

Universidade de Lisboa
Faculdade de Letras
Departamento de Estudos Anglísticos



Shell nouns
in a Systemic Functional Linguistics perspective

Ann Cecelia Henshall

Doutoramento no ramo de conhecimento de Linguística
Especialidade de Análise do Discurso

2015

Universidade de Lisboa
Faculdade de Letras
Departamento de Estudos Anglísticos



Shell nouns
in a Systemic Functional Linguistics perspective

Ann Cecelia Henshall

Tese orientada pelas Prof^{as} Doutoradas Emília Pedro e Merja de Mattos-Parreira especialmente elaborada para a obtenção do grau de doutor em Linguística, especialidade de Análise do Discurso

2015

Resumo

“Shell nouns” numa perspectiva da gramática sistémica funcional

Esta tese visa desenvolver uma explicação de “shell nouns” (Schmid, 2000) com base na teoria da gramática sistémica funcional (Halliday & Matthiessen, 1999; 2004; Martin, 1992, Martin & White, 2005). Considera-se a teoria apropriada para o estudo de “shell nouns” dado que a língua é vista como um sistema semiótico e estratificado, construindo simultaneamente o significado em três metafunções (a ideacional, a interpessoal e a textual) e em estratos diferentes (e.g. o da léxico-gramática e do discurso semântico), e sendo um texto uma instância do sistema linguístico. A teoria fornece as ferramentas adequadas para analisar “shell nouns” uma vez que já se demonstrou que estes últimos podem exercer funções avaliativas (e.g. Schmid, 2000; Charles, 2003; 2007), funções textuais que promovem a coesão do texto (e.g. Schmid, 2000; Francis, 1994; Tadros, 1994), e funções cognitivas (Schmid, 2000). Define-se o “shell noun” pela capacidade do nome de entrar em um ou dois padrões léxico-gramaticais: construir Valor numa oração cujo Processo é realizado pelo verbo ‘ser’ e cuja Característica é realizada por uma oração de facto; e/ou construir Cabeça num grupo nominal, projectando uma oração como Qualificador. Quando um desses nomes encapsula informação complexa realizada por uma oração ou mais, considera-se que está a funcionar como um “shell noun”. As proposições encapsuladas no “shell noun” denominam-se lexicalização.

Opta-se por uma abordagem de corpus paralelo. O corpus é composto por cinco artigos científicos que estão prontos a submeter a publicação, escritos em inglês por professores portugueses na área de economia, e cinco artigos científicos escritos em inglês e já publicados sobre matérias comparáveis. A análise pretende responder a duas perguntas gerais:

1. Qual a contribuição do “shell nouns” para o significado ideacional, interpessoal e textual?
2. Qual a diferença no uso dos “shell nouns” nos artigos científicos publicados e nos manuscritos de submissão escritos por professores portugueses na área de economia?

As duas perguntas permitem uma análise qualitativa e quantitativa, sendo feita a análise no estrato da léxicogramática e no estrato da semântica do discurso. Utilizando a versão 2.7.4 do *Corpus Tool* (O'Donnell, 2008), o corpus é anotado de acordo com sistemas linguísticos das metafunções nos dois estratos.

A conclusão principal do estudo é que os “shell nouns” são um recurso sistémico importante para um académico, que precisa de construir um argumento, posicionando-se e apresentando o estudo para convencer a comunidade do discurso de que o artigo faz um contributo ao conhecimento científico na sua área. Uma das maneiras em que os “shell nouns” contribuem a esse propósito é através da construção da posição dialógica do autor. São maioritariamente os “shell nouns” mentais, linguísticos e factuais que contribuem para esse efeito. Enquanto os “shell nouns” mentais e linguísticos, que muitas vezes são nominalizações de Processos mentais e de fala, normalmente expandem o espaço dialógico, os “shell nouns” factuais podem enquadrar uma proposição, fechando o espaço dialógico, muitas vezes construindo Valor numa oração de Processo relacional cuja Característica é realizada por uma oração de facto. Mais, a junção entre as seleções dos sistemas interpessoais e textuais permite que a informação da lexicalização seja avaliada de uma forma positiva ou negativa, assim aliando o leitor com uma visão em particular, e afastando o de outras. O elemento avaliativo pode ser isolado ou incorporado no “shell noun”. Enquanto outros estudos já notaram essas funções interpessoais (e.g. Francis, 1994; Charles, 2003; 2007; Kanté, 2010), a abordagem da gramática sistémico funcional explica como é construído o significado.

Uma segunda função dos “shell nouns” é textual. Funcionam para resumir informação realizada por uma oração ou mais e que se encontra noutra sítio do texto, podendo esta ser antes ou depois do “shell noun”. O acto de resumir é que permite o desenvolvimento do texto e contribui para a coesão do texto. O uso anafórico e catafórico torna-os um recurso importante na estrutura do texto, aparecendo nos macro-Temas, nos híper-Temas, nos híper-Novos, e na conjunção interna e externa no estrato da semântica do discurso. Nos híper-Temas e nos híper-Novos, é frequente que se junte com significado interpessoal. Num Adjunto de conjunção, um “shell noun” que

funciona como uma anáfora resumativa pode acrescentar significado interpessoal ou significado ideacional.

No estrato léxicogramatical, um “shell noun” com a função de anáfora resumativa tipicamente constrói Tema, sendo um ponto de partida partilhada da oração. Quando um “shell noun” com a função de anáfora resumativa constrói Característica numa oração relacional descodificadora cujo Processo é realizado por um verbo tal como ‘reveal’, ‘confirm’ ou ‘suggest’, funciona para construir explicitamente o argumento do escritor. É muito frequente encontrar nesse uso “shell nouns” que constroem o campo de investigação, tais como ‘resultados’.

Outra conclusão do estudo é que, ao contrário do que muitas vezes se assume, alguns “shell nouns” constroem um campo específico. É o caso dos “shell nouns” que pertencem ao campo de investigação: e.g. resultado, conclusão, análise. Mais ainda, embora a maioria dos “shell nouns” não construa um campo específico, as lexicalizações funcionam para tal. O “shell noun” pode projectar uma oração, criando entidades semióticas específicos a um campo, e instanciadas no discurso. Por exemplo, a junção de selecções do sistema de JULGAMENTO e uma figura do campo específico que constrói Qualificador ao “shell noun” permite que o investigador construa entidades que formam a base da investigação empírica porque podem ser testadas, quantificadas ou manipuladas. Assim, torna-se possível definir um valor para “a **probabilidade** de que *w* seja juntado a um parceiro pelo menos tão bom como *v*” ou medir “a **capacidade** de uma família de gerir as suas despesas gerais”¹.

Sendo um elemento do grupo nominal, o “shell noun” consegue entrar em estratos e funções diferentes da léxicogramática. Embora nesse estudo se verifique que constroem mais frequentemente Participantes nos Processos relacionais, também se verifica que constroem Meta e Circumstância com frequência. Sendo um elemento funcional da oração, podem beneficiar dos sistemas de estrutura Temática e de estrutura Informacional.

¹ “A probabilidade de que *w* seja juntado a um parceiro pelo menos tão bom como *v*” e “a capacidade de uma família de gerir as suas despesas gerais” são exemplos tirados do estudo.

A capacidade do “shell noun” de funcionar nas maneiras descritas deriva do seu estatuto de abstracção semiótica, que tem a capacidade de se referir ao texto como um facto. Essa capacidade, por seu turno, tem origem no seu estatuto de metáfora gramatical, sendo um “shell noun” uma nominalização de um Processo ou Qualidade de projecção, uma nominalização de modalidade, um nome de expansão, um nome de facto ou uma nominalização de um verbo de conação. Quer isto dizer, os “shell nouns” juntam-se como uma classe através do processo semogénico da nominalização, onde a metáfora gramatical neutraliza algumas das distinções na gramática entre as relações logico-semânticas de projecção e de expansão (Halliday & Matthiessen, 1999; 2004), mas também entre a modalidade e a projecção e entre a modalidade e a expansão. Pode-se representar os três sistemas – projecção, expansão e modalidade - por três círculos que se intersectam, as zonas do cruzamento sendo a zona da metáfora gramatical, e a zona central resultada do cruzamento dos três sistemas, o processo de nominalização, que cria um “shell noun”.

A tese visa também estabelecer se existem diferenças no uso dos “shell nouns” entre o corpus dos artigos científicos publicados e as submissões dos académicos portugueses. Embora se verifique uma maior frequência de “shell nouns” no corpus das submissões, os escritores dos dois corpus tiveram recurso a todas as funções dos “shell nouns”. Não obstante, destaca-se uma diferença: a posição dialógica construída pelos “shell nouns”. Os autores portugueses utilizam os “shell nouns” na maior parte das vezes para expandir o espaço dialógico, apresentando as suas afirmações e a sua voz como uma voz entre muitas outras. Em contraste, os autores publicados utilizam os “shell nouns” na maior parte para reduzir o espaço dialógico, reconhecendo outras vozes quando atribuem os argumentos aos outros investigadores. Outra diferença é que o corpus publicado revela mais avaliação negativa do que o corpus das submissões. Juntando estes resultados, pode afirmar-se que a reacção do leitor possa ser influenciada pelas escolhas, os textos publicados podendo parecer ter mais autoridade, e os das submissões podendo parecer menos assertivos e mais experimentais.

PALAVRAS-CHAVE: gramática sistémica funcional; “shell nouns”; discurso académico; Inglês como língua adicional; metáfora gramatical

Summary

Shell nouns in a Systemic Functional Linguistics perspective.

The aim of this thesis is to develop an account of shell nouns (Schmid, 2000) in a Systemic Functional Linguistics (SFL) perspective. Using a parallel corpus comprising five article submissions by Portuguese academics in the field of economics and five published articles on comparable topics, the ideational, interpersonal and textual functions of shell nouns are tagged at the strata of the lexicogrammar and discourse semantics using Corpus Tool version 2.7.4 (O'Donnell, 2008). The systems networks used to tag the corpus are grounded in SFL theory.

The analysis shows that shell nouns constitute an important systemic resource for the writers of research articles, who need to build an argument, positioning themselves and their study to convince the discourse community that their paper makes a contribution to knowledge in their disciplinary field. They enable a text to unfold by compacting information realised as a clause or more elsewhere in the text. Thus they can help scaffold a text through hyper-Themes, hyper-News and internal conjunction. At the stratum of the lexicogrammar, anaphorically referring nominal groups with a shell noun as Head often compose Theme, where they constitute a shared point of departure for the clause. In a decoding relational clause whose Process is realised by a verb such as *reveal*, *confirm*, or *suggest*, an anaphorically referring shell noun that construes Token helps to explicitly build the writer's argument. Shell nouns that construe the field of research, such as **results** and **findings** are common in this function. Mental, linguistic and factual shell nouns contribute to construing dialogic position, and coupling between interpersonal systems and textual systems enables the writer to align the reader with certain positions and disalign with others. Although most shell nouns are not field specific, because they can project a figure that instantiates an entity, they contribute to construing field, for example instantiating entities as the object of study of the empirical research.

The capacity of shell nouns to function as described above derives from their status as semiotic abstractions, which can refer to text as fact or report and are grammatical metaphors. They can be seen as lying at the intersection of modality and

the logico-semantic relations of projection and expansion, brought into being by the semogenic process of nominalisation.

The writers of the published articles and article submissions are found to use shell nouns in all of the functions above, but there are differences in the relative shares of the functions, which may affect reader reactions to the text.

KEYWORDS: Systemic Functional Linguistics; shell nouns; academic discourse; English as an additional language; grammatical metaphor

Table of contents

1.	Introduction	1
1.1.	A preview of shell nouns	1
1.2.	Motivation for and context of the study.....	2
1.2.1.	The local context	3
1.2.2.	Writing research articles in English for publication	4
1.3.	The study.....	5
1.3.1.	Data for the study and research questions.....	6
1.3.2.	Significance of the study	7
1.3.3.	Organisation of the thesis	7
2.	A broad framework for the study.....	9
2.1.	A Systemic Functional Linguistics framework.....	9
2.1.1.	A very brief overview of Systemic Functional Linguistics.....	9
2.2.	The stratum of the lexicogrammar	10
2.3.	The stratum of discourse semantics	12
2.3.1.	Ideation at the discourse semantics stratum.....	12
2.3.2.	Interpersonal meaning at the discourse semantics stratum	13
2.3.3.	Textual meaning at the discourse semantics stratum	18
2.4.	Instantiation	20
2.5.	Individuation and affiliation	21
2.6.	The research article as genre	23
2.7.	Register and scientific discourse	26
2.8.	Grammatical metaphor	27
2.9.	Summary of chapter 2.....	35
3.	Shell nouns	37
3.1.	A brief definition of shell nouns.....	37
3.1.1.	The functions of shell nouns	38
3.1.2.	Lexicogrammatical patterns of shell nouns	40
3.1.3.	Semantic classification of shell nouns.....	41
3.1.4.	Pragmatic uses of shell nouns	43
3.2.	Other research on shell nouns	45
3.2.1.	Early approaches to shell nouns.....	45
3.2.2.	More recent studies on shell nouns.....	52
3.2.3.	Shell nouns in Portuguese	61

3.3.	A Systemic Functional Linguistic Perspective.....	65
3.3.1.	Shell nouns as a class	65
3.4.	Shell nouns in the lexicogrammar.....	74
3.4.1.	Shell nouns, projection and expansion	75
3.4.2.	Shell nouns as Value in intensive indentifying relational Process clauses.....	80
3.4.3.	Shell nouns and reference.....	83
3.4.4.	Shell nouns and text reference	87
3.4.5.	Summary of shell nouns at the stratum of the lexicogrammar	88
3.5.	Shell nouns in the discourse semantics stratum.....	89
3.5.1.	Shell nouns and textual patterning	89
3.5.2.	Shell nouns and interpersonal meaning.....	96
3.5.3.	Shell nouns and ideational meaning	100
3.5.4.	Summary of shell nouns at the discourse semantics stratum	101
4.	Methodology.....	103
4.1	Research questions	103
4.2.	Selection of the corpus.....	104
4.3.	Corpus Tool	109
4.4.	Identifying and tagging the shell nouns	109
4.5.	Tagging the lexicogrammatical stratum.....	111
4.5.1.	Tagging ideation at the lexicogrammatical stratum	111
4.5.2.	Tagging the interpersonal metafunction in the lexicogrammar	123
4.5.3.	Tagging the textual metafunction in the lexicogrammar.....	125
4.6.	Tagging the corpus at the discourse semantics stratum.....	130
4.5.1.	Basic systems.....	131
4.5.2.	Tagging the ideational function in discourse semantics	132
4.5.3.	Tagging the textual function in discourse semantics	134
4.6.4.	Tagging the interpersonal function in discourse semantics.....	139
4.7.	Summary of Chapter 4	144
5.	Results and discussion for shell nouns at the stratum of the lexicogrammar	145
5.1.	A brief global comparison	145
5.2.	Shell nouns at the stratum of the lexicogrammar	147
5.2.1.	Basic functions of shell nouns at the stratum of the lexicogrammar	147
5.3.	Participants realised by shell nouns.....	149
5.3.1.	Shell noun as Participants in material Process clauses	150

5.3.2.	Shell nouns as Participants in verbal Process clauses	154
5.3.3.	Shell nouns as Participants in mental Process clauses.....	156
5.3.4.	Shell nouns as Participant in existential Process clauses	159
5.3.5.	Shell nouns as Participant in relational Process clauses	161
5.3.6.	Shell nouns as Medium	175
5.4.	Modification to Participants realised with shell nouns.....	177
5.4.1.	Modification realised by Post-Deictic	177
5.4.2.	Modification by Numerative	178
5.4.3.	Modification by Epithet.....	180
5.4.4.	Modification by Classifier	181
5.4.5.	Modification by Qualifier	182
5.5.	Circumstances realised with shell nouns	187
5.6.	Relators realised with shell nouns	193
5.7.	Textual Themes with shell nouns	193
5.8.	Summary and discussion	194
6.	Results and discussion of shell nouns at the discourse semantics stratum.....	201
6.1.	Shell nouns and textual meaning at the discourse semantics stratum	201
6.1.1.	Reference	201
6.1.2.	Macro-Themes, hyper-Themes and hyper-News.....	222
6.1.3.	Internal conjunction	227
6.2.	Shell nouns and experiential meaning at the discourse semantics stratum	230
6.2.1.	Shell nouns as a resource for construing a specific field.....	230
6.2.2.	Non-field-specific shell nouns	232
6.3.	Shell nouns and interpersonal meaning at the discourse semantics stratum	247
6.3.1.	Shell nouns and dialogic positioning	247
6.3.2.	Shell nouns as a resource for grading propositions	262
6.3.3.	Shell nouns as a resource for Attitude	265
6.3.4.	Modification of shell nouns by means of GRADUATION.....	281
6.4.	Summary of results and discussion	285
6.4.1.	The contribution of shell nouns to textual meaning.....	285
6.4.2.	The contribution of shell nouns to ideational meaning	289
6.4.3.	The contribution of shell nouns to interpersonal meaning	291
6.4.4.	Coupling among textual, ideational and interpersonal metafunctions	297
6.4.5.	Revisiting Schmid's three functions of shell nouns	300

7. Conclusions	303
Bibliography	309
Appendices	323
Annexes	391

List of figures

Figure 2.1	Ideational, interpersonal and textual metafunctions (Martin & White, 2005: 8)	9
Figure 2.2	Modelling of language in context adapted from Martin & White (2005) in Hood (2010: 23)	10
Figure 2.3	The system of Attitude (adapted from Martin & White, 2005) ..	14
Figure 2.4	Cline of interpersonal grammar and appraisal (adapted from Martin & White, 2005:55)	15
Figure 2.5	The system network of Graduation (Martin & White, 2005, extended by Hood, 2010)	16
Figure 2.6	The system network of Engagement (Martin & White, 2005, adapted)	17
Figure 2.7	Sandwich texture in abstract written discourse (Martin 1992:456)	19
Figure 2.8	The system network of Identification – primary delicacy (Martin, 1992: 120)	20
Figure 2.9	Cline of instantiation (Martin & White, 2005: 25)	20
Figure 2.10	Individuation and affiliation (Martin, 2010: 24)	22
Figure 2.11	The three semohistories related (Halliday & Matthiessen, 1999: 18)	26
Figure 2.12	Grammatical metaphor: sequence reconstrued as figure (adapted from Halliday, 1998: 190,191)	27
Figure 2.13	Grammatical metaphor: figure reconstrued as element (adapted from Halliday, 1998: 191)	27
Figure 2.14	Evolution of grammatical pattern in scientific English (Halliday, 1988/1993: 66)	29
Figure 2.15.	<i>Indicate</i> interpreted as identifying Relational Process and as Mental Process (adapted from Halliday, 1988/1993: 65)	31
Figure 2.16	The tension between literal and figurative readings of Gould’s nominalisation of Darwin’s theory (Halliday, 1993b: 31)	33
Figure 2.17	The tension between literal and figurative readings of Gould’s nominalisation of Darwin’s prediction	34
Figure 3.1	The converging balance of shell nouns (Schmid, 2000: 19)	37
Figure 3.2	A general system for nouns (Halliday & Matthiessen, 1999: 190)	66
Figure 3.3	Organising vocabulary (Ravelli, 2004: 117)	68
Figure 3.4	Congruent and metaphorical wordings of N- <i>be</i> -cl (Halliday & Matthiessen, 2004: 639)	70

Figure 3.5	Congruent and metaphorical wordings of <i>N-cl</i>	71
Figure 3.6	Shell noun as Value / Identified in an encoding identifying clause	81
Figure 3.7	Shell noun as Token / Identified in a decoding identifying clause	83
Figure 3.8	The system of DETERMINATION (Halliday & Mathiessen, 2004 :313).....	85
Figure 3.9	Dual facing of semiotic abstraction/shell noun in hyper-Theme (adapted from Ravelli, 2004: 120)	90
Figure 3.10	Reconstrual of clause as Qualifier to shell noun	92
Figure 3.11	Shell nouns and internal conjunction (adapted from Martin, 1992: 416-417)	93
Figure 4.1	Primary systems for analysis at the stratum of the lexico-grammar	111
Figure 4.2	Basic systems for the analysis of the ideational metafunction at the lexico-grammar stratum	112
Figure 4.3	Shell noun in Qualifier realised as prepositional phrase (P_CEO)	113
Figure 4.4	System network for experiential function at group level	114
Figure 4.5	Primary systems for tagging experiential elements of the clause.	114
Figure 4.6	System network for PARTICIPANT	115
Figure 4.7	System network for CIRCUMSTANCE	115
Figure 4.8	Transitivity analysis showing tagging of elements in an embedded clause containing a shell noun (P_Econ)	116
Figure 4.9	Transitivity analysis showing tagging of elements in a hypotactic bound clause containing a shell noun (P_Econ)	116
Figure 4.10	Transitivity analysis showing tagging of elements in a Qualifier containing a shell noun (S_Rand)	116
Figure 4.11	Transitivity analysis showing tagging of elements in a conjunction (P_Pov)	117
Figure 4.12	Transitivity analysis showing tagging of elements in a conjunctive Adjunct (P_Rand)	117
Figure 4.13	Identifying Intensive clause with agency	120
Figure 4.14	Verbs realising intensive identifying relational processes graded by degree of commitment to proof	120
Figure 4.15	System network for tagging PROCESS	122
Figure 4.16	Transitivity analysis showing unmarked identifying intensive relational Process clause (P_Pov)	122

Figure 4.17	Transitivity analysis showing marked identifying intensive Process relational clause (P_Net)	123
Figure 4.18	System for tagging the interpersonal metafunction in the lexico-grammar	123
Figure 4.19	Interpersonal analysis showing shell noun as Subject (P_Pov)	124
Figure 4.20	Interpersonal analysis showing shell noun as Complement (P_CEO)	124
Figure 4.21	Interpersonal analysis showing shell noun as Adjunct (P_CEO) ...	124
Figure 4.22	Basic systems for the analysis of the textual metafunction at the stratum of the lexico-grammar.	125
Figure 4.23	Textual analysis showing elements tagged unmarked Theme	126
Figure 4.24	Textual analysis showing elements tagged marked Theme	126
Figure 4.25	Textual analysis showing tagging of shell noun in Theme and in Qualifier (S_Rand)	126
Figure 4.26	System network for THEME SELECTION	127
Figure 4.27	Textual analysis showing clausal Theme (S_Rand)	127
Figure 4.28	Textual analysis showing Shell noun in Conjunction in clausal Theme in clause nexus (S_Pov)	127
Figure 4.29	Textual analysis showing Non-finite clausal Theme in clause nexus (S_Pov)	128
Figure 4.30	System network for tagging N-RHEME	128
Figure 4.31	Textual analysis showing elements tagged N- RHEME in clause simplex (S_Rand)	129
Figure 4.32	Textual analysis showing elements tagged N- RHEME, clause complex internal (S_Rand)	129
Figure 4.33	Textual analysis showing elements tagged (other)	129
Figure 4.34	System for tagging the clausal environment in which a shell noun appears	129
Figure 4.35	Analysis showing tagging of dominant clause in hypotactic sub-complex nested in paratactic clause nexus (P_CEO)	130
Figure 4.36	Primary systems for tagging at the discourse semantics stratum.	131
Figure 4.37	Systems network for tagging the lexicalisation of a shell noun ...	131
Figure 4.38	Tagging the elements in the system SEGMENT	132
Figure 4.39	Systems network for tagging ideational function in discourse semantics	133
Figure 4.40	Analysis showing tagging of elements construing external conjunction	133

Figure 4.41	Analysis showing tagging of elements in SEGMENT and IDEATION	134
Figure 4.42	Analysis showing tagging of elements at and below clause rank in IDEATION	134
Figure 4.43	Primary systems network for the textual function	134
Figure 4.44	Systems network for macro-Themes and hyper-Themes/News	135
Figure 4.45	Analysis showing tagging of macro-Theme	135
Figure 4.46	Analysis showing tagging of hyper-Theme	135
Figure 4.47	Analysis showing tagging of hyper-New	135
Figure 4.48	Analysis showing tagging of shell noun in INTERNAL CONJUNCTION ..	135
Figure 4.49	Systems network for analysis of IDENTIFICATION	136
Figure 4.50	Systems network for tagging the relation of the shell noun to its lexicalisation	137
Figure 4.51	Analysis showing tagging of element containing shell noun in IDENTIFICATION	138
Figure 4.52	Analysis showing tagging of element containing esphorically referring shell noun in IDENTIFICATION	139
Figure 4.53	Primary systems network for the interpersonal function in discourse semantics	139
Figure 4.54	Analysis showing tagging of element with shell noun in ENGAGEMENT	141
Figure 4.55	Analysis showing tagging of shell noun as APPRECIATION	142
Figure 4.56	Analysis showing tagging of shell noun as AFFECT	142
Figure 4.57	Analysis showing tagging of element with shell noun as invoked ATTITUDE	142
Figure 4.58	Analysis showing tagging of element with shell noun as downscaled actualisation	143
Figure 4.59	Analysis showing tagging of element with shell noun as downscaled completion	143
Figure 4.60	Analysis showing tagging of element within nominal group with shell noun as Head as GRADUATION: FOCUS	143
Figure 4.61	Analysis showing tagging of element within nominal group with shell noun as Head as GRADUATION	143
Figure 7.1	The class of shell nouns in a Systemic Functional Linguistics perspective	306

List of tables

Table 2.1	General options for external and internal conjunction (Martin and Rose, 2003)	13
Table 2.2	Cline of instantiation – evaluation (Martin & White, 2005: 164).....	21
Table 2.3	Stages of the research article (Martin and Rose, 2008:207)	24
Table 2.4	Generic structure of the research article introduction	24
Table 2.5	Domains of elemental metaphors (Halliday & Matthiessen, 1999: 245).	28
Table 3.1	Shell noun patterns (Schmid, 2000)	41
Table 3.2	Semantic relations between classes of shell nouns (adapted from Schmid, 2000:300, examples added)	42
Table 3.3	Semantic features functioning as classifiers (Schmid, 2000:88, examples added)	44
Table 3.4	Alternative names for shell nouns	51
Table 3.5	Grammatical metaphor: nominalisation (adapted from Halliday & Matthiessen, 1999: 245)	69
Table 3.6	Kinds of entities (Martin & Rose, 2003:108)	73
Table 3.7	Nouns of projection and nouns of fact (adapted from Halliday and Matthiessen, 2004: 469)	78
Table 3.8	Evaluation nouns appearing in <i>it is</i> Attribute: <i>noun that</i> (adapted from Halliday & Matthiessen (2004: 475), added nouns in italics)	79
Table 3.9	Nouns of expansion (Halliday & Matthiessen, 2004:435)	80
Table 3.10	Other nouns of expansion that function as shell nouns	80
Table 4.1	Parameters for defining a corpus as specialised (Flowerdew, 2004: 21)	105
Table 4.2	Parameters of the corpus for the study of shell nouns	106
Table 4.3	Summary of characteristics of the two sub-corpora	108
Table 4.4	Shell noun patterns (Schmid, 2000)	109
Table 4.5	Verbs serving as intensive identifying relational Processes, adapted from Halliday and Matthiessen (2004) and Martin (1992)	118
Table 4.6	Cline of caused modality in relational Processes and fact nouns of the ‘evidence’ type	121
Table 4.7	Elements tagged NO INTERPERSONAL FUNCTION	124
Table 5.1	Frequency of shell nouns in Published and Submissions	146
Table 5.2	Distribution of shell noun types in Published and Submissions	146

Table 5.3	Ideational, interpersonal and textual functions of shell nouns at clause rank	148
Table 5.4	Distribution of Process types in clauses in which shell noun construes a Participant	150
Table 5.5	Distribution of shell noun types construing Actor in Published and Submissions	152
Table 5.6	Distribution of shell noun types construing Goal in Published and Submissions	153
Table 5.7	Distribution of shell noun types construing Scope in Published and Submissions	154
Table 5.8	Distribution of shell noun types construing Verbiage in Published and Submissions	156
Table 5.9	Distribution of shell noun types construing Phenomenon in Published and Submissions	159
Table 5.10	Distribution of shell noun types construing Existent in Published and Submissions	160
Table 5.11	Distribution of shell noun types construing Token in Published and Submissions	164
Table 5.12	Distribution of shell noun types construing Value in Published and Submissions	170
Table 5.13	Distribution of shell noun types construing Carrier in Published and Submissions	172
Table 5.14	Distribution of shell noun types construing Attribute in Published and Submissions	174
Table 5.15	Distribution of shell noun types construing Medium in Published and Submissions	176
Table 5.16	Distribution of shell noun types among Participants realised with Post-Deictic in Published and Submissions	178
Table 5.17	Distribution of shell noun types among Participants realised with Numerative in Published and Submissions	179
Table 5.18	Distribution of shell noun types among Participants realised with Epithet in Published and Submissions	180
Table 5.19	Distribution of shell noun types among Participants realised with Classifier in Published and Submissions	182
Table 5.20	Distribution of realisation types of projected Qualifiers for shell nouns construing Participants	183
Table 5.21	Modification in cataphorically referring or lexicalised shell nouns in Circumstances composing marked Theme	189
Table 5.22	Modification to shell nouns in Circumstances composing N-Rheme ..	190

Table 5.23	Shell noun types in Circumstantial element at clause rank	192
Table 5.24	Textual Themes realised by conjunctive Adjuncts containing shell nouns	194
Table 6.1	Distribution of direction of LEXICALISATION of shell nouns in Published and Submissions	203
Table 6.2	Anaphorically lexicalised shell nouns as Attribute	206
Table 6.3	Textual reference to a superset with anaphorically lexicalised shell nouns in Published and Submissions	208
Table 6.4	Anaphorically referring shell noun by shell noun type in Published and Submissions	211
Table 6.5	Cataphorically lexicalised shell nouns that couple with resources from APPRAISAL: FORCE in Published and Submissions	216
Table 6.6	Cataphorically lexicalised shell nouns coupling with resources from GRADUATION: FOCUS	217
Table 6.7	Cataphorically lexicalised shell nouns coupling with ATTITUDE	217
Table 6.8	Cataphorically lexicalised shell nouns infused with prosodic value in Published and Submissions	218
Table 6.9	Cataphorically lexicalised shell nouns with reference to a superset ...	220
Table 6.10	Distribution of shell noun types in cataphorically lexicalised shell nouns in major clauses in Published and Submissions	221
Table 6.11	Distribution of shell nouns in anaphoric and cataphoric functions in major clauses across whole corpus	221
Table 6.12	Shell nouns in headings composing macro-Themes in Published and Submissions	223
Table 6.13	Esphorically referring shell nouns in hyper-Themes in Published and Submissions	225
Table 6.14	Cataphorically lexicalised shell nouns in hyper-Themes in Published and Submissions	225
Table 6.15	Shell nouns as a resource for internal conjunction in Published and Submissions	228
Table 6.16	Anaphorically referring shell nouns combined with internal conjunction in Published and Submissions	230
Table 6.17	Distribution of shell noun types as contribution to field in whole corpus	230
Table 6.18	Distribution of field construed through shell nouns in Published and Submissions	231
Table 6.19	Shell nouns that construe field as the object of study in Published and Submissions	232

Table 6.20	Realisation type of lexicalisations that construe field as the object of study in Published and Submissions	239
Table 6.21	Lexicalisations that construe field as methodology by realisation type in Published and Submissions	240
Table 6.22	Shell nouns whose lexicalisations construe field as methodology	240
Table 6.23	Shell nouns whose lexicalisation construes the field of research projecting the field of the object of research in Published and Submissions	244
Table 6.24	Field-specific modification to shell nouns in Published and Submissions	246
Table 6.25	Shell nouns that expand discourse by attribution in Published and Submissions	250
Table 6.26	Shell nouns endorsing other writers' propositions in dialogically contracting discourse in Published and Submissions	251
Table 6.27	Esphorically referring shell nouns (lexicalisation realised as fact clause) that function to contract	253
Table 6.28	Shell noun matrices in expanding discourse: entertain	257
Table 6.20	Shell nouns in headings dialogically expanding discourse in Published and Submissions	260
Table 6.30	Distribution of shell nouns functioning to contract or expand the discourse in whole corpus	260
Table 6.31	Distribution of dialogic position by means of shell nouns in Published and Submissions	261
Table 6.32	Shell noun types enacting CONTRACT in Published and Submissions	262
Table 6.33	Shell noun types enacting EXPANDING discourse in Published and Submissions	262
Table 6.34	Instances of shell nouns that grade FULFILMENT of a proposition in Published and Submissions	264
Table 6.35	Distribution of shell nouns contributing to GRADUATION: FULFILMENT in Published and Submissions	265
Table 6.36	Attitudinally inscribed shell nouns enacting APPRECIATION in Published and Submissions	267
Table 6.37	Inscribed ATTITUDE in shell nouns realised by Epithets in Published and Submissions	270
Table 6.38	Distribution of shell nouns to enact APPRECIATION, inscribed and infused, in Published and Submissions	272
Table 6.39	Shell nouns enacting JUDGEMENT: SOCIAL ESTEEM: capacity in Published .	274
Table 6.40	Shell nouns enacting JUDGEMENT: SOCIAL ESTEEM: capacity in Submissions	275

Table 6.41	Shell nouns enacting JUDGEMENT: SOCIAL SANCTION: propriety in Published and Submissions	276
Table 6.42	Shell nouns enacting JUDGEMENT: SOCIAL SANCTION: veracity in Published and Submissions	278
Table 6.43	Distribution of shell nouns enacting JUDGEMENT in Published and Submissions	278
Table 6.44	Shell nouns as AFFECT in Published and Submissions	280
Table 6.45	Distribution of shell nouns in APPRECIATION and JUDGEMENT in whole corpus	281

Chapter 1. Introduction

1.1. A preview of shell nouns

In the Andrew V. McLaglen western, *Chisum*, (1970), the following dialogue between the hero, Chisum, played by John Wayne, and the baddie, Murphy, played by Forrest Tucker, takes place at a point where Chisum is about to leave town for a few days, leaving his ranch and the store unprotected.

Chisum: And if one of your men cross my land or even touch one of my cows, or do anything to that store, I'm not going to the sheriff, the governor or the president of the United States. I'm coming to see you.

Murphy: Mr Chisum, that sounds like a threat. [Chisum punches Murphy and knocks him down]

Chisum: Wrong word. Fact.

In this dialogue both 'threat' and 'fact' can be seen as referring to Chisum's first utterance that he would be going to see Murphy if any of his men encroached on Chisum's land or the store. While Murphy interprets the utterance as a threat, Chisum refutes the interpretation and recharacterises it as a fact; there is no doubt about the outcome. A little later in the film, another dialogue – this time between the Sheriff and Murphy – also involves interpretation of an utterance, but unlike the first excerpt where the interpretation follows the utterance, the interpretation precedes it:

Murphy: You just leave him [Chisum] alone. I'll take care of Mr Chisum.

Sheriff Brady: That ain't gonna be easy.

Murphy: You're right. But see, there's a fundamental difference between Mr Chisum and me.

Sheriff Brady: Yeah, what's that?

Murphy: Mr Chisum is a man who respects the law. [Murphy chuckles.] Around here, I'm the man who owns it.

Here, Murphy explains how, or perhaps why, he will be able to deal with Chisum; differences in their characters determine how they will behave. The difference between Chisum and Murphy is that Chisum respects the law but Murphy does not need to because he owns it.

Nouns like 'fact', 'threat' and 'difference' are the subject of this thesis. Other examples of nouns that function in similar ways include 'idea', 'evidence' or 'possibility' to name but a few. These nouns have been called 'shell nouns' because

they encapsulate another proposition and function as a ‘conceptual shell’ (Schmid, 2000: 6). In the excerpts above ‘threat’ and ‘fact’ encapsulate the idea that Chisum will come and see Murphy if anything happens to the store or his cows, while ‘difference’ encapsulates the idea that Chisum respects the law but Murphy owns it. In each case, the specific meaning of the noun – fact, threat, difference – depends on recovering meaning from the co-text. This feature is one of the reasons why shell nouns are interesting. Unlike technical vocabulary, shell nouns depend on the context in which they appear in order to be understood. They contribute to the cohesion of the text and offer insight into the way the writer builds the discourse.

1.2. Motivation for and context of the study

The motivation for this study of shell nouns stems from the context in which I work. As in many other universities that are undergoing a process of internationalisation (Brock-Utne; Gazzola, 2010; Ferguson, Perez-Llantada, & Plo, 2011; Curry and Lillis 2004:680), academics at Lisbon School of Economics and Management (ISEG) are required to carry out research and publish their results. Publishing research articles in high-standing journals in the field of economics is tremendously competitive (Klamer and Hendrik, 2002; Labarde, Tollison & Karahn, 2002), and, as in other academic fields, it usually involves a time-consuming process of negotiating the content and language of the text with various literacy brokers (Lillis & Curry, 2010), including peer reviewers, the editor, and, in some cases, copy editors (McAfee, 2010; Berkenkotter & Huckin, 1995; van Bergeijk et al. 1997; Curry & Lillis, 2004; Barbin, 2008; Flowerdew, 2000; Lillis & Curry, 2006). Even then, there is no guarantee of being published (Barbin, 2008). In the field of economics it has been found that the probability of being published depends on the perceived relevance of the topic of the study (Klamer & Hendrik, 2002; Barbin, 2008; McAfee, 2010), the methodological rigour of the study (Barbin, 2008; Stewart, 2008), and writing or language-related concerns, (Klamer & Hendrik, 2002; van Bergeijk et al. 1997; Barbin, 2008; Svensson & Helgerson, 2008). These critical features are evident in the information for reviewers in the publication process for Elsevier, a major editor of scientific articles in economics (Henshall, 2012). An initial screening can reject submissions on the basis of incorrect references or basic science. Accepted manuscripts are then reviewed by peers, and if

approved, they pass to a reviewer for English grammar and language. After authorial revision, the manuscript is reviewed again by the peers and language reviewer, and it may still be rejected at this point. The full process is detailed in Annex A. In addition to the hurdles inherent in the publication process, it has been noted that the United States of America is dominant in setting standards for the field of economics (Coupé, 2003; Frey & Eichenberger, 1997; Diamond, 1989).

1.2.1. The local context

At ISEG publishing research articles is linked to the performance evaluation of academics. The number of publications, whether they are co-authored or not and the standing of the journal in which they are published are some of the performance evaluation criteria. There is a list of journals, FIISEG, which are ranked and weighted for their contribution to the performance evaluation. Moreover, a prize is awarded annually to the academic who publishes the greatest number of research articles in the highest ranking journals. The FIISEG list is not without its critics. In particular it has been criticized for a bias towards the field of economics, and for not being representative of all the areas of research undertaken by ISEG academics. In fact, there is currently a move to restructure the list, but it is only in its early stages.

The FIISEG list in operation during the course of this study was the object of analysis in Henshall (2012). I found that the overwhelming majority of the journals publish in English only. This bias towards English in scientific publication has been well documented in many other fields (Barbin, 2008; Diamond, 1989; Swales, 1990; Hamel, 2007), even to the extent that English has been called 'Tyrannosaurus rex' (Swales, 1997). In addition, many of the economic journals – often those based in the United States – operate on unstated assumptions about the English language variety to be used while other journals hold normative language policies, where they accept British English norms or American English norms, but not a mixture. In many journals there is advice for non-native speakers to have their research articles edited by a native speaker to enhance the probability of acceptance (Henshall, 2012). Such a service is usually paid for and represents an added expense for the Portuguese academic (Barros, 2014), which, in the current economic downturn, may not always be possible to secure. Dueñas (2012), for example, has documented that Spanish academics do not

have their papers edited for language prior to submitting them to publication for reasons of expense. Advice to revise the English of a manuscript before submission implies potential discrimination against non-standard English in research articles, which makes the task of the English-as an additional language (EAL) academic more difficult.

1.2.2 Writing research articles in English for publication

The plight of scholarly writers whose first language is not English as they struggle to get their research results published has been highlighted in a lively debate between Flowerdew (2008; 2009a) and Casanave (2008). Despite their different views, both acknowledge that there may be real and perceived discrimination against EAL scholars partly as a result of their non-standard use of English, and certainly the feeling that having to write in English represents an additional burden for the EAL scholar has been reported in case studies of academics in different fields and countries (e.g. Curry & Lillis, 2004; Flowerdew, 2000; Lillis & Curry, 2006; 2010; Pérez-Llantada, Plo & Ferguson, 2011) as well as in a recent survey in Spain (Moreno et. al., 2012). Further support for this view comes from bibliometric studies that identify the onus of writing in English as a barrier to publication for some EAL countries (Barbin, 2008; Svensson, & Helgerson, 2008). Barbin, for example, posits that it may explain the US advantage in publishing in top-ranking journals in the field of marketing. Notwithstanding, it has been noted that academic literacy is not merely a question of English language competence for EAL scholars; it must be acquired by native speakers and EAL speakers alike (Ferguson, Perez-Llantada & Plo, 2011). Moreover, editors may be supportive of EAL scholars' manuscript submissions (Lillis & Curry, 2010; Flowerdew, 2001).

Southern Europeans – Spanish and Portuguese – writing in the social sciences have in particular come under attack for language related issues. They have been accused of “severe wordiness” (Barbin, 2008: 379), writing “arguments [that] are vague and confusing” (Dueñas, 2012:148) or showing a “lack of elegance” (Parada, 2009), writing in a style that is “a little bit over the top and too pretentious” and “too Latin for a North-West European” (Lillis & Curry, 2010: 150), and using “weasel words (‘phenomena’, ‘approach’, ‘consideration’) [in a way that] is useless to the reader” (Lillis & Curry, 2010: 150). The last reviewer in question goes on to state that “Their

[the weasel words'] sole utility is to fill a void in the author's mind". Bennet (2012) offers a potential explanation for such criticisms. She argues that the Portuguese and English academic writing norms derive from different epistemological paradigms. The English style can be traced back to the ideology that emerged with the Scientific Revolution in the 17th century while the Portuguese retains the characteristics of the humanist traditions that came from Cicero and the classics, receiving further input with influences from the French poststructuralists in the early 20th century. She likens the English style to a 'window pane' in so far as it suggests clarity and transparency, unlike the aforementioned wordy, confusing style attributed to Southern European academics.

It is these 'weasel words' that are the object of this thesis. Like the nouns from the excerpt from Chisum, fact, threat and difference, the abstract nouns—phenomenon, consideration and approach can function as shell nouns. Perhaps somewhat perversely, words that function in the same way as these 'weasel words' are often encouraged in advice for students (Gray, 2010) as well as in submissions guidelines for scientific journals. For instance, the following advice is included in a list of requirements which was elaborated by the editor for publishing in the *Journal of Business Research*.

Avoid writing, "This results in three conclusions." Add a noun after writing, "This." For example, "This finding results in three conclusions."

The requirement to "Add a noun after writing, "This"" will often result in the use of a shell noun. In the advice given, **finding** would refer to a preceding clause, and in doing so, it is functioning as a shell noun.

1.3. The study

The purpose of this study is to explore how shell nouns contribute to meaning in academic discourse. While shell nouns have been comprehensively identified and classified, and their functions described in Schmid (2000), his work draws on different linguistic theories, and he does not attempt to account for their use within a single theoretical framework. This thesis uses Systemic Functional Linguistic (SFL) to do so. SFL arises as an appropriate theory to examine shell nouns because it offers a functional view of discourse in which meaning is explored through a trinocular

perspective: the ideational metafunction explores the ways in which people, entities and processes are construed; the interpersonal explores the ways in which social relations are enacted; and the textual explores how the ideational and interpersonal functions are organised to create text. In addition, the theory is multi-level: it provides systematic ways of looking at language at the level of the clause – the stratum of the lexicogrammar – as well as beyond the clause – the stratum of discourse semantics. The latter is particularly important as many shell nouns, like those in the excerpts from Chisum, operate across sentence boundaries.

1.3.1. Data for the study and research questions

The study uses a parallel corpus of research articles from the field of economics. The corpus is made up of un-edited article submissions written by Portuguese speaking scholars and published articles on comparable topics. The motivation behind the use of un-edited article submissions is that it is not always possible to have an article proof-read before submitting it to a journal (Parada,2009), so the use of un-edited papers reflects the context in which many Portuguese scholars write. Moreover, unlike many studies of shell nouns (e.g. Aktas & Cortes, 2008; Hasselgård, 2012), the writers are not novice to writing for academia. Published articles have been taken as the reference for comparative purposes, given that Portuguese scholars are writing for an international audience, not only native speakers, and the peer reviewers are similarly international, not necessarily from an Anglophone background.

The analysis of shell nouns in the corpus is approached using two broad research questions:

1. In what ways do shell nouns contribute to ideational, interpersonal and textual meaning?
2. Are there significant differences in their use between published articles and article submissions by Portuguese academics in the field of economics?

The first research question is further broken down to elucidate the functions of shell nouns at the stratum of the lexicogrammar and that of discourse semantics.

- i. What textual, ideational and interpersonal functions do shell nouns carry out in the lexicogrammar?
- ii. What textual, ideational and interpersonal functions do shell nouns carry out in the discourse semantics?

To carry out the study, CorpusTool 2.7.4 (O'Donnell, 2008) is used to tag the corpus. The software incorporates tools for carrying out statistical analyses so that it is possible to make comparisons across texts, as well as compare the statistical likelihood that items are used in specified data segments. As a result, the study is both qualitative and quantitative.

1.3.2. Significance of the study

The study makes a contribution to knowledge by a) using an SFL framework to define shell nouns; b) examining their ideational, interpersonal and textual functions at the strata of the lexicogrammar and discourse semantics; c) using a corpus of established academics rather than novice or apprentice writers; and d) using a parallel corpus of published articles and article submissions by Portuguese academics. The latter contribution is important because in today's university system, academics' career prospects are dependent on publishing in internationally peer-reviewed journals and lack of mastery of a field's preferred discourse may hinder access to publication. Greater knowledge about the use of shell nouns by Portuguese academic writers can then be used to inform academics writing for scientific publication.

1.3.3. Organisation of the thesis

The thesis is organised in six more chapters. Chapter 2 provides an overview of Systemic Functional Linguistics, which constitutes a framework for the analysis. The description is approached via the stratum of the lexicogrammar and discourse semantics. Special attention is given to the notion of grammatical metaphor, given its relevance to the study of shell nouns. Chapter 3 reviews the literature on shell nouns. After summarising Schmid's (2000) treatment of shell nouns, I describe and compare other studies' findings. I also include a brief description of how the concept of shell nouns is applicable to Portuguese. The final section of the chapter addresses how shell nouns fit into a Systemic Functional Linguistic framework. Chapter 4 explains the

methodology used for the analysis of shell nouns in the corpus. It details the characteristics of the corpus, and describes and exemplifies the metafunctional systems used to tag the corpus at the strata of the lexicogrammar and discourse semantics. Chapters 5 and 6 present and discuss the results of the study. Chapter 5 focuses on the analysis at the stratum of the lexicogrammar and Chapter 6 on that of discourse semantics. The results include both qualitative and quantitative analyses of the functions of shell nouns. At the end of each chapter there is a summary of the results and a discussion of how they relate to those of other studies. The conclusions are presented in Chapter 7.

Chapter 2. A broad framework for the study

2.1. A Systemic Functional Linguistics framework

This thesis is concerned with a functional class of nouns, shell nouns, and it studies them from a Systemic Functional Linguistics perspective. The aim of this chapter is to provide an initial theoretical framework that will underpin the study of the nouns. I shall start by overviewing the tenets of systemic functional linguistics, then look more closely at academic discourse, in particular the research article, and finish with a note on the relevance of grammatical metaphor as a resource in academic writing.

2.1.1. A very brief overview of Systemic Functional Linguistics

Systemic Functional Linguistics (SFL) considers language to be a social semiotic system; that is to say: “A language is resource for making meaning, and meaning resides in systemic patterns of choice” (Halliday & Matthiessen, 2004: 23). Through language we simultaneously construe experience and enact social relationships. These two functions are theorised in SFL as the ideational and interpersonal metafunctions respectively. A third metafunction – textual – enables the ideational and interpersonal metafunctions to be organised as text. Thus any text can be analysed from the perspective of each of the metafunctions, offering a trinocular perspective (figure 2.1), and the systemic choices are contrasted with other options from the system to explore meaning.

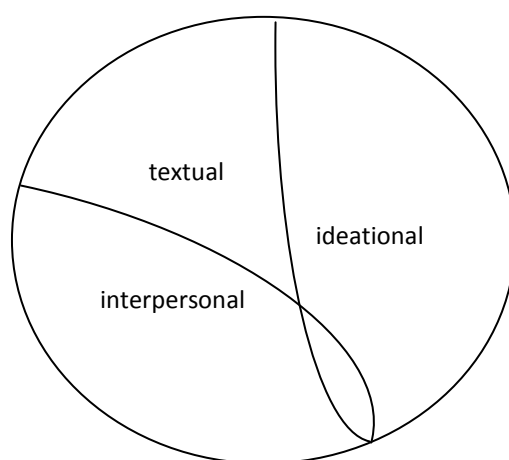


Figure 2.1 Ideational, interpersonal and textual metafunctions (Martin & White, 2005: 8)

In addition to the three kinds of meaning of the metafunctions, SFL adopts a stratified view of content in which all strata realise meaning (figure 2.2). That is to say, each stratum encodes meaning at a different level of abstraction (Martin & White, 2005), and realisation is “the process of linking one level of organization with another” (Halliday & Matthiessen, 2004: 26). Although at first sight this model may suggest that there are clear boundaries between the different strata, there are not. Rather, as Halliday and Matthiessen (1999: 32) put it, “the ideation base is construed as a multidimensional, elastic semantic space”, organised as meaning potential. Of particular interest to this thesis are the stratum of the lexicogrammar, which focuses on meaning at the level of the clause, and that of discourse semantics, which explores meaning choices across longer stretches of discourse, i.e. beyond the clause. Beyond language lies the context, which encompasses register and genre. A brief overview of the systemic approaches to the strata of the lexicogrammar and discourse semantics follows.

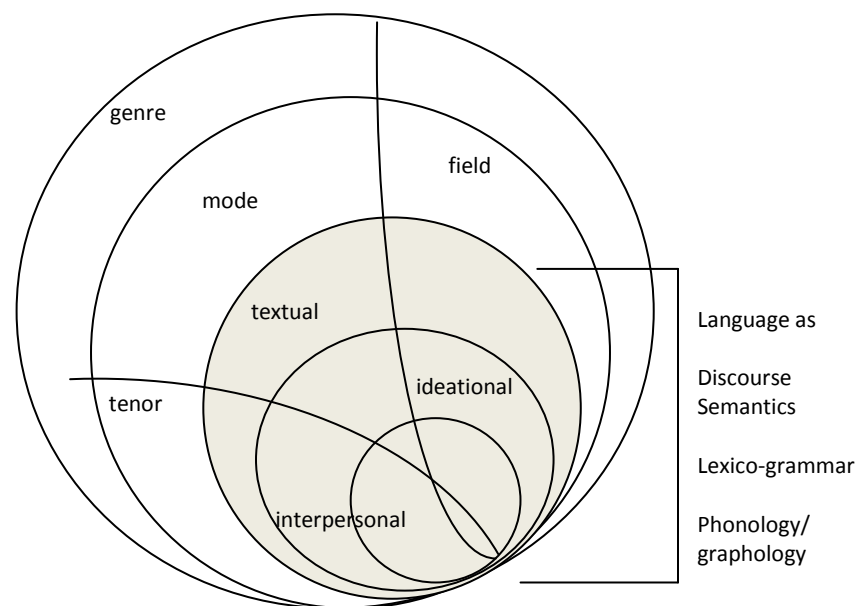


Figure 2.2 Modelling of language in context adapted from Martin & White (2005) in Hood (2010: 23)

2.2. The stratum of the lexicogrammar

The lexicogrammar focuses on meaning at the level of the clause. At this stratum, ideational meaning is construed through the system of transitivity, which identifies Processes (the kind of happening – that of doing, being or having, existing,

sensing, or saying), Participants (who/what is involved in the process) and any attendant Circumstances (e.g. when, where or how) (Halliday & Matthiessen, 2004).

The interpersonal metafunction is concerned with the semantic space between the speaker and listener that is opened up through language (Halliday & Matthiessen, 2004). The relevant system here is that of Mood. The grammar of the Mood includes the Subject and the Finite. The Subject is considered the one responsible for the validity of the proposition, and it can be identified by means of a tag question. The Finite indicates whether it is information (propositions) or goods-&-services (proposals) that are being exchanged. It also indicates reference with regard to the time of speaking and the speaker's assessment of the likelihood or desirability of the proposition.

The textual metafunction is composed by means of the systems of Thematic structure and Information structure. Together these systems distribute the flow of information in the clause. Thematic structure is composed of the Theme and Rheme, where the Theme is taken to be the first element with a role in the transitivity structure (Halliday & Matthiessen, 2004; Thompson & Thompson, 2009). In some analyses Theme is considered to be up to and including the Subject (e.g. Cummings, 2005; Hood, 2010), as proposed by Ravelli (1995: 226). The Theme represents the starting point or departure for the clause (Halliday & Matthiessen, 2004), and in its unmarked status, it conflates with the Subject. A marked Theme might consist of an Adjunct, for example. Thematic structure also allows for multiple Themes where there can be a textual Theme (e.g. *and*), an interpersonal Theme (e.g. *surely*) and a topical Theme (the first element with an experiential function). While the function of Theme is to provide the point of departure for the message, the function of Information structure is to organise the information in terms of Given (already known by the listener) and New (what the listener is to pay attention to). Typically Given information conflates with Theme, and New is presented in the Rheme.

The selection of Theme is important in a text because it contributes to the method of development of the text (Fries, 1981 in Halliday & Matthiessen, 2004: 99), which Fries (1995, a; b) stresses is a semantic notion, not structural. For example,

selecting a Theme that is not subject – i.e. a marked theme – can serve to re-orient the development of the text by creating a local context for what follows in the Rheme (Halliday & Matthiessen, 2004). Notwithstanding, Fries (2009) warns against including too much in the concept of method of development at the risk of it becoming meaningless.

2.3. The stratum of discourse semantics

The stratum of discourse semantics looks at meaning-making resources beyond the level of the clause. Stratification of language makes it possible to deal with units larger than the clause such as semantic motifs like evaluation, which are realised by different kinds of resources of the lexicogrammar, as well as the layering of meaning introduced by means of grammatical metaphor and aspects of coherence and cohesion (Martin, 1992). Following the early work on cohesion (Halliday & Hasan, 1976), a significant body of work that lays the foundations for analyses at the discourse semantics stratum has been carried out by Martin (1992), Martin and Rose (2003), and Martin and White (2005). These works will be the main sources for this section. The main systems that are useful for the analysis of shell nouns will be examined from the perspectives of the three metafunctions.

2.3.1. Ideation at the discourse semantics stratum

Ideational meanings are characterised by a particulate structure (Martin & Rose, 2003). Martin (1995b: 13) describes the structure of ideational meaning as involving “a segmental construal of reality”. Experiential meaning is structured along a principle of part/whole relations while logical meaning follows a part/part structure, where the parts are dependent on each other (Halliday, 1993b: 38).

Analysis of experiential ideational meaning at the discourse semantics stratum focuses on the taxonomic relations between elements (entities and Processes) from clause to clause: nuclear relations and activity sequences (Martin & Rose, 2003). Nuclear relations refer to configurations of elements within each clause to show the roles of people and things in activity sequences. Activity sequences are glossed as “series of events that are expected by a field” (Martin & Rose, 2003: 101), the unmarked relation between events being ‘and’. Events that are related by cause and effect such as those found in science fields are termed implication sequences.

The logical relations between activity and implication sequences are construed through the resources of Conjunction. While external conjunction contributes to field by organising the activity or implication sequences, internal conjunction contributes to the organisation of the text and organises the rhetoric (Halliday & Hasan, 1976; Martin, 1992). Thus, while external conjunction enters into the logical system of ideation, internal conjunction is considered textual, not ideational. The four main categories of conjunction are given in table 2.1.

External and Internal	Addition	additive alternative
	Comparison	similar different
	Time	successive simultaneous
External	Consequence	cause means purpose condition
Internal	Consequence	concluding countering

Table 2.1 General options for external and internal conjunction (Martin and Rose, 2003)

2.3.2. Interpersonal meaning at the discourse semantics stratum

There is a considerable body of research outside SLF that studies interpersonal meaning in academic discourse. Hood (2010: 7) notes that it is impossible to reconcile the different approaches because they derive from different underlying theories of language, research interests, and developments in research and theory. She provides a synthesis of this research, in which she groups studies according to the way evaluation is treated: as grammar, lexis or discourse (Hood, 2010: 13-22); and although some of this work has informed studies on shell nouns, it will not be reviewed here. In SFL interpersonal meanings can be realised by different selections from the lexicogrammar. It is “strung throughout the clause as a continuous motif or colouring” (Halliday, 1979 cited in Martin & White, 2005: 19), it is diffused (Halliday & Matthiessen, 1999), it spreads (Martin, 1992), and it radiates (Hood, 2006) across clause boundaries and stretches of text. This patterning is referred to as prosody (Halliday, 1979 in Martin & White, 2005; Martin, 1992; Martin & Rose, 2003). Martin and White (2005: 19-24) identify three types of prosody: saturating prosody, where evaluative language is peppered throughout a stretch of discourse; intensifying prosody, where there is a build-up by means of repetitions of various kinds; and

dominating prosody, where the interpersonal meaning spreads to the surrounding discourse, imbuing it with interpersonal meaning. At the stratum of discourse semantics interpersonal meaning draws on the resources of the system of APPRAISAL (Martin, 2000; Martin & White, 2005).

The system of APPRAISAL is a discourse semantic resource that construes power and solidarity. There are three interacting domains: ATTITUDE, GRADUATION and ENGAGEMENT. The domain of ATTITUDE is concerned with feelings: AFFECT deals with resources for expressing emotional reactions; APPRECIATION with how things – material or semiotic – are evaluated; and JUDGEMENT with assessments of behaviour or people (figure 2.3).

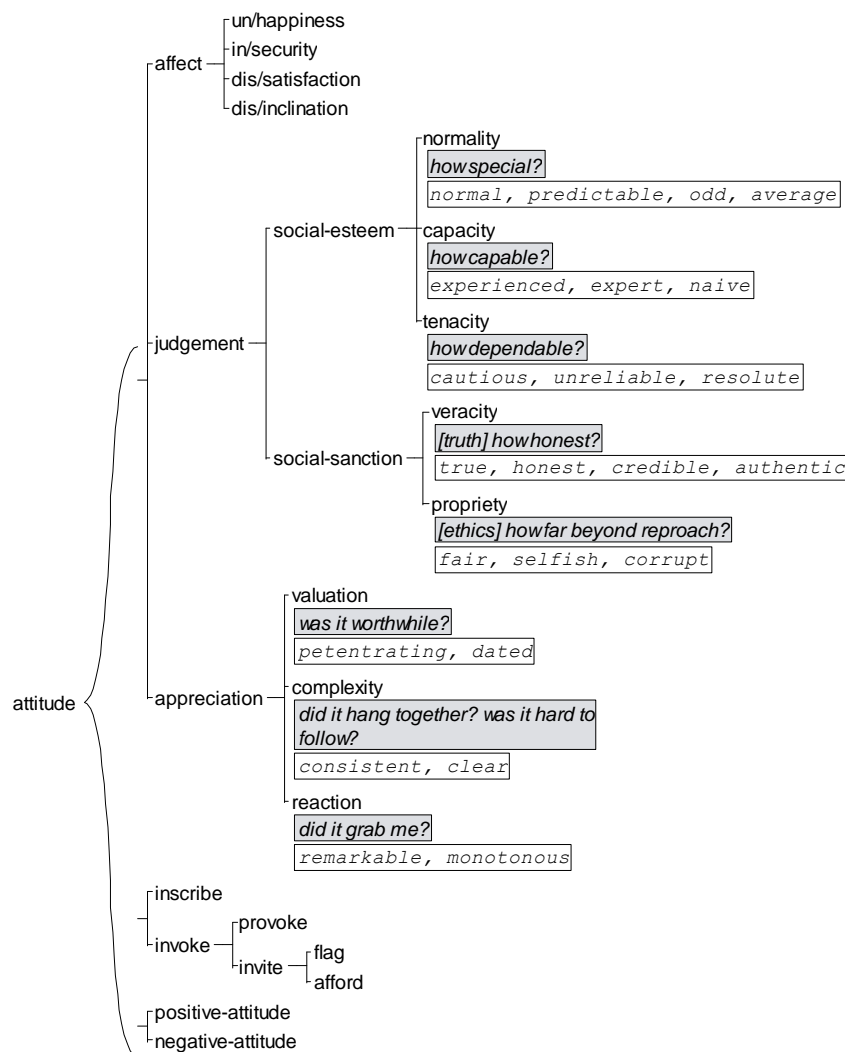


Figure 2.3 The system of ATTITUDE (adapted from Martin & White, 2005)

The system of JUDGEMENT is further broken down in delicacy such that it reflects grammatical distinctions in the systems of modalisation and modulation (Martin & White, 2005: 54). Within the subsystem of social esteem, normality is related to modalities of usuality, capacity to ability, and tenacity to modulations of inclination. Within social sanction, veracity is related to modalisations of probability, and propriety to modulations of obligation. (See Martin & White, 2005: 52-56.) On the basis of these relations Martin and White (2005: 55-56) posit a cline with grammaticalised realisations at one end and lexicalised realisations at the other. In between are modality metaphors. An example of such a cline for ability/capacity is given in figure 2.4.

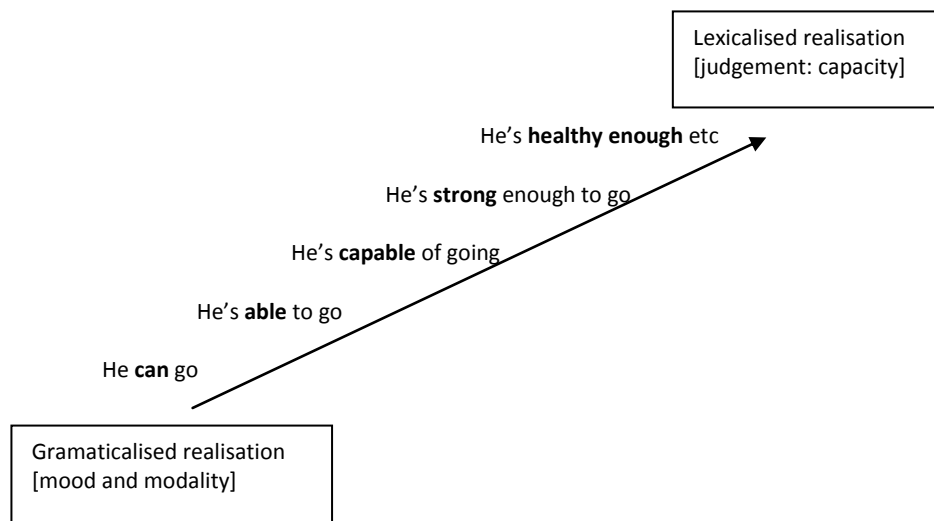


Figure 2.4 Cline of interpersonal grammar and appraisal (adapted from Martin & White, 2005:55)

While the system of JUDGEMENT construes evaluations of people or people's behaviour, the system of APPRECIATION deals with the evaluation of things, both semiotic and material, abstract and concrete. There are three sub-types: reaction, composition and valuation. Realisations for reaction can be probed by prompts about the impact (did it grab me?) or quality (did I like it?) of the thing, while realisations for composition can be probed by prompts about the composition (did it hang together?) or the complexity (was it hard to follow?) of the thing. Martin and White (2005: 57)

claim that valuation (was it worthwhile?) is field sensitive because it depends on the writer's institutional focus; what is valued in one field is not necessarily so in another.

Positive or negative evaluations may be inscribed in the lexical realisation of attitude or they may be invoked. In the former, the positive or negative appraisal is inherent to the lexical choice. In the latter, meanings that would not normally be read as evaluative are imbued with positive or negative values by means of prosody – saturating, intensifying or dominating. Similarly, inscribed appreciation of an activity realised as a Thing may invoke a judgement of the person who did the activity (Martin and White, 2005; Ventola 1998), and vice versa. Martin and White (2005: 67) suggest reading the options for inscribing and invoking attitude as a cline, from 'inscribe' to 'afford'; the lower down the cline, the greater the freedom allowed to the reader to align himself with the values of the text or not.

Attitudinal lexis is inherently gradable, i.e. its meanings can be made more or less intense by selecting resources from the system of GRADUATION. Martin and White's (2005) original network for GRADUATION has since been extended by Hood (2010) (figure 2.5).

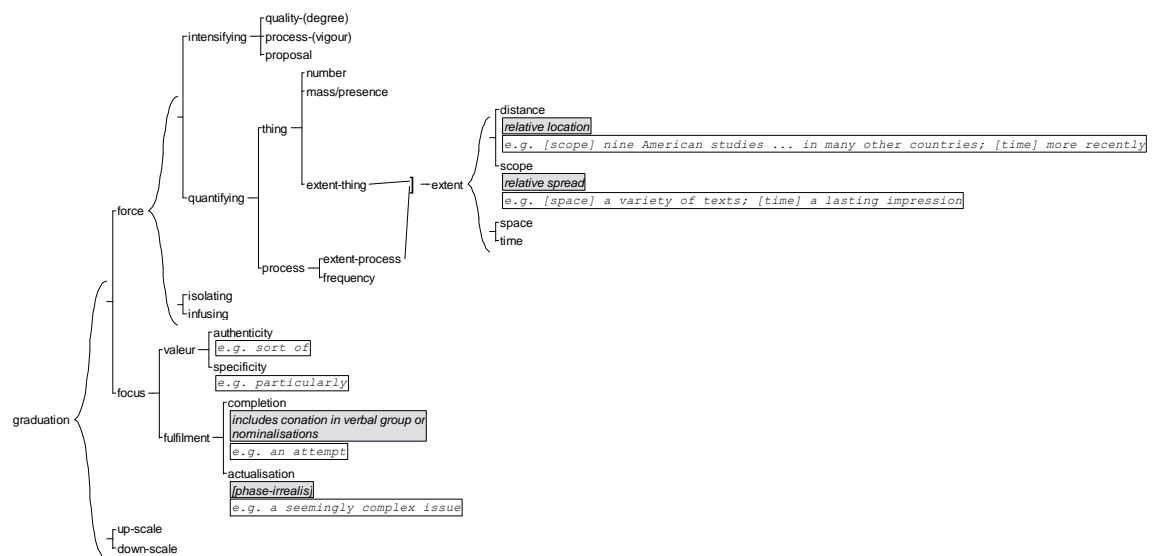


Figure 2.5 The system network of GRADUATION (Martin & White, 2005, extended by Hood, 2010)

The resources of the system of GRADUATION can be used to heighten or lessen attitudinal meanings in two ways: either by intensifying an inscribed value or by adjusting the experiential boundaries so as to sharpen or blur them (Martin & White,

2005). The subsystem of Force is used for the former and that of Focus for the latter. Hood (2010: 86) notes that while premodification of an attitudinal quality intensifies the quality, premodification of a nominalised quality quantifies rather than intensifies it. Thus, processes and things may be quantified and as a result invoke an attitudinal reading.

The third domain of APPRAISAL is that of ENGAGEMENT. Unlike approaches that focus on the writer's evaluation of or commitment to the truth of a proposition (e.g. Hyland, 2000; Conrad & Biber, 1999; Hunston, 1999; 2011), ENGAGEMENT is concerned with the ways in which writers position themselves on a topic with regard to other writers' positions. In other words writers are assumed to be engaged in a dialogue with others, even though they are not present as they are in a conversation. By means of the resources of ENGAGEMENT writers can choose the extent to which they acknowledge the contestability of what they say as well as align themselves with other writers' views. The system network for ENGAGEMENT is given in figure 2.6.

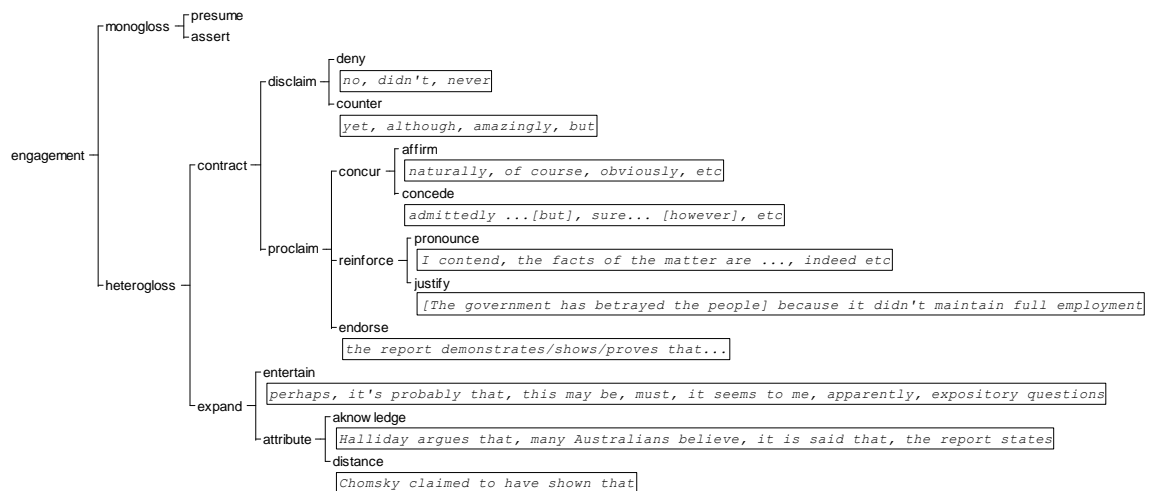


Figure 2.6 The system network of ENGAGEMENT (Martin & White, 2005, adapted)

The ENGAGEMENT network presents the writer with choices as to how much dialogic space is given to authorial voices other than the writer's. That is to say, the writer chooses whether to open up the dialogic space or close it partially or even completely. In monoglossic text, only the authorial voice of the writer is present; propositions are presented as if it were unnecessary to dispute them. In doing so, no other viewpoints are acknowledged. By contrast, in heteroglossic text, the author's

voice is seen to operate against a backdrop of other voices. The dialogic space allowed to these other voices varies, and thus contributes to the stance of the writer. Through the resources of the subsystem of Expansion, the writer may choose to open up the dialogic space to acknowledge other voices, while through the resources of the subsystem of Contract, the writer chooses to close the dialogic space, presenting propositions as warrantable and suppressing alternative views.

Martin and White (2005: 124-126) note that resources may be paired for rhetorical effect and to build solidarity with the reader. For example options from Concur or Concede often precede Counter as the writer first aligns himself with a shared view, or presents the view as reasonable, but then counters it with another of his own. The resources of the system of Engagement contribute to creating the writer's stance by helping position him with regard to the information in the text and alternative views.

2.3.3. Textual meaning at the discourse semantics stratum

Textual meaning at the discourse semantics stratum follows a wave-like pattern, with peaks and troughs. This rhythm is referred to as periodicity (Martin, 1992; 1995b; Martin & Rose, 2003) or information flow patterns (Halliday & Matthiessen, 2004: 588). In abstract discourse, positions of prominence are associated with macro- and hyper-Themes, and macro- and hyper-News (Martin, 1992), which establish a hierarchical structure for the text. Martin (1992: 447) suggests that the hyper-Theme is a metaphorical marked theme for the text that signals a shift in the conceptual development of text as a whole. It is not "an orthographic place, that is, the beginning of the paragraph, but a structural place, namely a step in the argument" (Ravelli, 2004: 113), and it corresponds more or less to a topic sentence of a paragraph, predicting the thematic development of the paragraph. A macro-Theme is a larger unit still, such as a paragraph or heading, that predicts a set of hyper-Themes. Macro- and hyper-Themes serve as compositional scaffolding (Martin, 1992: 444) for the text and contribute to its coherence. By contrast, the hyper-New and macro-New bring together meanings that have been developed through the text. In short, while a hyper-Theme *predicts* phases of discourse, a hyper-New *distils* information, and they work together in what Martin (1992: 456) calls a "textured sandwich" (figure 2.7).

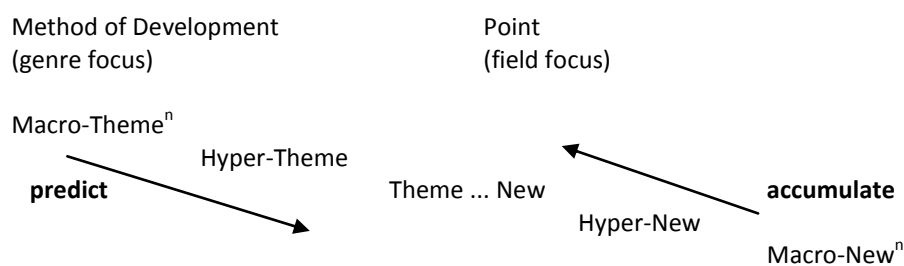


Figure 2.7 Sandwich texture in abstract written discourse (Martin 1992:456)

In addition to the structural scaffolding afforded by hyper- and macro-Themes and hyper- and macro-News, participants can be tracked through a text by means of the system of IDENTIFICATION (Martin, 1992). The system of IDENTIFICATION draws on the notion of phoricity, where phoric items are marked as recoverable or known and non-phoric items are presented as new to the reader/listener. Non-phoric entities are introduced with Presenting reference, which can be realised by non-specific Deictics such as 'a' or 'some'. Martin distinguishes between three kinds of phoricity: reminding, relevance and redundancy. Reminding phoricity identifies a particular entity as known, and it draws on the system of Presuming Reference, which is realised by specific Deictics such as 'the', 'this', proper names, or pronouns. Relevance phoricity establishes the entity as known by means of its relation to another. Generally this relation is one of comparison. Redundancy phoricity signals that the experiential identity of the entity must be recovered in the co-text, and it is realised by means of substitution and ellipsis. The three kinds of phoricity can combine freely, and the system accounts for this possibility and that of generic reference by positing three systems that function simultaneously (figure 2.8). The systems are then extended in delicacy to capture differences in meaning.

An initial distinction is drawn between effected reference, described above, and neutralised reference. Neutralised reference is used in 'little texts' (Martin, 1992: 120) such as headings. In such texts the deixis is not fully effected.

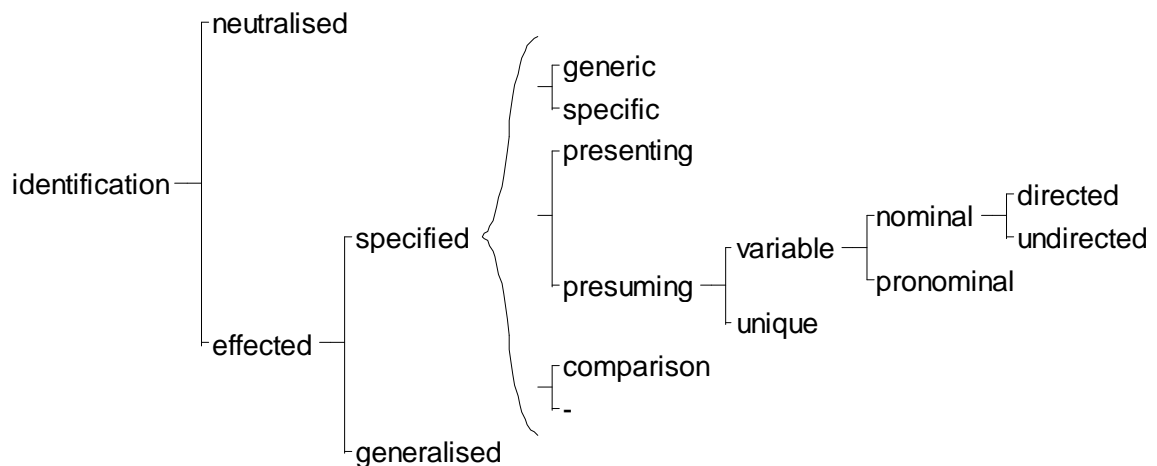


Figure 2.8 The system network of IDENTIFICATION – primary delicacy (Martin, 1992: 120)

2.4. Instantiation

The system networks of the lexicogrammar and discourse semantics reflect a realisational hierarchy of abstraction, in which meaning is simultaneously encoded at all strata. Recently, SFL has shifted its lens to look at how texts construe meaning in terms of instantiation, although this aspect is said to be under-theorised (Martin, 2008; 2010). Instantiation is a scale of generalisation that relates system to text, and it involves all levels of realisation (Martin, 2010) (figure 2.9). At one end of the cline is the meaning potential of the system and at the other a particular reading of an instance of text (Martin & White, 2005: 23). In between are patterns of typical instantiation (Halliday & Matthiessen, 1999; 2004; Martin & White, 2005).

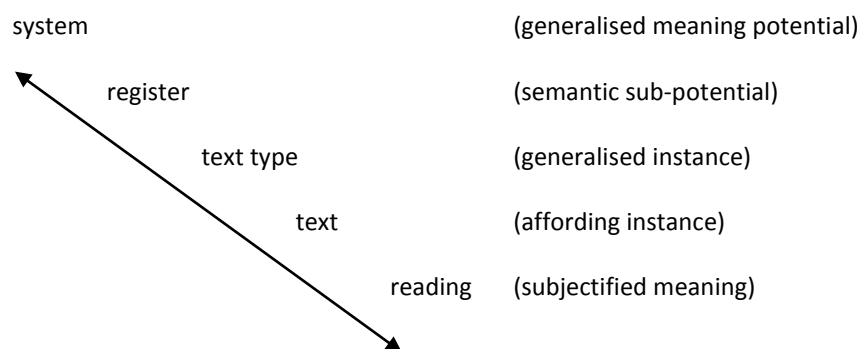


Figure 2.9 Cline of instantiation (Martin & White, 2005: 25)

Martin (2008) argues that there is a need for an instantiation perspective in analyses to complement that of realisation. An instantiation perspective can show, for

example, the extent and ways in which a text both conforms to yet differs from typical expectations of a text type. Martin and White (2005) have used an instantiation perspective to examine the typical choices at the level of ‘key’ in newspaper articles and history texts, positing a cline of instantiation for evaluation, as shown in table 2.2.


	<ol style="list-style-type: none"> 1. appraisal (system) – the global potential of the language for making evaluative meanings, e.g. for activating positive/negative viewpoints, graduating force/focus, negotiating intersubjective stance 2. key (register) – situational variants or sub-selections of the global evaluative meaning making potential – typically reconfiguration of the probabilities for the occurrence of particular evaluative meaning-making options or for the co-occurrence of options 3. stance (text-type) – sub-selections of evaluative options within text; patterns of use of evaluative options within a given ‘key’ associated with particular rhetorical objectives and the construction of authorial personae 4. evaluation (instance) – instantiation of evaluative options in text 5. reaction (reading) – the take-up of evaluative meanings in a text according to the listener/reader’s subjectively determined reading position; the attitudinal positions activated by the reader as a result of their interaction with the text
---	---

Table 2.2 Cline of instantiation – evaluation (Martin & White, 2005: 164)

Martin (2008) proposes two dimensions through which the cline of instantiation can be explored: coupling and commitment. Coupling refers to the ways that meanings from different systems combine in pairs, triplets or larger numbers. As Martin (2008: 40) puts it: “What realisation freely combines, instantiation specifically couples”. For example Martin & White’s (2005) analysis of key for the different voices used in newspaper articles is based on patterns of exclusion/inclusion of particular options from the Appraisal system. Commitment refers to the amount of meaning manifested in terms of ideational meaning or interpersonal meaning. For instance afforded attitude is less committed than inscribed attitude. Hood (2008), for example, uses commitment to ideational meaning to explore the kinds of changes in meaning that are implicated in writing a summary of a source text.

2.5. Individuation and affiliation

Related to instantiation are individuation and affiliation. Individuation and affiliation are concerned more with the language user and the options that he/she mobilises from the system, with which he/she construes a social identity as belonging to a community (Martin, 2010). At one end of the individuation cline are the resources of the system, at the other the selections used by the individual. In particular, the coupling of ideational meaning with interpersonal meaning contributes to the

construal of communal identities, aligning the reader into a shared community of values. Individuation/affiliation as a cline can be seen as two trajectories: from the perspective of individuation, the analysis reveals how the writer individuates his/her own self, while from the affiliation perspective, it reveals how the writer aligns with the community (Martin, 2010). The two trajectories are shown in figure 2.10. Hao and Humphrey (2012), for example, have drawn upon the individuation/affiliation hierarchy to show how students in biology use ideational and interpersonal coupling strategically to “burnish”, i.e. evaluate positively, their own study and those with which they align, but “tarnish”, i.e. evaluate negatively, other studies with which they do not wish to align in order to align themselves with a community yet individuate themselves.

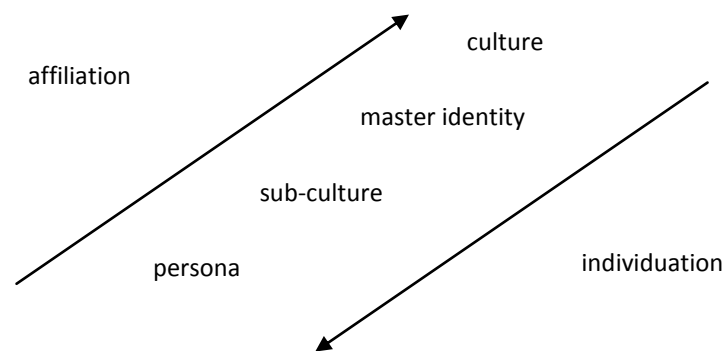


Figure 2.10 Individuation and affiliation (Martin, 2010: 24)

An instantiation perspective seems to offer rich ground for this study given that it is concerned with a community of writers – Portuguese and non-Portuguese – writing for each other in what has been considered an important knowledge-producing text type: the research article (Hyland: 2000; 2009; Klamer & McCloskey, 1988). As argued in the introduction, publishing in English has become as crucial for career advancement for academics in many non-English speaking countries as it is in many English-speaking countries. A brief review of insights into the research article follows. However, given the vast array of studies, which are underpinned by a wide range of theoretical frameworks, only a selective account can be given here. The aspects

deemed relevant for the study of shell nouns in this thesis include the research article as genre, and grammatical metaphor.

2.6. The research article as genre

Genre is a multi-faceted term that has been used in a myriad of ways within the study of discourse. Hyon's (1996) classification of genre studies into three traditions – English for Specific Purposes (ESP, e.g. Swales, 1990), New Rhetoric (e.g. Freedman, 1994) and SFL – has helped to situate studies.

In the ESP tradition the research article is considered a genre. This tradition has produced descriptions of the organisation of sections (Myers, 1994), the moves of abstracts (Martín, 2003; Samraj, 2005) the introduction section (Berkenkotter & Huckin, 1995; Samraj, 2005; Swales, 1990); the methods section (Lim, 2006); the transition from the results section to the conclusions (Yang & Allisson, 2003), the results section (Brett, 1994; Hopkins & Dudley-Evans, 1988), and the discussion section (Berkenkotter & Huckin, 1995, Holmes, 1997). More recently attention has been drawn to the wide range of macro-structures of research articles (Lin & Evans, 2012). In particular Lin and Evans note that writing guides are limited in their explanations of research article macro-structures and fail to explain sections such as the literature review. Studies carried out in the ESP Swalesian tradition of moves have been criticised by Hood (2010) because the separation of text into moves is based upon intuitive, common sense criteria rather than linguistic.

In SFL genre is defined as a staged goal-oriented social process (Martin & Rose, 2003). Martin and Rose (2008: 6) reformulate this definition in functional linguistic terms: “what this means is that genres are defined as a recurrent configuration of meanings and that these recurrent configurations of meaning enact the social practices of a given culture”. Thus genres are seen as operating at the stratum of the culture. At the stratum below genre, the resources of register – field, tenor and mode – are seen as generalising across genres (Martin & Rose, 2008:16) (figure 2.2, p 10).

Within SFL there is some inconsistency in the way the research article is treated in relation to genre. Martin and Rose (2008: 200), for example, classify research articles written by research scientists as procedural recounts. On the other hand Hood

(2010: 37) considers the research article to be a macro-genre, i.e. a larger text made up of shorter genres (Martin & Rose, 2008: 218). Despite the ambiguity of the status of the research article in relation to genre, Martin and Rose (2008: 207) propose a six-stage structure for a research article in the field of science (table 2.3), and they break down some of the linguistic resources used at each stage in an analysis of one from the scientific field of metallurgy. However, these stages clearly would not apply in a non-empirically oriented research article such as in the field of history, as evidenced by results from work on the research article in the ESP tradition (e.g. Lin & Evans, 2012) and implied by Martin and Rose's (2008: 99-139) topography of genres in history. Perhaps the research article is best seen from an instantiation perspective as a text type, the genres that constitute it depending on field.

Abstract	A brief summary of the experimental method and the results and discussion
Introduction	Locates the text in the development of the field by reference to previous research. Establishes a problem that previous research has not dealt with. States intention of current research.
Experimental Details (i.e. Method)	Lists experimental methods used, including equipment and procedures.
Results and Discussion	Presents experiment results in graphic and mathematical form. Interprets these results verbally. Reasons about the probable cause of the problem.
Conclusions	Summary of reasoning
References	Previous research

Table 2.3 Stages of the research article (Martin and Rose, 2008:207)

Hood (2006; 2010) has identified the generic stages in research article introductions from an SFL perspective. The macro-genre establishes a warrant for research, and the stages are shown in table 2.4. The shift in genre is marked by a shift in field, the boundaries of which are signalled by a hyper-Theme. At the same time, considerable interpersonal meaning is being developed at each stage as the writer draws on lexicogrammatical resources first to persuade the reader that the object of study is of significance, then that there is room for new knowledge in the field, and finally that the writer's own study will contribute to new knowledge.

Macro-genre	Establishing a warrant for research
Genre	Descriptive report of object of study
Genre	Descriptive report of other research
Genre	Description of writer's own study

Table 2.4 Generic structure of the research article introduction

While the control of objective structures to replace overtly subjective or spoken registers is often considered the mark of competent professional academic discourse (Hood, 2010), also important is the need to engage critically with the material, which implies subjective positioning. These two apparently conflicting needs need to be met if the writer is to be deemed professional (Hood, 2010). Hood's (2006; 2010;) work on evaluation in research article introductions illustrates the variety of resources that are drawn upon to position the writer as a credible member of the discourse community. For example she finds that when reporting other people's research, writers tend to invoke attitudinal values by means of the resources of Graduation rather than inscription, and this contributes to a more objective sounding text. A second insight from her research is that there is an interplay between two fields in the introduction of a research article: the field of research can be seen to be projecting the field of the object of study, each field favouring different resources from Appraisal. While the field of research favours invoked evaluative resources, the field of the object of study draws on both invoked and inscribed attitude.

The importance of interpersonal choices coupled with ideational selections in research articles as the writer seeks to align himself with like-minded academics and disalign himself with others has also been highlighted by Ventola (1998). She proposes a cline of alignment/bashing similar to the concepts of burnishing and tarnishing (Hao & Humphrey, 2012), which could equally be applied to the research article.

Other work on research articles in an SFL tradition has highlighted differences in Thematic choices according to the section. Martínez (2003) notes a higher proportion of multiple Themes and marked Themes in the discussion section of research articles in biology than in the methods section. She attributes this difference to the purpose of each section. The purpose of the methods section is to recount procedures, so it makes greater use of the object of research as Theme. The discussion section is more argumentative, and the differences in Theme reflect movement from fact to interpretation, due to differences of purpose. She also finds that the discussion section makes greater use of grammatical metaphor as Theme than does the methods section.

2.7. Register and scientific discourse

Despite the acknowledged importance attributed to the research article and the wealth of research on it from non SFL theories, there seems to be comparatively little work focussed on it from an SFL genre perspective. By contrast, much work has been undertaken at pre-tertiary and tertiary educational levels to establish the nature of specific disciplinary discourses (e.g. Halliday & Martin, 1993; Martin & Veel, 1998) and genres (e.g. Martin and Rose, 2008). Work on the register of scientific discourse has highlighted the importance of grammatical metaphor (Halliday & Matthiessen, 1999; 2004) in disseminating scientific knowledge. This work has been grounded in different time frames, which include the evolution of scientific discourse, i.e. a phylogenetic perspective (e.g. Halliday, 1988/1993; 1998; Halliday & Matthiessen, 1999; Banks, 2005), the development of the genres used in schools as children's language develops, i.e. an ontogenetic perspective (e.g. Christie & Cléirigh, 2008; Martin, 1991; Martin & Rose, 2003: 38-44; Veel, 1998), and the unfolding of text, i.e. a logogenetic perspective (e.g. Halliday, 1990/2002; Guillén Galve, 1998). The three time frames are shown in figure 2.11. A brief explanation of grammatical metaphor as it relates to this study follows.

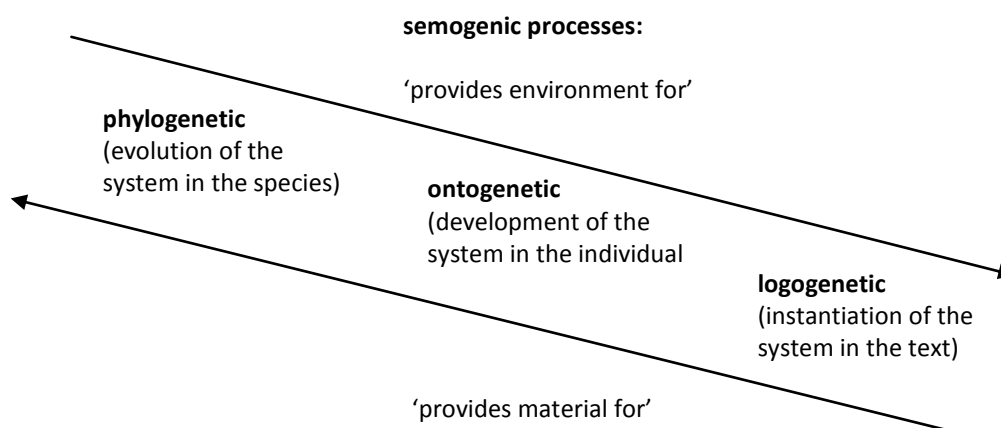


Figure 2.11 The three semohistories related (Halliday & Matthiessen, 1999: 18)

2.8. Grammatical metaphor

Grammatical metaphor involves a shift in the realisational domain of a phenomenon such that a sequence, which is typically realised as a clause nexus joined by a conjunction, is reconstrued as a figure, typically realised as a clause, a figure is reconstrued as an element of clause structure, or an element of clause structure is reconstrued as a different transitivity role. Notwithstanding, the grammatical metaphor does not lose the status of the congruent phenomenon (Halliday & Matthiessen, 1999; 2004). Examples are given in figures 2.12 and 2.13.

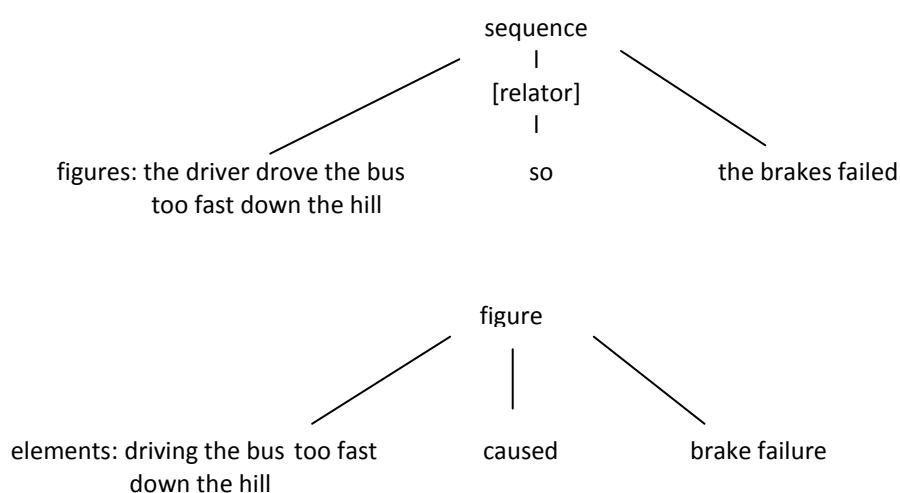


Figure 2.12 Grammatical metaphor: sequence reconstrued as figure (adapted from Halliday, 1998: 190,191)

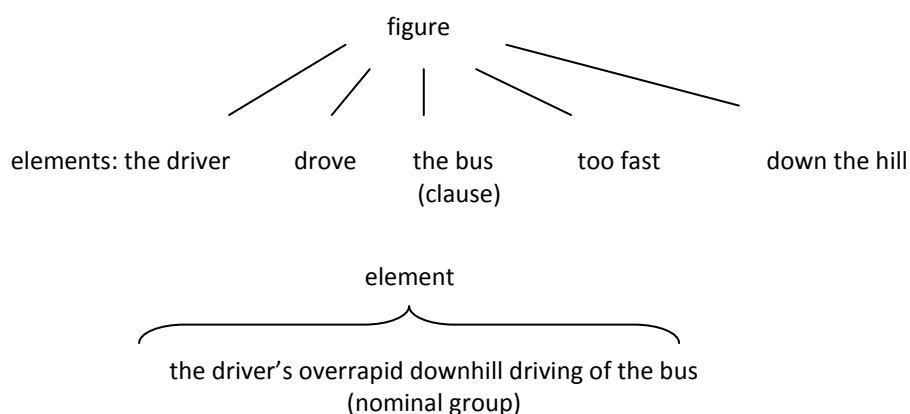


Figure 2.13 Grammatical metaphor: figure reconstrued as element (adapted from Halliday, 1998: 191)

Various transformations have taken place in the examples above. In the first, the logico-semantic Relator ‘so’ has been reconstrued as Process, and the clauses have been shifted down in rank to function as Participants. In the second, ‘the driver’ [Actor] has been reconstrued as Deictic, ‘too fast’ [Circumstance] as Epithet, ‘down the hill’ [Circumstance] as Classifier, ‘drove’ [Process] as Thing, and ‘the bus’ [Goal] as Qualifier, such that the elements functioning at the rank of the clause are now functioning as constituents within a nominal group. Typical clausal realisations of phenomena as Process + Participant + Circumstance are referred to as being congruent, while metaphorical realisations, i.e. those that involve grammatical metaphor, are referred to as incongruent (Halliday, 1990/2002; Halliday & Matthiessen, 1999). Halliday and Matthiessen (1999: 537) argue that congruent realisations typically precede metaphorical realisations across phylogenesis, ontogenesis and logogenesis.

There are 13 types of elemental grammatical metaphor (table 2.5), of which five reconstrue the congruent realisation as Thing. A grammatical metaphor of this type is termed nominalisation. As Halliday and Matthiessen are careful to point out, the examples given in the table are examples, not glosses for the class.

congruent:	metaphorical:			
	=» circumstance	=» process	=» quality	=» thing
quality =» <i>unstable</i>				1 <i>instability</i>
process =» <i>absorb</i>			3 <i>absorptive</i>	2 <i>absorption</i>
circumstance =» <i>instead of;</i> <i>on the surface</i>		6 <i>replaces</i>	5 <i>alternative;</i> <i>superficial</i>	4 <i>replacement</i> <i>surface</i>
relator =» <i>for/because [b,</i> <i>for/because a]</i> <i>so [a, so b]</i>	10 <i>because of;</i> <i>as a result</i>	9 <i>causes,</i> <i>proves;</i> <i>ensues,</i> <i>follows from</i>	8 <i>causal;</i> <i>consequent</i>	7 <i>cause,</i> <i>proof;</i> <i>result</i>
∅ =»		12 <i>occurs;</i> <i>imposes;</i> <i>does, has</i>		11 <i>phenomenon,</i> <i>fact</i>
thing, circumstance =» <i>driver [be safe]</i> <i>decided [today]</i>	13 expansion of thing <in environment of 1 or 2> <i>driver [safety], driver's [safety], [safety] of the driver</i> <i>today's [decision], [decision] of today</i>			

Table 2.5 Domains of elemental metaphors (Halliday & Matthiessen, 1999: 245)

Grammatical metaphors can be oriented towards ideational (experiential or logical), interpersonal (Halliday & Matthiessen, 1999) or textual meaning (Martin, 1992). An example of each kind of meaning orientation is given in 2.1 – 2.5. The metaphor is underlined. While examples 2.3 and 2.4 are both oriented towards interpersonal meaning, 2.3 is explicitly subjective, and 2.4 is explicitly objective (Halliday & Matthiessen, 2004: 613-616), the latter contributing to a more objective sounding text.

Ideational: experiential

2.1 Beliefs arise from perceptions. (Halliday & Matthiessen, 1999: 588)

Ideational: logical

2.2 Beliefs arise from perceptions. (Halliday & Matthiessen, 1999: 588)

Interpersonal

2.3 I don't believe that pudding will ever be cooked. (Halliday & Matthiessen, 2004: 613)

2.4 The Granada move has also focused attention on the possibility that other companies will become targets for stake-building. (Schmid, 2000: 237).

Textual

2.5 That point is just silly. (Martin, 1992: 417)

Grammatical metaphor has become an important resource in modern scientific discourse for representing and explaining physical phenomena (Halliday, 1988/1993; Halliday & Martin, 1993; Banks, 2005). Halliday (1988/1993) schematises its evolution through time (figure 2.14), where the first set corresponds to a relation between two events, *in rebus*, and the second corresponds to a relation between two stages in the discourse, *in verbus* (Halliday & Matthiessen, 1999: 262,263).

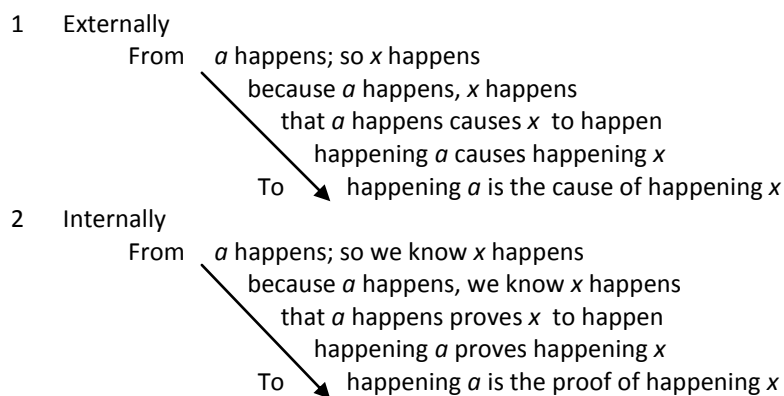


Figure 2.14 Evolution of grammatical pattern in scientific English (Halliday, 1988/1993: 66)

The scheme indicates a drift towards nominalisation such that nominalised Processes functioning as Participants come to be related through relational Processes, and these relational clauses, in particular intensive or circumstantial relational clauses, have become a prevalent feature in scientific or academic writing (Halliday, 1998; Ravelli, 2004). According to Halliday (1988/1993) many verbs other than 'be' have come to realise the relational Process. Verbs realising external relations (i.e. *in rebus*) in intensive relational clauses include, among others, *be*, *become*, *form*, *equal*, *represent*, *constitute*, *symbolise*, and *mean*, while verbs realising circumstantial relational Processes, themselves examples of grammatical metaphor, include *cause*, *lead to*, *accompany*, *follow*, *produce*, *arise from*, *result from*, to name but some. Examples 2.6 and 2.7 illustrate an identifying intensive relational Process clause and a circumstantial relational Process clause respectively. The grammatical metaphor is in bold.

2.6 These people **constitute** a reservoir for the transmission of the virus. (Halliday & Matthiessen, 2004: 235)

2.7 This situation **is** apparently **caused** by anomalous low temperatures. (Halliday & Matthiessen, 2004: 243)

Internal intensive relations - *in verbis* - may be realised by verbs such as *prove*, *show*, *predict*, *illustrate*, *suggest*, *indicate*, and *confirm*. Such verbs are an important resource for developing an argument in academic writing (Halliday, 1988/1993; Halliday & Matthiessen, 2004; Christie & Cléirigh, 2008; Martin, 1992). Because many of the verbs that realise internal relations also function as sources of projection, Halliday (1988/1993) suggests that when they function to relate nominalisations, they are more clearly relational, but when they introduce a projected clause, they may also be interpreted as mental Processes, glossed by 'makes us think', such as in example 2.8.

2.8 Our discovery of the importance of molecular diffusion near the crack tip **indicates** that surface coatings might be designed to block the opening of the crack. (Halliday, 1988/1993:65)

Thus, two interpretations are possible for 2.8 given that the *Senser* may be realised by a product of human consciousness (Halliday & Matthiessen, 2004: 203). The two analyses are given in figure 2.15.

Our discovery of the importance of molecular diffusion near the crack tip **indicates** [[that surface coatings might be designed to block the opening of the crack]].

Token	Pro. Rel	Value
-------	----------	-------

Our discovery of the importance of molecular diffusion near the crack tip **indicates** that surface coatings might be designed to block the opening of the crack.

Senser	Pro. mental	Phenomenon
--------	-------------	------------

Figure 2.15. *Indicate* interpreted as identifying Relational Process and as Mental Process. Adapted from Halliday, 1988/1993: 65

Christie and Cléirigh (2008: 17) refer to this kind of fuzziness as ‘symbolic slippage’, and they demonstrate how it is possible to produce alternate analyses for verbs such as *show*, *indicate*, or *reveal* when the Subject is abstract and the Complement is a projected clause. They argue that the verb may be parsed as an intensive identifying relational Process, in which case the projected clause is embedded as a fact clause, or as a verbal Process, in which case the projected clause is hypotactic and construes the Locution. Their choice to interpret these verbs as verbal rather than mental Processes is based on the principle of systemic indeterminacy (Halliday & Matthiessen, 2004: 173), on whose basis the grammar of experience is interpreted as a continuous space within which verbal Processes shade into relational Processes (Halliday & Matthiessen, 2004: 172). Interpreted relationally, these verbs are a valuable resource for construing abstract relationships because Token and Value are ‘stratally distinct’ (Halliday & Matthiessen, 2004: 230), Value being understood to be more abstract than Token.

Halliday (1998) argues that the grammatical metaphor of nominalisation in particular unlocks meaning potential for the discourse. When Processes or qualities are nominalised, they can then be expanded using all the resources of the system available to Thing, and the potential for modification is extremely useful in the construal of technical taxonomies. In addition, nominalisation renders the meaning less negotiable. In Halliday’s (1993b: 39) famous words: “you can argue with a clause but you can’t argue with a nominal group”. This reification (Halliday, 1998; Halliday & Matthiessen, 1999), i.e. reconstruing the meaning of a figure as a nominal group, creates the impression of a stable entity, which makes nominalisation particularly suitable for scientific writing as the apparent stability of the phenomena in the discourse is in

keeping with scientific goals (Banks, 2005). It also identifies the occasion of a process as unique (Halliday & Matthiessen, 1999: 266).

Another benefit stemming from nominalisation is that the nominalisation can enter into environments that a clause cannot. This textual feature can be strategically used to structure the reasoning of the argument, carrying it forward (Halliday & Matthiessen, 2004). For example, choosing a nominalisation to compose Theme can serve to orient the flow of information. Halliday (1998: 228) refers to this feature as 'compacting'. Nominalisations in Theme position that follow a more congruent realisation have also been called "ensuing intratextual dynamic grammatical metaphors" (Guillén Galve, 1998). Similarly, nominalisations can become the focus of a clause when they compose New. As Halliday (1998: 205) states: "The complex interplay of Theme + Rheme in the clause with Given + New in the information unit constitutes an immensely powerful discursive resource: it is the primary source of energy for the dynamic of scientific and technical argument." Nominalisation enables the writer to take advantage of these systems in order to present information as foregrounded or backgrounded.

The textual benefits of compacting information by means of a nominalisation at the level of the clause also hold for larger units of information such as sections of a research article. Ravelli (2004) has noted their usefulness in hyper-Themes, and Guillén Galve (1998) has found them prevalent in the final sections of abstracts and research article introductions. He has called these "culminating grammatical metaphors".

Despite the payoffs of nominalisation, there may also be costs. The nominalisation may result in some loss of meaning such that logical relationships are left implicit rather than explicit (Halliday, 1989/1993; 1993b; Halliday & Matthiessen, 1999). In addition, nominalisations also affect the density of a text with the result that a dense text may be more difficult to understand. Nonetheless, it is possible to adjust the density of the nominal group in which a nominalisation construes Head. The three examples from Gould's critique on Darwin, which are discussed in Halliday (1993b), will provide the starting point for this discussion. The grammatical metaphors of interest – all nominalisations – are highlighted in bold.

- 2.9 Yet the peculiar character of this evidence has not matched **Darwin's prediction of a continuous rise in complexity towards Cambrian life**. (Gould, 1989/1991: 57 in Halliday, 1993b: 39)
- 2.10 Thus, instead of **Darwin's gradual rise to mounting complexity**, the 100 million years from Ediacara to Burgess *may have witnessed* three radically different faunas ... (Gould, 1989/1991: 57 in Halliday, 1993b: 34)
- 2.11 He insisted that any complex Cambrian creature must have arisen from a lengthy series of Precambrian ancestors with the same basic anatomy. (Gould, 1989/1991: 57 in Halliday, 1993b: 38)

Halliday notes, in particular, that 2.10 is ambiguous and that it illustrates the loss of precision arising from the grammatical metaphor. While examples 2.9 and 2.10 contain nominalisations, 2.11 is a projecting clause and is clearly the most congruent. The projected clause is still up for negotiation, so to speak. Halliday uses a projecting clause to unpack 2.10 (figure 2.16).

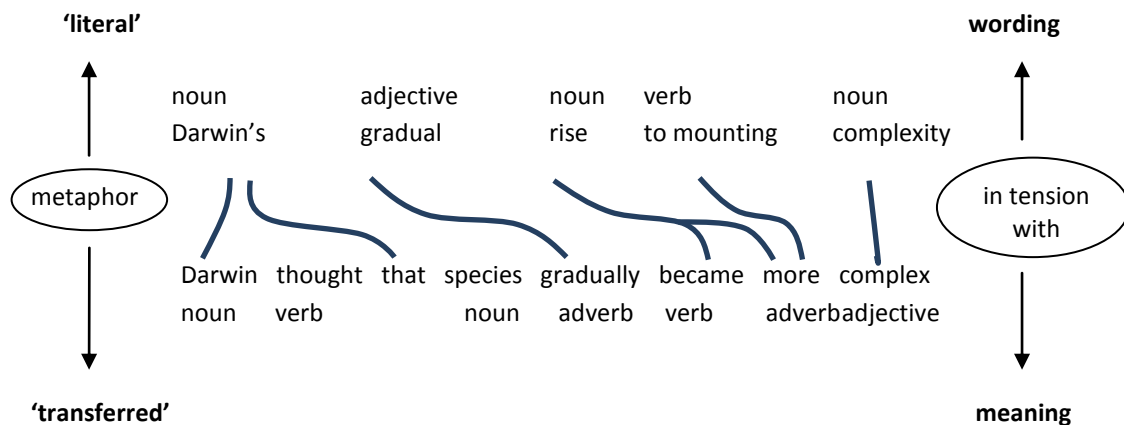


Figure 2.16 The tension between literal and figurative readings of Gould's nominalisation of Darwin's theory (Halliday, 1993b: 31)

2.9 represents an intermediate step between the others. Assuming a congruent version similar to that of 2.10 (figure 2.16), the verbal Process is nominalised, the Sayer is reconstrued as Deictic, and the projecting figure is reconstrued as Qualifier (figure 2.17).

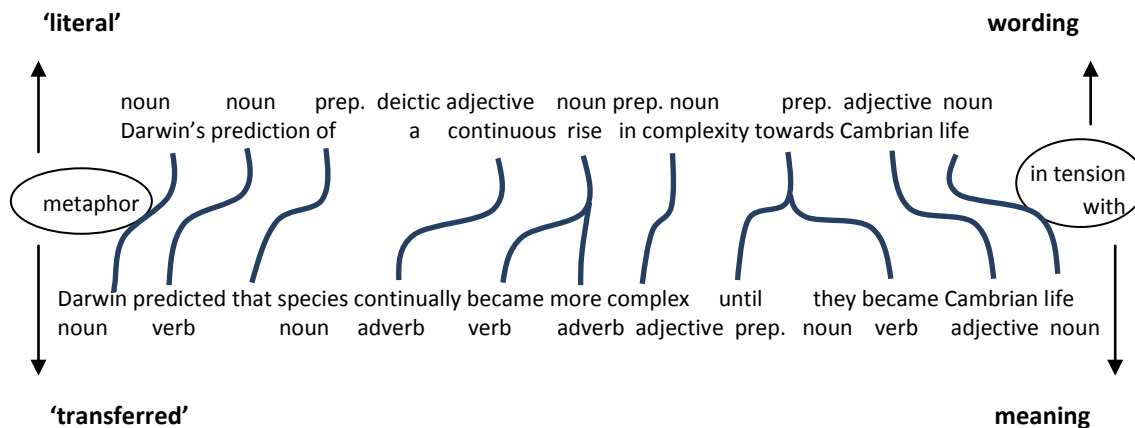


Figure 2.17 The tension between literal and figurative readings of Gould's nominalisation of Darwin's prediction²

Because the grammatical metaphor is a nominalised verbal Process, it is also possible to realise the Qualifier as a projected clause rather than prepositional phrase, and as a result, it is possible to adjust the realisation of the projected clause. For example, less information is lost in 2.12 than 2.13, both variations on 2.9. The first rewrite construes a more congruent version, the second retains the grammatical metaphor, which construes Existent.

2.12 Yet the peculiar character of this evidence has not matched **Darwin's prediction that species continually became more complex until they became Cambrian life.**

2.13 Yet the peculiar character of this evidence has not matched **Darwin's prediction that there was a continuous rise in complexity until Cambrian life.**

Nouns of projection such as 'prediction' seem to represent an intermediate step in the formation of nominalisations analogous to the fact clauses in the evolution of patterns of grammatical metaphor featured in figure 2.14 (p.29). While nominalisations of verbal and mental Processes all have the ability to project, i.e. be modified with a projected clause as Qualifier, other grammatical metaphors – fact nouns such as *evidence*, *problem*, or *fact* – also have this ability. Such nouns have been called 'shell nouns' (Schmid, 2000). Other shell nouns include *concept*, *possibility*, *surprise*, or *knowledge*, to name but a few. It is these nouns that are the subject of this

² Due to space constraints and because it is not the focus of the example, 'towards Cambrian life' has not been fully unpacked.

study. Examples of shell nouns in sentences are given in 2.14 – 2.19. In each case the shell noun – a grammatical metaphor – is underlined.

- 2.14 This is consistent with the concept that the Antarctic ozone hole phenomenon causes a dilution effect throughout much of the Southern Hemisphere (Halliday & Matthiessen, 2004: 645)
- 2.15 A standard empirical hypothesis is that one component of the mind/brain is a *parser*, which assigns a percept to a signal (abstracting from other circumstances relevant to interpretation). (Halliday & Matthiessen, 2004: 645)
- 2.16 The Bill is short and modest in scope, and it is doubtful whether the other Private Members' Bills in the offing will fill all the gaps. This fact may give the government an extra excuse for counselling patience until the next report from the Molony committee. (Halliday & Matthiessen, 2004: 471)
- 2.17 You've said that one of your editorial rules is not to publish your buddies. (Halliday & Matthiessen, 2004: 479)
- 2.18 Warwick Town Council originally decided to build its own crematorium, but in April last year it abandoned the idea and entered into a joint scheme with Leamington Town council and Warwick Rural District Council. (Halliday & Matthiessen, 2004: 471)
- 2.19 The talks lasted for three hours. This was a surprise, for they had only been scheduled to last two hours. (Halliday & Matthiessen, 2004: 468)

2.9. Summary of chapter 2

This chapter has sought to provide a theoretical backdrop of Systemic Functional Linguistics for the study of shell nouns. It has presented the theoretical tenets of SFL and briefly introduced some of the systems at the lexicogrammatical and discourse semantics strata, which will underpin the analyses undertaken in this thesis. In addition, it has provided a short account of SFL insights into the research article from the perspectives of genre and register. Grammatical metaphor – nominalisation above all – has been highlighted as a strategic resource in academic writing. Textually, it contributes to the structuring of the unfolding argument by means of its interplay with the systems of Theme and Information structure. Ideationally, it construes phenomena as stable entities rather than a quantum of change. Interpersonal modality metaphors can also be nominalised, which contributes to a more objective sounding

text. Finally, a relation between grammatical metaphor and shell nouns has been posited. This relation will be explored more fully in the following chapter, which focuses specifically on shell nouns.

Chapter 3. Shell nouns

This chapter aims to provide a theoretical basis for the study by reviewing the literature on shell nouns. The review will comprise two parts: the first will focus on how shell nouns have been described and how they contribute to the unfolding discourse. This section will not be restricted to a single theoretical perspective. While its starting point is the position of Schmid (2000), I shall discuss other scholars' positions. The second part will focus on how shell nouns fit into a Systemic Functional Linguistics perspective. It will examine shell nouns at the stratum of the lexicogrammar, then move to their use at the discourse semantics stratum.

3.1. A brief definition of shell nouns

Schmid (2000: 4) defines shell nouns thus:

Shell nouns make up an open-ended functionally-defined class of abstract nouns that have, to varying degrees, the potential for being used as conceptual shells for complex, proposition-like pieces of information.

He places them at the intersection of three functional properties, as shown in figure 3.1.

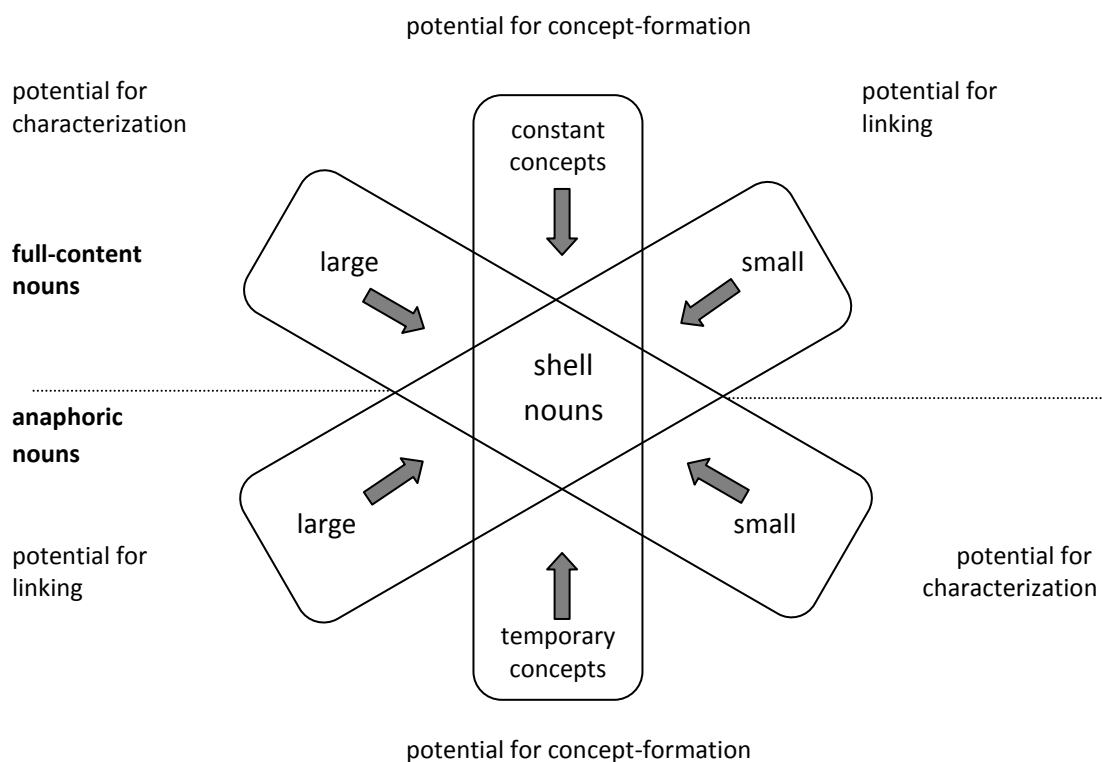


Figure 3.1: The converging balance of shell nouns (Schmid, 2000: 19)

For Schmid, shell nouns are neither full content nouns nor anaphoric nouns. While part of their meaning is stable, another part must be retrieved from the co-text. It is this 'semantic gap' (Schmid, 2000:76) that endows shell nouns with a degree of unspecificity, and it must be filled for the shell noun to gain communicative effectiveness. He argues that the gap is often visible in dictionary definitions. For example, a *fact* is "*a piece of information* that is known to be true" (Schmid, 2000: 77), where 'a piece of information' represents the semantic gap and 'that is known to be true' is the stable part of the meaning. At the same time, the stable part of the meaning makes it possible to characterise or label the discourse segment that the shell noun refers to. In this way, the noun can be seen to be a "conceptual shell" that holds or encapsulates information realised elsewhere in the co-text.

3.1.1. The functions of shell nouns

Although shell nouns may share the feature of a semantic gap, the class of shell nouns is not defined by this property. Schmid argues that it is a functional class, i.e. nouns become shell nouns because of the way they are used and can only be considered shell nouns if used in that way. Three functions define their use:

1. "Shell nouns serve the **semantic function** of *characterizing* and perspectivizing complex chunks of information which are expressed in clauses or even longer stretches of text.
2. Shell nouns serve the **cognitive function** of *temporary concept-formation*. This means that they allow speakers to encapsulate these complex chunks of information in temporary nominal concepts with apparently rigid and clear-cut conceptual boundaries.
3. Shell nouns serve the **textual function** of *linking* these nominal concepts with clauses or other pieces of text which contain the actual details of information, thereby instructing the hearer to interpret different sections of a text together." (Schmid, 200:14)[my bold, italics in original]

The following examples taken from Schmid illustrate the three functions. The shell noun group is in bold, and the content that it encapsulates is underlined.

Semantic function - characterising

- 3.1 (I won the freshmen's cross-country. – Mm) That was a great achievement wasn't it? (Schmid, 2000: 22)

The speaker has chosen to characterise his winning the cross-country as an *achievement* rather than, say, a *surprise* or a *fact*.

Cognitive function – temporary concept formation

3.2 **The fact** that the head of one of Nigeria's northern-based regiments, based in Kaduna, has already pledged support for President Babangida reinforces the view that the coup is ethnically-based. (Schmid, 2000: 362)

This example was spoken. Although it is possible to leave out 'the fact', Schmid argues that to do so would increase the cognitive burden of processing the information for the listener. By heading the nominal group with 'the fact', it becomes a bounded entity, which facilitates its processing by the listener.

Textual function – linking

3.3 Yet another theory suggests that the goal of the welfare state, within a society in which economic competition under capitalism dominates, must be to effect gradual reform. **This Fabian approach** argues for ... (Schmid, 2000: 344)

The shell noun provides a link that enables the reader to identify the antecedent easily whereas demonstrative reference might not. The textual link may function anaphorically or cataphorically. In sections 3.4 and 3.5 it will be argued that the semantic and textual functions can be considered to correspond roughly to the interpersonal and textual metafunctions of Systemic Functional Linguistics respectively.

Schmid rejects identity of reference as the relation underlying the link between shell nouns and shell content in situations involving anaphoric reference. He argues that a "substitution view" (Schmid, 2000: 28) inherent in identity of reference is inappropriate when "the antecedents of the anaphora are normally not referring expressions but clauses, extended stretches of discourse, or even pieces of information which must be inferred from the context" (Schmid, 2000:28). To overcome this he coins "experiential reference" as the means by which the shell noun and the shell content are related. In doing so he incorporates a more cognitive view in which "items with referring potential are seen as being related neither to the text itself [endophoric] nor to the world outside the text [exophoric] but to the cognitive models that are created in the minds of language users" (Schmid, 200:28). Schmid is able to take this

approach because he does not restrict himself to one theoretical framework in his study. However, it seems difficult to justify given that Halliday and Hasan (1976) stress that reference is a semantic relation, i.e. one of meaning, and not a grammatical one:

Reference is presupposition at the semantic level. A reference item signals that the meaning is recoverable though not necessarily in the form of the actual word or words required. For this reason a reference item cannot necessarily be replaced by what it presupposes; even if the presupposed item is present in the text, the reference to it may require an item of a different function in structure." (Halliday and Hasan, 1976:145)

Their view clearly runs counter to Schmid's claim that identity of reference is based on substitution. Claiming experiential reference as the link between the shell noun and its antecedent does have one advantage. The same relation can be used to explain cases where the shell content follows the shell noun. An abridged example is given in 3.4.

3.4 As we will see later, there are basically **two ways** of going about studying conversation and other basic linguistic interaction: **in one, we can just study what's going on Another, more theoretical approach** tries to go 'behind conversation', as it were ... (Schmid, 2000: 358)

In this example, 'two ways' sets up the expectation that the writer will explain what the 'two ways' are. The shell content is further identified by the marked Theme 'in one', which is an example of ellipsis (in one way), and 'Another, more theoretical approach', which is an example of comparative reference (Halliday and Hasan, 1976). I shall argue in section 3.5.1 that such cataphoric uses of shell nouns can be accounted for by Martin's analysis of hyper-Themes and macro-Themes at the discourse semantics stratum.

3.1.2. Lexicogrammatical patterns of shell nouns

For Schmid, to be classified as a shell noun, the noun must be able to appear in specific lexicogrammatical environments: with a postnominal (non-relative) *that-*, *to-* or *wh*-clause (as in example 3.2) and/or "as subjects of copula clauses with *that-*, *to-* or *wh*-clauses as subject complements" (Schmid, 2000:80). An example of the latter is given in 3.5. In the example the shell noun is in bold, and the content it encapsulates – the lexicalisation – is underlined.

3.5 **The fact** is that neither of these institutions have significant social or political research arms. (Schmid, 2000: 97)

Schmid uses these two lexicogrammatical environments to retrieve shell nouns from the British section of the COBUILD Bank of English in order to identify them, and then looks for other uses on the basis of the nouns retrieved with the former patterns. This second search found that shell nouns are also used cohesively with anaphoric reference (as in example 3.3) and in the pattern *th-be-N*, where *th-* stands for a demonstrative reference item (as in example 3.1). Thus, he describes 4 patterns for use as a shell noun which are based on the lexicogrammatical environments in which they appear. These patterns and an example of each are given in table 3.1. In the table the shell noun is marked in bold and its lexicalisation is underlined.

Shell noun Pattern	Example
N-cl	The fact <u>that the head of one of Nigeria’s northern-based regiments, based in Kaduna, has already pledged support for President Babangida</u> reinforces the view that the coup is ethnically-based. (Schmid, 2000: 362)
N-be-cl	The fact is <u>that neither of these institutions have significant social or political research arms.</u> (Schmid, 2000: 97)
th-N	Yet another theory suggests that <u>the goal of the welfare state, within a society in which economic competition under capitalism dominates, must be to effect gradual reform.</u> This Fabian approach argues for ... (Schmid, 2000: 344)
th-be-N	<u>(I won the freshmen’s cross-country. – Mm) That was a great achievement</u> wasn’t it? (Schmid, 2000: 22)

Table 3.1: Shell noun patterns (Schmid, 2000)

3.1.3 Semantic classification of shell nouns

Schmid’s analysis of the British section of the COBUILD Bank of English found 670 lexemes that were used as shell nouns. He classifies them into five classes – factual, linguistic, mental, eventive, and circumstantial – depending on the semantic features of their use, across which cut four broad semantic groups: general, specific, attitudinal, and modal (table 3.2).

	Factual	Linguistic		Mental		Eventive		Circumstantial
General	Neutral <i>thing, fact</i>	Propositional <i>news, gossip argument</i>		Conceptual <i>idea, theory, mystery</i>		General eventive <i>event, act, situation</i>		General circumstantial <i>situation, position</i>
Specific	causal <i>reason</i> evidential <i>evidence</i> comparative <i>difference</i> partitive <i>example</i>	(fact-related) assertive <i>statement</i> rogative <i>question</i>	(event-related) directive <i>suggestion</i> commissive <i>promise</i>	(fact-related) creditive <i>knowledge</i> dubitative <i>doubt</i>	(event-related) volitional <i>belief</i>	specific eventive <i>attempt</i>		specific circumstantial <i>procedure</i>
Attitudinal	attitudinal factual <i>problem</i>	expressive <i>complaint</i>		emotive <i>surprise, worry</i>		attitudinal eventive		
Modal	epistemic <i>possibility likelihood</i>	(epistemic)	(deontic)	(epistemic)	(deontic)	deontic <i>need</i>	dynamic <i>Ability</i>	dynamic

Table 3.2: Semantic relations between classes of shell nouns (adapted from Schmid, 2000:300, examples added)

Each class is further subdivided into ‘groups’ and each group into ‘families’ on the basis of the cognitive components inherent in the shell noun and the relations between them, which Schmid interprets as a frame. The notion of ‘frame’ is borrowed from Talmy, who describes it as “a set of conceptual elements and relations that [...] are evoked together or co-evoked together” (Talmy, 1996: 238 cited in Schmid, 2000: 91). For example, while *reason* and *result* both evoke a relation of cause and effect, the former noun directs the listener’s attention to the cause, and the latter to the effect. The frame for each noun would reflect the similarity [CAUSAL], as well as the difference: focus on cause/focus on effect. Membership of a family is not exclusive; a shell noun may appear in more than one family because it is used in more than one way. Because this study is not concerned with the finer categorisation of shell nouns, I shall restrict my analysis to the six main classes. These classes and examples of each are given in table 3.3 below.

Semantic feature	Type of experience being described	Examples
[FACTUAL]	facts, states of affairs	thing, reason, evidence, difference, aspect, problem
[MENTAL]	ideas, cognitive states and processes	belief, doubt, wish, fear
[LINGUISTIC]	utterances, linguistic acts and products thereof	news, argument, statement, suggestion
[MODAL]	possibilities, abilities, permission, obligations, etc.	possibility, permission, ability
[EVENTIVE]	activities, processes, states	event, attempt, trouble mistake
[CIRCUMSTANTIAL]	events	situation, position, case

Table 3.3 Semantic features functioning as classifiers (Schmid, 2000:88, examples added)

In his classification of shell nouns, Schmid recognises that some shell nouns are more marginal than others. While he considers some to be prime shell nouns (e.g. *fact*, *thing*, *cause*, *difference*, *problem*), or good shell nouns (e.g. *proposition*, *observation*, *claim*, *gripe*), others are more marginal (e.g. *event*, *act*, *situation*, *position*, *time*). In general, the circumstantial and general eventive shell nouns fall among the latter.

3.1.4. Pragmatic uses of shell nouns

Schmid (2001) examines *N-be-cl (that)* pattern, as in example 3.6, for its pragmatic functions. He argues that the nominal group in Theme position has the pragmatic function of focalising in that it enables the *that*-clause to receive end focus, i.e. be presented as New. It can also act as a hesitation device. A second pragmatic function – that of topicalising – is achieved by means of the nominal group acting as a

cataphoric signpost for the information in the *that*-clause. He claims that “[t]he construction triggers the expectation that the topic is highly accessible and represents information which is shared by the discourse participants” (Schmid, 2001: 1539).

3.6 But **the greatest fear** is that the new arrangement will deplete the numbers of Britain’s researchers severely and permanently. (Schmid, 2001: 1543)

Because the stable semantic component of the noun confers the *that*-clause with a semantic presupposition, the pattern can be strategically employed by speakers for different purposes. Some nouns can reduce or heighten the speaker’s commitment to the embedded clause. For example a shell noun such as *danger* or *risk* expresses epistemic possibility while *fact* or *feeling* express epistemic certainty. In addition, use of the shell noun makes the proposition sound more objective (Schmid, 2001: 1544) and so enables speakers to present personal beliefs as factual information. With evaluative nouns such as *problem* or *advantage*, two semantic presuppositions are triggered: that the evaluative nature of the noun is uncontested and that the embedded clause is true. Schmid argues that in such a case it is the truth value of the embedded clause that will be contested rather than the semantic categorising of the noun. He summarises the benefits of the pattern:

“What speakers gain from using the nouns in this construction is in fact quite a lot, then. They manage to sell their own personal views and opinions as objective truths and facts. And what is more, by exploiting the pragmatic presuppositions of the *N-be-that* construction, they are able to create the impression that their views-disguised-as-truths represent given knowledge apparently shared by all discourse participants anyway.” (Schmid, 2001: 1545)

His view suggests that shell nouns in this pattern may be related to the discourse semantics system of ENGAGEMENT. The contribution of shell nouns to ENGAGEMENT will be addressed in section 3.5.2.

This section has provided an overview of what Schmid means when he talks about shell nouns. He was, of course, not the first to notice their use, but he was perhaps the first to provide a systematic account of their functions in discourse as well as a full description of what nouns can be used as shell nouns. The next section will review other scholars’ treatment of shell nouns.

3.2 Other research on shell nouns

Research prior to Schmid (2000; 2001) has tended not to focus exclusively on shell nouns although they are included in the analyses. Several scholars have noted the cohesive function of certain nouns in referring to a stretch of discourse rather than a participant (Winter, 1977; 1992; Ivanič, 1991; Francis, 1994; Tadros, 1994). Further, it has also been noted that these nouns can provide a label for the chunk of text being referred to (Francis, 1994; Tadros, 1994; Charles, 2003). These two functions correspond to Schmid's functions of linking and characterisation respectively. The third function – temporary concept formation – has received less attention in the research literature. Section 3.2 will provide an overview of the different studies.

3.2.1 Early approaches to shell nouns

Perhaps the first to recognise that shell nouns play a role in discourse was Winter (1977). He identified three closed-set classes of lexical items in written discourse that function to anticipate information in subsequent clauses. The first two classes, Vocabulary 1 and Vocabulary 2, roughly correspond to SFL Conjunctions and conjunctive Adjuncts, respectively. The third class, Vocabulary 3, includes shell nouns although the class is not restricted to nouns. He argues that because the lexical items predict that there is a semantic relation with another clause, the clause containing these lexical items remains communicatively incomplete unless the relevant information to complete the relation is retrieved from an adjacent clause (Winter, 1977: 7). The second clause is said to provide the 'lexical realisation' (Winter, 1977: 7-8) of the relation predicted by the Vocabulary 3 item. The relation is either a logical sequence relation or matching relation. Logical sequence relations include those of time and space, such as premise-conclusion, i.e. enhancing relations, while matching relations are those of similarity or difference, i.e. relations of extension or enhancement (manner). The clause providing the lexical realisation for the Vocabulary 3 item may precede or follow the item, that is, the Vocabulary 3 item can refer anaphorically or cataphorically. Winter places a default condition such that the lexical realisation of the second part of the semantic relation is expected to follow the clause

if it has not appeared in the previous one.³ In such a case, the Vocabulary 3 item can be seen as a 'signpost', a term later picked up by Ivanič (1991) and Schmid (2000; 2001), and it anticipates the following clause. Anticipation is defined as "the forward reference of the contextual function of lexical realisation" (Winter, 1977: 28). Winter notes that Vocabulary 3 items are abstract, and he implies a distinction between the cohesive systems of substitution and reference that parallels Halliday and Hasan (1976), where substitution is seen as a grammatical relation and reference a semantic one. He notes that Vocabulary 3 items refer to "the information" of the other half of the clause pair, rather than "repeating" it by means of a "substitute clause" (Winter, 1977: 70). For Winter, then, Vocabulary 3 nouns can refer anaphorically or cataphorically to another clause, and it is this clause that provides the lexical realisation of the noun.

It is quite clear from this description of Vocabulary 3 items that the nouns within the class are performing the same function as shell nouns. In fact, Winter identifies and gives numerous examples of two of the patterns described by Schmid (2000): *Th-N* in anaphoric uses and cataphoric uses; and *Th-N-be-cl*, which is one of Schmid's function-defining lexicogrammatical patterns for shell nouns. Furthermore, he notes that Vocabulary 3 items do not only "represent the notions underlying a clause relation..., but also that they themselves may be used to make the relation explicit by *saying* what the relation is" (Winter, 1977: 22). In other words, they can be used to label the referent clause, as later argued by Flowerdew (2003a;b), Francis (1994), Charles (2003), Schmid (2000; 2001), Tadros (1994), and Yamasaki (2008). And finally, he notes the potential benefits that the nouns bring to the discourse due to the fact that they can be modified with adjectives. This latter claim is in line with the benefits of nominalisation as argued by Halliday (1988/1993; 1998), and exemplified in studies on shell nouns (cf. Willis, 1993; Francis, 1994; Flowerdew, 2003a; Dueñas, 2003-2004). Winter (1992) later refined his analysis of Vocabulary 3 nouns, calling them 'unspecific metalanguage nouns'. He claimed that they can be made specific by

³ Sinclair (1993) also argues that there is a basic discourse structure where by default the previous sentence is encapsulated in the current one. In addition, he shows that various lexicogrammatical choices, some of which include shell nouns, can explicitly encapsulate retrospectively or prospectively the information in the adjacent clause(s).

means of modification that restricts their identity or by lexicalisation of a clause. The latter corresponds to shell noun use.

If Winter can be seen as setting the grounds for a description of two of the functions of shell nouns, textual – linking and semantic – characterising, subsequent research has both corroborated his claims and extended them. Like Winter’s study, many of the studies are not restricted to shell nouns. They may focus on one of the patterns identified by Schmid (2000), or on different lexical realisations of a function, as did Winter. In particular the studies by Francis (1994) and Charles (2003; 2007) have been influential.

In her study of the textual function of shell nouns, which she calls retrospective and advance labels, Francis (1994) found that retrospective labels play a discourse organising function and are used to “signal that the writer is moving on to the next stage of his/her argument, having disposed of the preceding stage by encapsulating or packaging it in a single nominalization” (Francis, 1994: 86). Thus, they play a textual function and contribute to the cohesion of the text. This function is borne out by Shaw (2000), who found that shell nouns were often used in the transition of statement of results to the next stage in the cycle to create an argument in the discussion section of dissertations and results and discussion section of research articles. In line with Winter (1977) and Schmid (2000), Francis claims that “[t]he relationship between a label and the clause(s) it replaces is not a random process of naming, but an encoding of shared, or sharable, perceptions of the world” (Francis, 1994:100). The discourse organising function of labels has been contested by Flowerdew (2003c), who argues that when the shell noun ‘problem’ is used sentence internally, i.e. in the pattern *Th-be-N*, e.g. *This is a serious problem*, the shell noun/label is not organising discourse but evaluating it. Further, the burden of anaphoric reference is carried by the Deictic *This*, and not the shell noun (Flowerdew, 2000c: 505). The preponderance of this pattern to carry out an evaluative function is corroborated by Yamasaki (2008). Although Flowerdew’s critique seems valid, Francis (1994) only considers labels that have a definite article, so she does not examine such uses, and strictly speaking, Flowerdew’s critique does not apply. Francis, however, does note that interpersonal meaning may be prospective, that is, it may be taken up and continued. The two uses – discourse

organising and evaluative – and their lexicogrammatical environments are accounted for in Schmid’s analysis by the textual and semantic functions respectively.

Francis claims that a retrospective label is always presented as given information in the clause structure even though the head noun is a new lexical item and does not have a synonym in the preceding discourse. It often acts as Carrier for an attribute. Schmid (2000: 12) surmises that this use is perhaps behind Ivanič’s (1991) term for shell nouns: carrier nouns. By contrast, advance labels are presented as new information. Francis notes that both the label and its modifiers can carry ideational, interpersonal and textual meaning. While ideational modifiers can enable the noun to enter into taxonomic relations with other entities, interpersonal modifiers evaluate the encapsulated clause, creating stance as Charles (2003; 2007) argues. Francis posits that the economy with which evaluation can be included may constitute a reason for using retrospective labels: “Writers seem to choose the labelling device because of the modification options it offers: by choosing a nominalization ... they can get in their evaluation without having to make a special point of it” (Francis, 1994: 97). It is these modifiers that will be brought out in an analysis at the discourse semantics stratum, which will be discussed in more detail in section 3.5. Evaluative modifiers such as *important* can be captured in an analysis of Appraisal (Martin & White, 2005). Textual modifiers, such as *same*, *different*, or *another* have a metalinguistic function and “contribute directly to the organizational role of labels” (Francis, 1994: 98). As will be seen later, in section 3.5.1, these textual modifiers are operating in the system of internal conjunction (Martin, 1992). Hoey (1993) makes similar claims about the motivations for using the shell noun ‘reason’ as Subject in relational clauses (*Th-N be cl*). In his study of *reason* as a ‘lexical signal’ (i.e. shell noun) to express cause, he found that this was the most common pattern in the written part of the Birmingham Corpus of English, and he argues that the possibility for comment on the reason by means of evaluative pre-modification could well motivate its use. Other motivations for the pattern are (1) textual: it allows the writer to count the reasons to help orient the reader; (2) ideational: post modification can serve to refocus or clarify the writer’s position; and (3) informational: by placing *reason* in Theme position, it allows the cause to receive end focus. The prevalence of this pattern for words like *question*,

problem and *difficulty* in the British National Corpus has been noted by Willis (1993), who argues that such commonly occurring patterns and the lexis that appears in them should be taught to students.

Charles (2003, 2007) studies the contribution of shell nouns, which she calls labels, to enact the interpersonal metafunction in dissertations by Masters and PhD students in the fields of politics and the material sciences. She focuses on the patterns *Th-N* in sentence initial position and *N cl (that)*. She argues that both the choice of shell noun and its modifiers can indicate the writer's stance through the shell noun's functions as a discourse organiser and the act of labelling the referent clause(s). The labels are interactive in so far as "their use affects the reader's perception of the proposition(s) put forward and so enables the reader to perceive the organisation and meaning that the writer intends" (Charles, 2003:318). Labels such as *claim*, *theory*, and *assumption* indicate epistemic stance while labels such as *problem*, *limitation*, or *atrocious* indicate attitudinal stance (Biber *et al.*, 1999; Conrad & Biber, 1999). Charles further notes that because sentence-initial, anaphorically referring shell nouns (i.e. Theme conflated with Subject) are presented as given information, the reader is constrained to accept the labels assigned, at least provisionally. Schmid (2001) similarly argues that shell nouns in the pattern *Th- N be cl* (e.g. The trouble is that it is not a forecast and only records the average level reached the day before (Schmid, 2001: 1547)) can trigger a presupposed semantic categorisation of the shell content that the listener must provisionally accept, although on reflection he may not in fact agree with it. This view implies that choosing an inappropriate shell noun will affect the reader's perception of the information and may prompt rejection of the writer's position by undermining the persuasiveness of the argument being constructed.

A second function identified by Francis (1994:84) – that shell nouns "assign a particular status in the argument" – is also taken by Charles (2003) to indicate stance. The example Charles gives is that sentence-initial 'This position' sets up an environment in which the reader expects the writer to contest, modify or endorse the position previously developed. Given that her corpus comprised Masters dissertations and Doctoral theses, such an expectation on the part of the reader is certainly warranted. Tadros (1994) similarly found that advance labels, i.e. shell nouns whose

lexicalisation is prospected by the noun, commit the writer to a discourse act. I will later argue that such instances contribute to establishing dialogic position.

Both Francis (1994) and Charles (2003; 2007) classify the labels/shell nouns. Their major distinction is between metalinguistic nouns and other nouns. Metalinguistic nouns are used to “label a stretch of discourse as being of a particular type of language. They are used by the writer to forge relationships which are located entirely with the discourse itself” (Francis, 1994:89). Although their terminology differs, Francis (1994) and Charles (2007) further subdivide the metalinguistic nouns into mental process nouns/idea group; illocutionary nouns/argument group; language activity nouns/evidence group; and text nouns. The first element in each pair corresponds to Francis’ terminology while the second corresponds to Charles’. With the exception of the last group, text nouns, each set has a clear counterpart in Halliday and Matthiessen (2004): projection nouns of the ‘ideas’ type; projection nouns of the ‘locutions’ type; and fact nouns of the ‘proof’ type respectively. Text nouns – for example *phrase*, *sentence*, *excerpt*, or *section* – strictly speaking are not shell nouns in that they cannot be used in the function-defining lexicogrammatical patterns identified by Schmid (2000). Francis herself notes that they are formal labels that involve no interpretation. Such nouns have also been called deictic nouns (Gray & Cortes, 2010). Within the ‘other’ group, there are non-metalinguistic nouns, such as *result*, *effect*, which, unlike metalinguistic nouns, exist outside the discourse (Charles, 2003); possibility group nouns (Charles, 2003; 2007); epistemic stance nouns; and attitudinal stance nouns (Charles, 2003). Epistemic and attitudinal stance are terms borrowed from Biber *et al* (1999: 972), who define their functions respectively as to comment “on the status of information in a proposition” to mark, for example, certainty, doubt actuality, or precision, and to “report personal attitudes or feelings” (Biber *et al*, 1999: 974). These stance nouns will be picked up in an Appraisal analysis at the discourse semantics stratum in section 3.5.2. Any abstract noun functioning as a label that is not included in the above groups is classified as ‘other’, and this group includes nouns such as *fact*, *case*, or *concern*. The ‘possibility’ group of nouns corresponds to Halliday and Matthiessen (2004) fact nouns of the ‘chances’ type, while the ‘others’ might be taken to be fact nouns of the ‘cases’ type. It is worth noting that

the subdivision in classes does not have a one-to-one correspondence with the six classes identified by Schmid.

Despite the similarities and complementary aspects of the functions attributed to shell nouns and their description in the above studies, there is one point on which there is no consensus: that is whether shell nouns belong to an open or closed set of nouns. Schmid (2000), and Francis (1994), for example, consider them to be an open class set of nouns. Schmid argues that any noun functioning as a shell noun can belong to the class. This view underpins his rejection of Halliday's position that projection nouns belong to clearly defined classes (Schmid, 2000: 12). For Winter (1977, 1992), the class is a closed set, albeit with some characteristics of an open set because the nouns may be modified like any other open-ended noun. Nonetheless, the fact that most Vocabulary 3 nouns can paraphrase the closed set vocabularies of conjunction and conjunctive Adjuncts, and that those that do not behave in the same way – i.e. they relate two clauses by means of lexical realisation – makes them a closed set (Winter, 1977: 27-28). On the other hand, Ivanič (1991) considers shell nouns to have the characteristics of both an open and closed class of vocabulary.

As has become clear through the review above, early studies on shell nouns have neither used the same terms to refer to the nouns under study nor classified them in the same way. Nor have the analyses been restricted to shell nouns only. Table 3.4 summarises the different terms given to shell nouns in both early and more recent studies, though in some cases the term may cover a wider or narrower range of nouns than shell nouns.

Recently, the term 'shell nouns' has gained more ground, and a broader set of studies have compared their use in different disciplines (Swales, 2005; Gray, 2010; Gray & Cortes, 2011); across languages (Moreno, 2004; Ribeira, 2007) and registers (Flowerdew, 2003a,b; Hirsh, 2010; Kanté, 2010; and Yamasaki, 2008); in learner writing in English by students from different language backgrounds (Flowerdew, 2009; Hasselgård, 2012); and in learner/professional writing (Flowerdew, 2003c; Atkas & Cortes 2008; Parkinson, 2013). In addition there are studies that compare sentence-initial anaphoric noun groups with extended or text reference (Swales, 2005, Gray,

2010; Gray & Cortes, 2011; Moreno, 2004; Álvarez-de-Mon y Rego). A brief review of the findings of these studies follows.

Shell nouns	Schmid (2000; 2001); Aktas & Cortes (2008); Gray & Cortes (2011); Gray (2010); Hasselgård (2012); Ribeira (2007); Hunston & Francis (2000)
Vocabulary 3	Winter (1977)
Unspecific metalanguage nouns	Winter (1992)
Labels (advance and retrospective)	Francis (1994); Tadros (1994); Moreno (2004); Charles (2003); Shaw (2000); Sinclair (1993)
Metalinguistic nouns	Francis (1994); Charles (2003; 2007); Moreno (2004)
Carrier nouns	Ivanič (1991)
Carrier words	Hirsh (2010)
Cognition nouns	Straker (2010)
Signalling nouns	Flowerdew (2003a; 2003b; 2009b)
Lexical signals	Hoey (1993)
Anaphoric nouns	Francis (1986); Bloor & Bloor (1995)
Unspecific anaphoric nouns	Yamasaki (2008)
Nouns taking complement that clauses	Kanté (2010); Parkinson (2013)
Summary nouns	Swales (2005); Bloor & Bloor (1995)
Status nouns	Hunston (2008; 2011)
Context inducing nouns	Price de Paiva & King (2008)
Textual nouns	Álvarez-de-Mon y Rego (2001)
<i>Rótulos</i>	Koch (2006); Bertucci (2006); Bezerra (2009); Martins (2009)
<i>Anáforas resumativas</i>	Figueiredo (2002)

Table 3.4: Alternative names for shell nouns

3.2.2. More recent studies on shell nouns

Unlike Francis (1994), whose corpus comprised complete editions of *The Times* newspaper from the Cobuild Bank of English corpus, most of the studies that involve shell nouns are concerned with academic vocabulary or academic writing. For example Hirsh (2010) classifies the 570 words of academic vocabulary from the NZ word list (Coxhead, 2000) according to Halliday and Matthiessen's metafunctions. Within the textual category, a subclass 'intratextual' is made up of shell nouns, which he calls carrier words, and conjunction. Hirsh compares the frequency of the academic vocabulary in a small corpus comprising academic textbook chapters, scientific articles and newspaper stories in the fields of chemistry, economics, law and sociology. In his study, he reports that academic textbooks use fewer carrier words/shell nouns than the scientific articles or the newspaper stories. He also finds that sociology and law are more likely to use academic words as shell nouns, and their use is more common in these fields than in economics.

His classification system resulted in some overlapping of categories. For example, *assumption*, which he identifies as scholarly process can in fact be functioning as a carrier word/shell noun, as in the example he gives:

3.7 In the regressions for *output* and inflation that follow, we assume that whether the exchange rate is pegged is a weakly exogenous *variable*. We test **this assumption*** using a Hausman test, adding the residuals from a Probit model or *regime* choice to the regressions. (Econ1) (Hirsh, 2010:162).

[* denotes a scholarly process (ideational) rather than carrier word (textual).]

He recognises this dual use, i.e. that carrier words can represent concepts at the ideational layer of the text (scholarly processes, states of affairs and relations between entities) as well as concepts at the textual layer (intratextual). Because Hirsh's analysis was restricted to items from the academic word list, his results for carrier words/shell nouns would underestimate the number of nouns that can be used in this way. However, they also suggest that some shell nouns - mental shell nouns like **assumption** - may be related to the field of research. Gledhill (2000), for example, finds shell nouns from the field of research – evidence – as well as factual, mental and linguistic shell nouns to frequently enter into the pattern *N be cl* in the introductions to research articles on cancer, with the clause that constitutes the lexicalisation of the shell noun construing a biochemical process. He also finds that the pattern *N-cl* is frequent in the introduction and discussion sections; however, while the former makes use of a wider range of shell nouns, the latter makes a greater use of the shell noun **fact**. He claims that the former is often evaluative in nature, either by means of modification or by means of the selection of shell noun, but considers the use of **fact** to be neutral, claiming that “nominal expressions of evaluation are more acceptable in Introductions than in Discussions”. This position seems untenable in a Systemic Functional Linguistics perspective, as will be argued in section 3.5.2. Gledhill also finds the shell noun **ability** is frequently used to project a clause construing a biochemical process. Again this suggests that exploration of the relation of shell nouns to field warrants further investigation.

Flowerdew (2003 a; b) also reports differences in frequencies of shell nouns across registers. Biology textbooks use more shell nouns than spoken lectures on the same topic. In addition he also finds differences in the nouns used: the spoken lectures

use **result** and **effect** less than the textbook or not at all. He attributes the differences to the fact that spoken lectures tend to express relations more congruently than written text. Like Francis (1994), Flowerdew (2003a) finds that anaphoric uses of shell nouns tend to coincide with Theme and Given information, while cataphoric uses tend to coincide with Rheme and New, and his findings corroborate Francis's (1994) claim that the shell nouns can function as context to introduce modification. While the nouns themselves may be neutral, evaluative modifiers such as *important* serve an interpersonal function by helping the listeners in the lectures to prioritise information (Flowerdew, 2003b). Evaluative modification of cataphoric shell nouns to help prioritise information has also been found in Brazilian history textbooks (Bezerra, 2009). Her analysis of anaphoric and cataphoric shell nouns across a range of history textbooks confirms the textual linking and semantic characterising functions (Schmid, 2000; Francis, 1994).

Comparing the use of shell nouns followed by a post-modifying *that* clause in legal transcriptions of felony trial reports and research articles, Kanté (2010) finds that there is a strong deontic relation between deontic nouns (e.g. intent, order) and the modal auxiliary in the *that* clause. He argues that the choice of noun reflects deliberate choice by the speaker and therefore indicates stance towards the projected clause. This view is in line with Charles (2003; 2007) and confirms Schmid's (2000) semantic characterising function.

The studies comparing the use of shell nouns in learner writing by university students and published or professional writing have found that learners use a more limited range of shell nouns than the professional writers (Aktas & Cortes, 2008; Parkinson, 2013) although the total frequency of nouns might vary little. It has also been found that the shell nouns are used in different ways (Flowerdew, 2003c). In her study of the realisations of the problem-solution pattern in recommendation reports, Flowerdew reports that the shell noun **problem** was used almost exclusively to denote causal relations (95%) by the professional writers whereas novice writers used it in that way only 32% of the time. Another study on learner writing in religious studies and history, which includes an analysis of English as Additional Language (EAL) university students' use of mental shell nouns over two years, reports that although

the students used more academic words from Coxhead's word list as their studies progressed, much of this was due to repetition rather than use of a wider variety of words, and the use of mental shell nouns was relatively infrequent (Straker, 2010). In particular it was found that students were not always able to use mental shell nouns successfully to refer anaphorically to another clause, and that some uses were both "redundant and arguably confusing" (Straker, 2010: 65). Hasselgård (2012) also finds differences in the ability to use some shell nouns appropriately in a study that compares advanced writers of English (university and upper secondary school) from different language backgrounds (Norwegian, French, German, and English). She argues that most of the difficulties that the Norwegian students had in using the shell nouns **idea** and **issue** are due to the fact that there are no semantic equivalents in Norwegian. By contrast, the French students had no problems with **idea** as a shell noun given its similarity to *l'idée de/que* in French. She finds that in general, with the exception of the shell noun *fact*, the Norwegian students underused shell nouns compared to the English. A similar result is found by Flowerdew (2009b), who reports that first year students from a Cantonese language background understood how to use shell nouns, although they did not always use them successfully. He finds a statistically significant correlation between writer competence and use of shell nouns: the more competent the writer, the greater the use of shell nouns.

A group of studies has compared the use of shell nouns, text nouns (Francis, 1994) and demonstratives in anaphoric reference. Anaphoric reference realised by a demonstrative pronoun has been called 'fuzzy' reference (Moreno, 2004) and 'unattended this' (Swales, 2005). Using Francis's (1994) distinction between metalinguistic nouns and non-metalinguistic nouns, Moreno (2004) analysed the metatext used to link a premise to a conclusion in English-medium and Spanish-medium research articles. She reports that the nominal groups with a noun as Head were more commonly used in the English-medium research articles than in the Spanish. The inverse was found to be the case for demonstrative reference, with extended or text reference⁴ accounting for 40.15% of the uses in Spanish and 16.73%

⁴ None of the studies on demonstrative reference reviewed in this chapter distinguish between text reference and extended reference (Halliday & Hasan, 1976: 52-53, 66-67). While many of the cases can

in English. In addition, she found that mental and linguistic shell nouns were used more frequently in the Spanish than in the English. Swales (2005) makes a similar study of sentence-initial demonstrative reference and anaphoric noun groups with noun as head using a corpus made up of research articles from 10 disciplines. Sentence-initial demonstrative reference made up between 25% to 42% of the referring expressions, with the exception of Dentistry, which showed higher values. In the anaphoric nominal groups, nouns dealing with research, methodology and results tended to be the most common. These findings indicate that in research articles reference to text by means of a nominal group with a noun as Head are an important means by which the writer moves the argument forward, and this corroborates previous findings (Francis, 1994; Shaw, 2000; Álvarez-de-Mon y Rego, 2001). Notwithstanding, it is difficult to determine the proportion of shell nouns used for this textual – linking function in these studies as the counts include head nouns that, although referring to stretches of discourse, are not operating as shell nouns, for instance, text nouns like ‘subsection’, found in Moreno (2004) and nominalised Processes that do not project, found in Swales (2005).

A similar problem is apparent in a study by Gray and Cortes (2011), who also compare the frequencies and lexical environments of *this* and *these* when used as a determiner in a nominal group or as demonstrative reference. Their corpus comprised research articles from Applied Linguistics and Materials and Civil Engineering. While they report that shell nouns account for 46% and 43%, respectively, of the uses, some of their analyses raise questions. For example, in 3.8 *this awareness* refers to a nominal group (underlined) rather than a clause, the head noun of which is a nominalised quality, *consciousness*. The relation between the two nominal groups could be seen as lexical cohesion (Halliday & Hasan, 1976), with *awareness*, also a nominalised quality, functioning as a synonym.

3.8 The next phase of research will consider ways in which teachers might best raise learner consciousness of the importance of theme in English information structure, and how **this awareness** may be activated to help learners produce fully coherent written discourse. (Gray & Cortes, 2011: 6)

be presumed to be text reference, it may not always be so. In this chapter, demonstrative reference will be taken to refer to either text reference or extended reference.

One of the difficulties of deciding whether *awareness* in example 3.8 is a shell noun or not stems from the fact that its nominal antecedent contains two nominalised qualities: *consciousness* and *importance*, both of which can function as shell nouns. A congruent reading could be ‘Theme is important in English information structure. Learners are/need to be conscious of this.’ In fact, the classification of anaphoric nominal groups with a potential shell noun as Head and whose antecedent is realised by a nominalisation, particularly a complex one with a Classifier and Qualifier such as in 3.8, remains obscure in the literature. Although Schmid (2000) gives an example of a shell nominal group whose Qualifier is a prepositional phrase rather than a clause (“a sad idea of a gay man”), he avoids specifically addressing this issue, and the treatment of such groups in subsequent studies is not always clear. For example, the shell noun group counts of Atkas and Cortes (2008) include antecedents realised by a nominalisation.

One study to specifically analyse the correlation between the kind of anaphoric referring item and its antecedent is Gray (2010). Using a corpus made up of education and sociology research articles, Gray classified sentence-initial anaphoric reference into five groups: pronouns; nominal groups with a concrete noun as head; nominal groups with an abstract noun as head; nominal groups with a shell noun as head; and nominal groups with other nouns as head. The antecedents of the referring expressions were classified into four types: simple NP (no post-nominal modification); complex NP (with post-nominal modification); sentence/clause (within sentence boundaries); or extended discourse (spanning sentence boundaries). Gray (2010) reports that nominal groups with a shell noun as Head make up 45% of the sentence-initial anaphoric noun groups, followed by nominal groups with an abstract noun as Head (43%). She also finds that the shell nominal groups tend to appear in relational clauses, and, as in Swales (2005), the most frequently used shell nouns are related to results and findings. Mental shell nouns and linguistic shell nouns make up a second major group. The shell nouns are not discipline specific, but are used similarly in both disciplines. Another finding is that while demonstrative reference refers overwhelmingly to a preceding sentence or clause, almost a third of the shell nouns refer to what Gray calls ‘extended discourse’, that is, information that spans more than

one sentence, just over a third referred to a clause, and a further quarter referred to complex a NP, that is, a nominal group with post-modification realised by a clause or a prepositional phrase, or apposition. Álvarez-de-Mon y Rego (2001) reports a similar trend for demonstrative reference to refer to the preceding sentence or clause and shell nominal groups to stretches of text spanning from a clause to a paragraph. Both Gray and Álvarez-de-Mon y Rego posit that the use of the shell noun enables the writer to alert the reader to the relevance of the information so that it can be commented on or developed in the Rheme.

A few studies do not use a corpus of academic discourse. Ribeira (2007) studies the use of anaphoric nominal groups with non-nominal antecedent-triggers in English and in Catalan, using a corpus made up of the first ten chapters of *Peter Pan*. He distinguishes between anaphoric reference, in which the antecedent is a noun group, and text deixis, in which there is no nominal antecedent but the trigger is an illocutionary act or the text itself. The latter could correspond to shell noun use. Ribeira (2007: 152) defines text deixis thus:

“a reference device which shares the referential properties of both deixis and anaphora. Like anaphora, a text-deictic expression has a textual antecedent-trigger; however, at the same time, it preserves the ability of space deixis to show the addressor’s position in relation to the referent, although this relationship is not physical, but metaphorical.”

Ribeira found that just over half the anaphoric nominal groups were examples of text deixis, that is, the head noun might be classed as a shell noun. He classified the relation between the text deixis expression and its antecedent trigger using Halliday and Hasan (1976) relations of lexical cohesion and reference: ellipsis, repetition, synonymy, hypernymy, general noun, comparative reference, which he calls associative anaphora, and deixis *am phantasma*. It is not quite clear how Ribeira actually did this, as he provides few examples other than those with a general noun, and the example of comparative reference that he does give (3.9) is not in fact text deixis because its antecedent trigger is a nominal group.

3.9 Once again *the stars_i* blew the window open, and *that smallest star_{j(<i)}* of all called out (Riberia, 2007: 162).

Nevertheless, his definition of a general noun can be taken to include shell nouns:

- a) The general noun is an abstract one, considering the degrees of abstractness expressed by entities other than first-order⁵.
- b) The antecedent-trigger is not an NP, but a complex discourse unit (Ribeira, 2007: 163).

Ribeira found that around half of the text deixis expressions use general nouns to refer to the antecedent trigger. He argues that the most important function of the text deixis expressions is “to categorise them [the non-nominal antecedent triggers] as nouns and therefore point them as topicalised in-focus objects of discourse” (Ribeira, 2007:161); in other words it is the act of naming the antecedent trigger and reconstruing it as an entity that brings it into focus. This represents the semantic, textual and cognitive functions of Schmid (2000). Ribeira classified the antecedent triggers of text deixis with a general noun as head into six classes: clause, sentence, text, illocutionary act, discourse knowledge and others. While the relative shares for clause, illocutionary act and others were similar in the two languages, English showed a higher proportion (almost double) of sentence and text antecedents. By contrast, Catalan had a larger share of discourse knowledge antecedents than English (43.5% and 34.6% respectively). Ribeira attributes the latter finding to the fact that the corpus was a narrative, and the translation in Catalan tended to use text deixis expressions to refer to narrative time, which English often construed by means of verbal tense. Demonstrative reference – i.e. extended or text reference – was also usually translated by means of text deixis.

Using the British National Corpus, Yamasaki (2008) compares four shell nouns (**shift**, **change**, **mistake** and **failure**) in two of the lexicogrammatical patterns identified by Schmid (2000): *Th-N* (i.e. shell noun group as Theme and Subject, Given) and *Th-be-N* (i.e. shell noun group as New in a relational Process clause). She argues that all shell nouns are inherently evaluative, and the latter lexicogrammatical pattern is more explicitly so than the former. She examines the pre-modification that co-occurs with each lexicogrammatical pattern and finds the shell noun is more likely to receive pre-

⁵ Ribeira (2007) classified the anaphoric demonstrative noun groups using Lyons (1977) and Dik’s (1997) classification of entities into orders. While first-order entities are discrete objects or individuals that are stable entities with a temporally or spatially bound existence, second-order entities refer to dynamic processes, third order entities to concepts and propositions, and fourth order entities to illocutions and perlocutions.

modification in the second pattern, i.e. when it appears as New, than in the former, although the attitudinally marked **mistake** and **failure** are less likely to receive evaluative pre-modifiers than **change** and **shift** in this pattern. Furthermore, the kind of pre-modifiers for **change** and **shift** tend to differ in the two lexicogrammatical environments. When composing Theme and Subject, the shell nouns tend to take Classifiers as pre-modifiers, but when they compose New information, the pre-modifiers tend to be evaluative, subjective Epithets. The difference can also be seen as correlating with a difference in Information structure: Given vs New. Yamasaki also studies the frequency with which anaphoric nominal groups with the shell noun **problem** as head appear in sentence-initial position or not across spoken and written subcorpora of the British National Corpus. She finds that the written corpus has higher frequencies of **problem** being used in sentence initial position than the spoken.

Shell nouns crop up in some studies as an analytical category, albeit in a restricted way, or as potential members of an analytical category. For example, in her study of thematic development in research articles Whittaker (1995) gives examples of shell nouns in Theme position in relational clauses, which correspond to one of the patterns identified by Schmid (2000): *Th-be-clause*. Godsen (2009), who also studied Theme but in referee reports on article submissions, included three domains that encompass mental and linguistic shell nouns. While mental shell nouns could be used to express a participant viewpoint or hypothesised and objectivised viewpoints, linguistic shell nouns were used in the discourse (textual) domain.

In the field of computational linguistics Price, de Paiva and King (2008) examine complement-taking nouns to determine the implied truth value of the clausal complement. Although their examples are handcrafted, one of the tests they used to identify complement-taking nouns parallels one of Schmid's (2000) defining patterns: "Bob's ...N... that Mary visited her mother reduced him to tears" (Price, de Paiva & King, 2008: 11). They identified 411 nouns that can take a finite clause (*that*) as complement and 173 that can take a non-finite clause (*to*) as complement. They classify the first group of nouns into those with an emotional implicature (e.g. grudge, happiness, problem); those with a factual implicature (e.g. absurdity, irony, miracle); those with an implicature of not true (e.g. lie, falsehood); and those with no implicature

(e.g. idea, option). The second group – those taking non-finite complements – were classed into ability nouns (e.g. ability, choice, heart); asset nouns (e.g. money, option, time); bravery nouns (e.g. audacity, chutzpah, cojones); chance nouns (e.g. chance, opportunity); effort nouns (e.g. initiative, trouble); and other nouns (e.g. aim, agreement, idea). It is to be noted that not all these nouns can be considered shell nouns. For example, not all the asset nouns do not meet Schmid's (2000) criteria for shell nouns. Further, many of the nouns such as heart, chutzpah or money, are used with empty lexical verbs and would similarly be excluded from being classified as shell nouns.

Shell nouns are apparent, albeit indirectly, in a study in the field of economics by Chevalier and Hudson (2001), who seek to determine whether economic agents in Theme position express intentionality. Their analysis includes nominalised forms of mental Processes, which they call psychological verbs, (belief, desire, expectation, and preference) in Theme position. In their frequency counts they exclude examples where the nominalised Process could be traced back to the author or a researcher as Senser. This step resulted in 21%, 78% and 5% of the nouns *belief*, *expectation* and *preference* being excluded. What this shows, but is not reflected on by the authors due to the aims of their study, is that a high proportion of some nouns construe the author(s) or a researcher, and not an economic agent, as Senser. It also suggests that the nouns *belief* and *preference* may construe nuclear relations, and thus construe the field of economics. While it is not possible to determine the proportion of these nouns that were being used as shell nouns, it is likely that some of them were. When traced back to a researcher, they would enter into an analysis of Engagement in Appraisal theory, as will be explained in 3.5.2.

3.2.3 Shell nouns in Portuguese

There is evidence that the concept of shell nouns is applicable to the Portuguese language. It has been argued that nouns that refer anaphorically to text – either to the propositional content or the illocutionary force, such as *ameaça* (threat), *preocupação* (worry/concern), *reflexão* (reflection) contribute to textual progression, establish new discourse referents as topics by condensing stretches of prior text, and enable the writer to evaluate propositions (Figueiredo, 2002; Koch, 2006; Bertucci,

2006; Raposo *et al.*, 2012). In Portugal an anaphorically referring noun with such uses has been called *anáfora resumativa* (summarising anaphoric noun) (Figueiredo, 2002; Raposo, *et al.*, 2012: 1709) or *anáfora conceptual* (conceptual anaphoric noun) (Raposo, *et al.*, 2012: 2185) while in Brazil, scholars have followed Francis (1994), calling it *rótulo* (label) (Koch, 2006; Bezerra, 2009; Martins, 2009; Martin & Silva, 2008) or *anáfora encapsuladora* (encapsulating anaphoric noun) (Bertucci, 2006). Despite the apparent equivalence to shell nouns, the concept of both *anáfora resumativa* and *rótulo* corresponds to the concept of ‘label’ as argued by Francis (1994) rather than shell noun (Schmid, 2000). This is because in the above-mentioned studies, like Francis, the authors include text nouns, i.e. metalinguistic objects such as ‘*as palavras*’ (the words) (Figuereido, 2002) or ‘sentence’ (Koch, 2006), neither of which have the ability to project a clause as Qualifier. Notwithstanding this difference, studies of Brazilian newspaper texts have found that labels are more frequently used in interviews about politics than in interviews about culture or sport (Martins, 2009), and that they carry out the three functions described above in editorials (Bertucci, 2006). In interviews anaphoric labels are typically introduced with a selective determiner (this) while cataphoric labels typically take the non-selective option ‘the’. Moreover, in Brazilian history textbooks there is a trend for cataphoric uses to include interpersonal modification of the label, which guides or shapes the student’s reading, highlighting what information is considered important, as also argued by Flowerdew (2003b) for English.

In addition to these anaphoric and cataphoric uses, the Gulbenkian grammar of mainland Portuguese includes a section on nouns that take clauses as complements (Raposo *et al.*, 2012: 1877-1886)⁶. Examples are given in 3.10 and 3.11.

3.10 *A proposta de que se avance com a greve*
(The proposal to go go ahead with the strike)

3.11 *A luta por que se fizesse justiça acabou por dar os seus frutos.*
(The fight for justice to be done brought results in the end.)

The nouns are divided into three classes depending on the kind of relation between the clausal complement and the noun.

⁶ The concept of nouns that take a clause as complement crops up in several places in the grammar and depending on the focus of the section, it is given various names. E.g. *complementos oracionais do nome* (p. 1052); *orações completivas de sujeito selecionada por nomes* (p. 1857).

In the first class, the complement clause specifies the noun. The class includes nouns such as *assunto* (matter/subject), *boato* (rumour), *fato* (fact), *hipótese* (hypothesis), *ideia* (idea), *maluquice* (craziness), *mentira* (lie), *notícia* (news), and *sorte* (luck). Membership of the class can be tested by an agnate structure:

3.12 *A proposta é que se avance com a greve.*
(The proposal is to go ahead with the strike.)

This agnate structure is, in fact, the second pattern that Schmid (2000) uses to define shell nouns: The **shell noun** *be* that + clause. A second criterion to determine whether a noun belongs to the class is that the noun can be replaced by a demonstrative pronoun (3.13) but the clause cannot (3.14):

3.13 *Votei contra isso de que se avance com a greve.*
(I voted against this to go ahead with the strike.)

3.14 **Ganhou a proposta disso.*
(*The proposal of this won.)

The final property of the first class is that the noun requires a definite determiner.

In the second class of nouns the relation between the noun and complement clause is agnate to that of verb and complement or adjective and complement. The relation can be tested by substituting demonstrative reference for the clause complement (3.15, 3.16).

3.15 *A luta por que se fizesse justiça acabou por dar os seus frutos.*
(The fight for justice to be done brought results in the end.)

3.16 *A luta por isso acabou por dar os seus frutos.*
(The fight for this brought results in the end.)

However, it is not possible to substitute the noun (3.17) or reconstrue the nominal group as subject and predicate in a copula clause (3.18).

3.17 **Isso de que se fizesse justiça acabou por dar os seus frutos*
(*This that the fight for justice to be done brought results in the end)

3.18 **A luta é de que se fizesse justiça*
(*The fight is that justice is to be done)

This class includes nouns like *certeza* (certainty), *confiança* (trust), *demonstração* (demonstration), *empenho* (effort), *estimativa* (estimate), *justificação* (justification), *pergunta* (question), *preferência* (preference), *prova* (proof), *fé* (faith), *medo* (fear),

vantagem (advantage) and *vontade* (willingness), many of which are nominalised qualities or Processes.

The third class of nouns consists of nouns whose complement ‘coincides’ with the noun. Nouns in this class have the same properties as the nouns in the first group, except that the determiner to the noun can be definite or indefinite. The class includes nominalised mental and verbal Processes such as *decisão* (decision), *desejo* (desire) or *promessa* (promise). The three classes of nouns clearly correspond to shell nouns.

The complement clause to these nouns can be finite or non-finite. In finite clause complements, the mode – indicative or subjunctive – depends on the relation of the clause to the noun (i.e. whether it complements or specifies), and then depending on that, on the noun itself, or the context. Non-finite infinitive complements to these nouns must be introduced with a preposition, e.g. *o meu interesse em/por desvendar o mistério* (my interest **in** unravelling the mystery). In the absence of another preposition, *de* (of) serves as the default preposition, e.g. *a minha promessa de comprar um presente para a Clara* (my promise of to buy a present for Clara) (Raposo *et al.* 2012: 1052). Note that the English translation does not always require a preposition. With finite clause complements, the use of *de* to introduce the clause is optional, e.g. *a sua declaração (de) que os impostos iam baixar* (His/her statement (of) that taxes would decrease) (Raposo *et al.* 2012: 1052).

The review of nouns that take complement clauses and anaphoric nouns above indicates that the concept of shell nouns is equally relevant for Portuguese as it is for English. It also suggests that their use in both languages might be similar.

The chapter so far has provided an overview of studies on shell nouns. It has shown that shell nouns perform textual, semantic and cognitive functions in discourse. It has also shown that although the concept of shell nouns is applicable to both English and Portuguese, the use of shell nouns varies across languages, registers, and writing competence. Moreover, shell nouns are not necessarily discipline specific. In particular, it has been shown that shell nouns are prevalent in academic writing, especially in highly valued texts such as research articles, which indicates that they provide a strategic resource for academics wishing to publish in English-medium journals. The overview has not been restricted to any one theoretical framework. I

shall now turn to how shell nouns are incorporated into a Systemic Functional Linguistics perspective.

3.3. A Systemic Functional Linguistic Perspective

A Systemic Functional Linguistic (SFL) perspective sees language as a semiotic system, simultaneously construing experience and enacting interpersonal relationships, both of which are enabled by the discourse itself (Halliday & Matthiessen, 2004). These functions correspond to the ideational, interpersonal and textual metafunctions of SFL, which are developed as systems of choices. It is through this trinocular perspective that any analysis using SFL will be developed. In SFL, however, there is no one category of nouns that corresponds to the class of shell nouns, as defined by Schmid (2000). The following section will provide an overview of SFL accounts that cover the use of shell nouns, and where useful, the discussion will contrast the SLF position with that of Schmid (2000).

3.3.1. Shell nouns as a class

Halliday and Matthiessen (2004) identify 3 vectors along which words are capable of functioning as Thing: 1. Countability – count/mass; 2. Animacy – conscious/non-conscious; and 3. Generality – general/particular. They claim that it is the location on each of the vectors that will suggest the functional potential of a noun or pronoun. In the case of shell nouns, shell nouns may be countable or mass; they are non-conscious and they probably tend more towards the generality than the particular. As a class, however, they cannot all be placed at the same position along each vector. While all shell nouns are abstract, some are countable (e.g. idea) and others are mass (e.g. knowledge). Some are extremely general (e.g. thing) while others are less so (e.g. mistake). Some are at one end of the spectrum of non-consciousness (e.g. fact) while others imply a conscious Senser who is absent (e.g. knowledge). It is to be noted that the three vectors do not correspond to Schmid's three functional properties.

A more useful framework to understand shell nouns is found in Halliday and Matthiessen's (1999:190) general system for nouns (figure 3.2 below). Within this system shell nouns would be classed as semiotic abstractions. Semiotic abstractions

are non-discrete, that is they are mass nouns. Halliday and Matthiessen claim that they typically function as Range in mental Processes and verbal Processes, or as possessed Attribute in relational Processes. As an unbounded semiotic substance, they may be qualified by projection. This latter feature is important, as will be argued in 3.3.2.

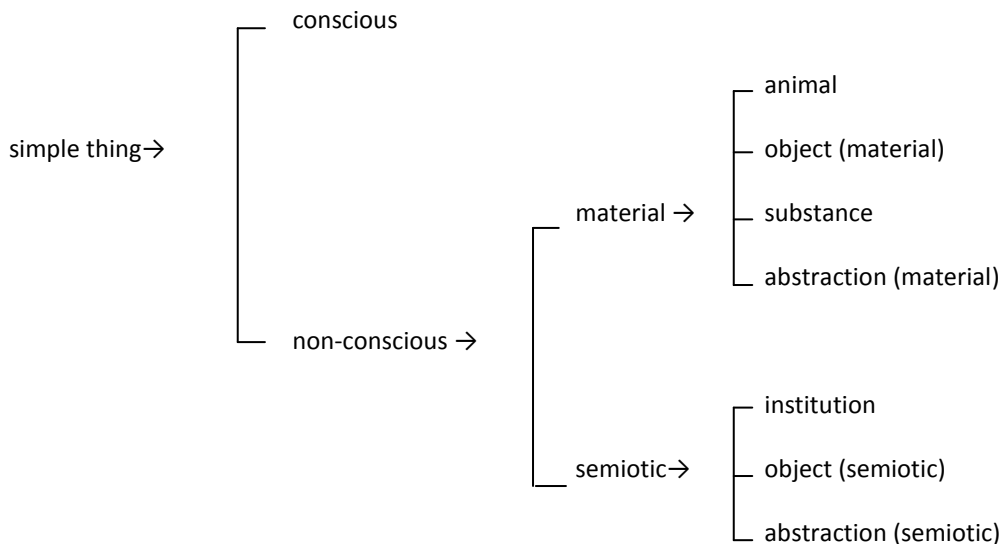


Figure 3.2: A general system for nouns (Halliday & Matthiessen, 1999: 190)

Within each category, Halliday and Matthiessen identify two additional subsets: pronouns and general nouns. The latter is of more interest to this thesis because many scholars have likened shell nouns to general nouns (e.g. Flowerdew, 2003c; Francis, 1994; Ribeira 2007). Schmid (2000: 6) rightly rejects the correspondence and demonstrates how some general nouns, for example *creature*, cannot be used as shell nouns because they cannot be modified with post-nominal clauses as complements. However, Halliday and Matthiessen's (1999: 190) general system of nouns (figure 3.2) accounts for the inclusion of some general nouns as shell nouns and the exclusion of others. Because general nouns are a sub-set in each category, it is the characteristics of the category that will determine the functional abilities of the general noun, not the other way round. According to Halliday and Matthiessen (1999: 189), general nouns "are used discursively to refer to instances of the category in question".

The categories of the system are not discrete, but blurred, and a noun may exhibit features of one category or another or be "equally at home in both" (Halliday & Matthiessen, 1999: 194). Halliday and Matthiessen identify three intermediate categories: natural forces, between animals and material objects; human collectives,

between conscious beings and institutions; and discrete semiotic abstractions, between semiotic objects and non-discrete semiotic abstractions. This latter category includes non-personalised 'facts' and 'cases' as well as mental entities such as 'beliefs' or 'ideas', all of which are typically used as shell nouns. In addition, it includes enhancing nouns such as 'way', 'time', 'place' and 'reason', which are also used as shell nouns. Nouns such as 'nuisance', 'mess', 'disaster', and 'shambles' are classed as *abstractions*, which means they could be either semiotic or material. Within this system, shell nouns would clearly be classed as semiotic abstractions, and they may be discrete or non-discrete.

The system also makes it possible to distinguish clearly between shell nouns and text nouns (Francis, 1994) such as *paragraph*, both of which Francis classed as metalinguistic nouns. In the framework, text nouns would be considered semiotic objects. The difference between them can be brought out in the following examples.

3.19 The study says that such a diversified village structure produces a dualistic pattern of migration (Halliday & Matthiessen, 2004: 254)

3.20 The regulations require that the following information must be conveyed to each subject: a statement that the study involves research ... (*answers.hhs.gov/ohrp/categories/1566*)

Both these clauses are verbal Process clauses. The Subject, conflated with Theme, is a semiotic object that construes the role of Sayer. However, only *regulations* can function as a semiotic abstraction if used in a certain colligational pattern, i.e. that of a shell noun such as Th- N *be* that. This is apparent in rewrites of the two examples where 3.19' is not possible but 3.20' is, even though there may be a change in meaning. The noun *regulation(s)* is thus "at home in both" categories; it can be used as a semiotic object and a semiotic abstraction.

3.19' *The study is that such a diversified village structure produces a dualistic pattern of migration.

3.20' The regulation is that the following information must be conveyed to each subject: a statement that the study involves research ...

Few SFL scholars have used Halliday and Matthiessen's (1999) general classification system for nouns in their analyses. Ravelli (2004) adapts part of it along a cline of abstraction moving from general noun through semiotic abstractions to

grammatical metaphors, realised as nominalisations (figure 3.3). Ravelli's interpretation suggests a relation between generic nouns, semiotic abstraction and grammatical metaphor. A similar relation has also been noted by Martin (1992: 376) and Hood (2008). Martin, for example, states that many of Winter's (1977) Vocabulary 3 items include metaphorical realisations of internal conjunctive relations, e.g. point, problem, reason, (consequence), result (consequence). He also notes that Carter's (1987:80) examples of Francis's A-nouns (anaphoric nouns) are all grammatical metaphors. Yet, while Ravelli classes shell nouns such as 'possibility' and 'variation' as grammatical metaphors rather than semiotic abstractions, both these nouns can be used as shell nouns and have the ability to project.

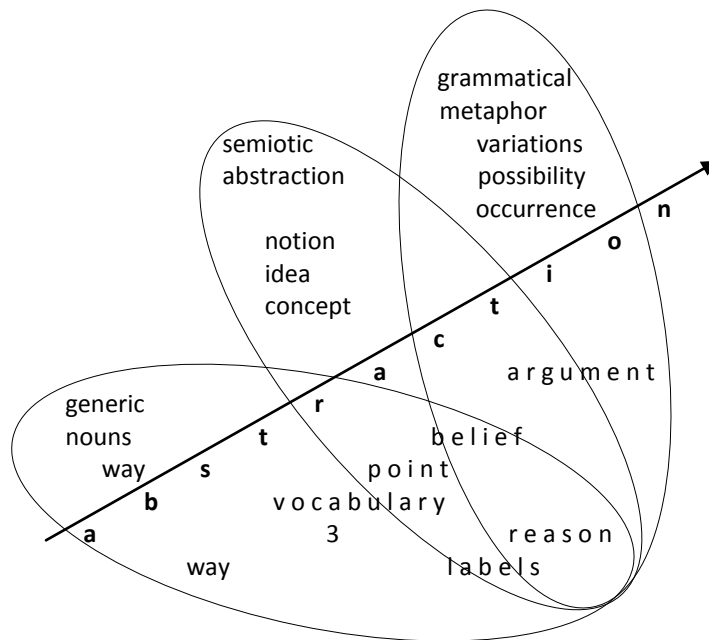


Figure 3.3 Organising vocabulary (Ravelli, 2004: 117)

Ravelli (2004: 122) posits that semiotic abstractions may be used “to mediate a grammatical metaphor which is analogous with, but not directly parallel to, a preceding figure or sequence”. Shell nouns certainly seem to fit this function, and because they remain unspecified unless the necessary propositional content is retrieved from the surrounding co-text, Ravelli's point is in keeping with Halliday's (1988/1993: 66) explanation of the drift towards nominalisation in scientific English, where one of the intermediary steps involves a fact clause. In Ravelli's model, shell

nouns would be classed as semiotic abstractions, some of them overlapping with general nouns (e.g. move) and others with grammatical metaphor (e.g. possibility). Although the overlapping categories in Ravelli’s model seem to account for the kind of difference between grammatical metaphor ‘proper’ (Ravelli, 2004: 117) and the use of shell nouns to construe a figure as an entity, the model mixes two distinct theoretical constructs: noun as a class, such as semiotic abstractions, or a sub-class, general nouns; and nouns resulting from the semogenic strategy of grammatical metaphor, e.g. an incongruent realisation of a Process, Quality or Relator. Given that Ravelli implies a distinction between grammatical metaphor ‘proper’ and semiotic abstractions, at this point a closer examination of the relation between grammatical metaphor and shell nouns is warranted.

As detailed in section 2.8, grammatical metaphor involves a shift in the realisational domain of a phenomenon. Of interest here is the semogenic process of nominalisation. Of the 13 types of elemental grammatical metaphor (see Table 2.5 p 28), five result in realisation as Thing. A summary of the types of nominalisation is reproduced here for convenience (table 3.5).

congruent:	metaphorical => thing
quality => <i>unstable</i>	1 <i>instability</i>
process => <i>absorb</i>	2 <i>absorption</i>
circumstance => <i>instead of;</i> <i>on the surface</i>	4 <i>replacement</i> <i>surface</i>
relator => <i>for/because [b,</i> <i>for/because a]</i> <i>so [a, so b]</i>	7 <i>cause,</i> <i>proof;</i> <i>result</i>
∅ =>	11 <i>phenomenon,</i> <i>fact</i>

Table 3.5 Grammatical metaphor: nominalisation (adapted from Halliday & Matthiessen, 1999: 245)

All the examples of the nominalisations reconstructed from the Relator (7) (cause, proof, result) and ∅ (11) (phenomenon, fact) can be used as shell nouns, and it is possible to find examples of shell nouns for all the others. For instance, the shell nouns **importance** and **tragedy** are metaphorical realisations of the qualities *important* and *tragic* respectively (1), **statement** and **belief** are metaphorical realisations of the

verbal and mental Processes *state* and *believe*, respectively (2), and **alternative** can be seen as a metaphorical realisation of the Circumstance *instead of* (7). Of note also is category number 11, which includes fact nouns, where there is no congruent realisation of the metaphor. The relations between fact nouns and shell nouns will be explored more fully in section 3.4.1.

In their explanation of grammatical metaphor, Halliday and Matthiessen (2004) provide several examples where the metaphorical wording is realised by shell nouns in the lexicogrammatical environments deemed essential for nouns to be functioning as shell nouns (*N-be-cl* and *N-cl*) (Schmid, 2000). The transfiguration of the grammatical metaphor of the two lexicogrammatical environments is explained here.

1. A sequence of figures realised by a projection clause nexus may be reconstrued as a figure in a relational clause. The Process is nominalised and the projected figure (hypotactic projected clause) is down-graded to an embedded fact clause. The Process *be* is added.

congruent	'(people)	most strongly	believe			that there is no ...'
	α				\rightarrow	' β
	clause: mental					clause: projected
	Senser	Manner: degree	Process			
	nom. gp.	adv. gp.	verbal gp.			
metaphorical	The	strongest	belief	of all	is	that there is no...
	clause: relational					
	Value				Process	Token
	nom. gp.					nom. gp.: clause
	Deictic	Epithet	Thing	Qualifier	Finite/event	Thing
	determiner	adjective	noun: nominalisation	prep. phrase	verb	Clause

Figure 3.4 Congruent and metaphorical wordings of *N-be-cl* (Halliday & Matthiessen, 2004: 639)

2. A sequence of figures realised by a projection clause nexus may be reconstrued as an element (nominal group). The Process is nominalised, and the projected figure is rankshifted to function as a Qualifier.

Congruent	(people)	claim	that inspection is about improving schools
	α		β
	clause: verbal		clause: projected
	Sayer	Process	
	nom. gp.	verbal gp.	
metaphorical	The	claim	that inspection is about improving schools
	nom. gp.		
	Deictic	Thing	Qualifier
	determiner	adjective	finite clause

Figure 3.5 Congruent and metaphorical wordings of N-cl

Halliday and Matthiessen (2004) also relate grammatical metaphor to the logico-semantic relation of expansion. Nominalisations of this kind include typical shell nouns such as *reason* and *cause*. This prompts the question of whether all shell nouns are grammatical metaphors. A reading of the index of shell nouns in Schmid (2000) shows that with remarkably few exceptions, they are all grammatical metaphors. The exceptions are circumstantial shell nouns (e.g. *area*, *age*, *era*), which Schmid recognises as being on the extreme fringe of shell nouns because they can be interpreted as derived from relative clauses (cf section 3.1.3, p 43)

Although it can be argued that except for some marginal, circumstantial shell nouns, all shell nouns are grammatical metaphors, it is not the case that all grammatical metaphors realised as nominalisations are shell nouns. This is no doubt what Ravelli (2004) meant when she distinguishes between grammatical metaphor and grammatical metaphor 'proper'. Nominalisations such as *displacements* and *failure* in 3.21 are not shell nouns although *failure* can function as one, as in 3.22, in which the shell noun is in bold and its lexicalisation is underlined. This is why Schmid defines the class of shell nouns as functional; it depends on the relation between the shell noun and its lexicalisation in a given text. From a Systemic Functional Linguistics perspective it can be argued that the meaning of the shell noun is instantiated in the text. Guillén Galve (1998) have called this feature intratextual dynamic grammatical metaphors.

3.21 Displacement along these faults caused failure of the Baldwin Hill Reservoir in 1963. (Halliday & Matthiessen, 2004: 643)

3.22 His remarks are being interpreted by many here as preparation for **his expected failure** to unseat Mr Patterson, Jamaica's first black leader. (Schmid, 2000:254).

Grammatical metaphors that are shell nouns can be ideationally oriented, e.g. *reason*, interpersonally oriented, e.g. *possibility* (Halliday & Matthiessen, 1999; 2004; Martin, 1992), or textually oriented (Martin, 1992), e.g. *point* as in *That point is just silly* (Martin, 1992: 406). The latter use has been referred to as ‘textual metaphor’ (Martin, 1992). Martin posits that such textual metaphors construe meta-message relations. As a result they add additional layers of meaning to a text that are best treated at the stratum of the discourse semantics (Martin, 1992).

Although Ravelli (2004) has noted the relation between shell nouns and grammatical metaphor, most other analyses that include shell nouns have focused on the experiential domain in the construal of technical taxonomies and their contribution to field. Halliday (1998) argues that technical taxonomies in scientific discourse are built on a relationship of generality, where a superordinate category is more general than its hyponyms, but, unlike folk or everyday taxonomies (domestic taxonomies, Martin 1992: 545), they also embody theoretical abstraction because “assigning a class to a larger, more general class is a theoretical operation” (Halliday, 1998: 198). Technical taxonomies thus distinguish between what belongs in the class from what does not, and, in doing so, recognise an entity’s “*place* in the taxonomy, and its *value* as a theoretical construct (Halliday, 1998: 198). Martin (1992: 545) proposes four grades of technicality - domestic, specialised, administration, and exploration – each of which construes field-related taxonomies. The fourth grade, exploration, which concerns instruction, is the most technical, and he argues that the “linguistically constructed taxonomies of humanities, social science and science are the ones that are most appropriately referred to as technical since they function as distillations of common sense or less technical experience into uncommon sense classifications of the world” (Martin, 1992: 545). Halliday (1998: 201) goes on to argue that “the terms created are not transient constructs that serve for one moment of discourse and disappear. They become part of a sub-system within the overall semantic space that constitutes the experiential domain of the grammar.”

This field-oriented view underlies many of the taxonomies for nouns used in SFL analyses. For example in their analysis of how the discourse of a personal exemplum, an argument by Desmond Tutu and an act of parliament construes experience, Martin

and Rose (2003) classify the nominal entities according to the following criteria (table 3.5).

Indefinite pronouns		<i>some/any/no thing/body/one</i>
Concrete	everyday	<i>man, girlfriend, face, hands, apple, house, hill</i>
	specialized	<i>matttock, lather, gearbox</i>
Abstract	technical	<i>inflation, metafunction, gene</i>
	institutional	<i>offence, hearing, applications, violation, amnesty</i>
	semiotic	<i>question, issue, letter, extract</i>
	generic	<i>colour, time, manner, way, kind, class, part, cause</i>
Metaphoric	process	<i>relationship, marriage, exposure, humiliation</i>
	quality	<i>justice, truth, integrity, bitterness, security</i>

Table 3.6 Kinds of entities (Martin & Rose, 2003:108)

They argue that the distinction between concrete and abstract ways of meaning reflects an important difference between the way everyday fields and the “uncommonsense” fields of social institutions and professions construe experience. They also note that semiotic entities are not field specific, and they are more frequently found in written discourse. These claims are in keeping with the findings that shell nouns are not discipline specific (Gray, 2010) and that written corpora show higher frequencies of shell nouns than spoken corpora (cf. Flowerdew, 2003a, b, Kanté, 2010). Nevertheless, while shell nouns are abstract nouns, they do not fit neatly into the class of semiotic entities; with the exception of abstract technical entities, it is possible to find examples of shell nouns for the abstract and metaphoric categories of the table.

The above discussion on technicality and the construction of taxonomies highlights a fundamental difference between shell nouns and technical terms: while technical terms aim to fix meanings permanently for a specialised audience, shell nouns do not. They rely on the listener/reader recovering the relevant information from the co-text of the noun to gain communicative effectiveness, and as such, their meaning is dependent on the surrounding discourse. In this sense they are closer to an instancial system (Matthiessen, 1995; Martin, 2006), which is built up over the course of the text.⁷ In addition, as shown by non SFL studies (cf Francis, 1994; Winter, 1977;

⁷ Matthiessen (1995:22) explains an instancial system as “created in the instantiation of the general system (general system potential) as a text unfolds; it is the product of logogenesis – the creation of meaning through instantiation of the system in text. From the speaker’s point of view, an instancial system is the system of selections s/he has to make in producing the text; from the listener’s point of view, an instancial system is the system that s/he can create out of the interpretation of the unfolding

Charles, 2003; 2007; Hoey, 1993; Schmid, 2000; Shaw, 2000), shell nouns tend to organise the discourse, rather than the field. This latter function will be developed in section 3.4.3.

The discussion on grammatical metaphor has shown that, with extremely few exceptions, shell nouns are grammatical metaphors and are strongly linked to projection. Because any embedded fact clause is a metaphorical realisation of a projected figure, even when there is no overt congruent realisation, following Schmid (2000), it thus follows that any noun appearing as Value in an identifying intensive relational clause whose Process is realised by the verb *be* and whose Token is realised by a fact clause is functioning as a shell noun. Similarly, any noun projecting a fact clause as Qualifier is functioning as a shell noun. These colligational patterns will be explored more fully in section 3.4.1. Seen from a different angle, it can be argued that because fact clauses belong not to the material realm but to the semiotic, i.e. they are on a higher level of abstraction than an ordinary thing or act (Halliday & Matthiessen, 2004: 205), nouns that can enter into the above colligational patterns must be semiotic as well, given that in identifying relational Process clauses the Token and Value must be of the same status – thing or fact (Halliday & Matthiessen, 2004: 301). For these reasons, the most insightful theoretical position within SLF classification systems for nouns is to class shell nouns as discrete or non-discrete semiotic abstractions whose full meaning derives from the instantial system as the text unfolds.

3.4. Shell nouns in the lexicogrammar

As a lexicogrammatical class there is no one category of nouns in Halliday and Matthiessen (2004) that corresponds to Schmid's shell nouns. Rather, they appear in several places. The class of general nouns includes some shell nouns e.g. *idea*, *thing*, *fact*, although, as argued in 3.3.1, some general nouns cannot be used as shell nouns e.g. *creature*. Shell nouns are also related to the lexico-semantic relations of projection and expansion. Nouns of projection include typical shell nouns such as *belief*, *suggestion* or *idea*, as well as fact nouns such as *fact*, *case*, *chance*, *confirmation*, or

text.” He goes on to say: “We can think of an instantial ideational system as the ‘knowledge’ created in a text; or more accurately, an instantial ideational system is the configuration of meanings created in a text that would be characterized as ‘knowledge’ in a cognitivist framework.” (Matthiessen, 1995:23)

problem. Enhancing nouns that can fulfil the function of shell nouns include *reason*, *purpose* or *result*, among others. The ensuing discussion will focus on how shell nouns function within the logico-semantic relations of expansion and projection at the stratum of the lexicogrammar.

3.4.1. Shell nouns, projection and expansion

As previously detailed in section 3.1.2, Schmid (2000) identifies shell nouns on the basis of certain lexicogrammatical patterns. He argues that in order to be classified as a shell noun, the noun must be able to appear in some, but not necessarily all, of the patterns. One of the distinguishing lexicogrammatical environments – *N-cl* – is given in example 3.23. The shell noun is in bold and its lexicalisation is underlined.

3.23 Mr Bush said Iraq's leaders had to face **the fact** that the rest of the world was against them. (Schmid, 2000:22)

In this example the shell noun is followed by a post-nominal clause. Schmid discusses the relation between 'that' clauses and infinitive clauses in terms of complementation and post-modification. He argues that "a strict general separation between post-modifiers and complements is impossible if one considers the whole range of possible sequences of abstract nouns followed by clauses" (Schmid, 2000:24). For that reason he uses the term 'post-nominal clause', which he considers neutral. For example, in his preliminary discussion (chapter 3), he provides the following examples of nouns with a post nominal clause. Again the shell noun is in bold and the lexicalisation underlined.

3.24 Mr Bush said Iraq's leaders had to face **the fact** that the rest of the world was against them.

3.25 But it is **a good idea** to stop and think about it.

3.26 Agnelli is due to step down soon as head of Fiat and already there is **speculation** that he might move into politics.

3.27 Britain is **a great place** to live and work in.

3.28 The lists of Indonesian communists were compiled by the CIA and State Department over two years, at **a time** when Washington believed that there was a real threat that southeast Asia would fall under communist rule.

3.29 Is there **any place** where you can go and play snooker or anything like that?

While he considers that 3.24 is an archetypical example of a noun functioning as a shell noun, he notes that 3.25 is a related pattern in which "the clauses

representing the shell contents ... make up the notional subjects of the clauses, which are moved to the end of the sentences” (Schmid, 2000:24), and that for many shell nouns this postponed pattern is the preferred one. Examples 3.25 – 3.28 he considers to be more marginal. In 3.26, he argues that “indefinite noun phrases do not create as strong conceptual boundaries as the definite noun phrases in which shell nouns tend to occur” (25). He contrasts the use of *place* in 3.27 with modal nouns such as *ability*, *chance*, *need*, or *willingness*. While a post-nominal clause acts as a complement to the modal nouns, with temporal or locative nouns a post-nominal clause tends to be seen as a variant of a relative clause. Examples 3.28 and 3.29 represent “a transition zone between complementing or appositive relations as adverbial post-nominal clauses” (26), which also tend to be treated as relative clauses.⁸

For Halliday and Matthiessen (2004), examples 3.24 – 3.26 are clear examples of projection. Projection refers to the logico-semantic relation “whereby a clause comes to function not as a direct representation of (non linguistic) experience but as a representation of a (linguistic) representation” (Halliday & Matthiessen, 2004: 441). A projection is a metaphephenomenon (Halliday & Matthiessen, 2004: 441). Both mental and verbal Processes can project, and their nominal equivalents, which can also project, are classed as projection nouns. As pointed out in the discussion on grammatical metaphor in section 3.3.1 (cf figure 3.5, p. 71), in a nominal group with a projection noun as Head, the projected element can be rankshifted, i.e. embedded, to act as Qualifier, and this projected clause restricts the meaning of the projection noun in the same way that a ‘restrictive’ relative clause defines a noun. This corresponds to Schmid’s pattern *N-cl*.

Example 3.26 involves a projection noun, *speculation*, which is qualified by the projected element *[[that he might move into politics]]*. The nominal group construes Existent in an existential clause. Halliday and Matthiessen specifically note the benefits of nominal groups with embedded projections; as a nominal group it can be used in grammatical environments in which a non-embedded tactically related projected

⁸ Schmid excludes the pattern *N of V-ing*, e.g. *the problem of raising money*, due to the difficulties in retrieving examples from the corpus; however, he recognizes that it appears with a restricted number of nouns.

clause cannot. An example of a projection noun with Qualifier construing Actor in a material Process is given in 3.30. The projection noun **knowledge** (in bold) corresponds to the shell noun and the embedded clause as Qualifier corresponds to the lexicalisation of the shell noun (underlined).

3.30 The **knowledge** that the experiment had failed destroyed his life. (Halliday & Matthiessen, 2004: 441)

In addition, the use of a projection noun, as in example 3.26 (There is speculation that ...), makes it possible to omit the Senser or Sayer. Halliday and Matthiessen highlight the importance of projection nouns in scientific writing where, for example, “proposals and demands are opposed, renounced and rejected” (Halliday & Matthiessen, 2004: 468).

Projection nouns are classified into those that project propositions, either stating or questioning, and those that project proposals, either offering or commanding. In the first group the projected element is realised by a finite clause and in the second it is realised by a non-finite clause. The class of projection nouns is illustrated in table 3.7, below.

3.31 Mr Bush said Iraq’s leaders had to face **the fact** that the rest of the world was against them.

Example 3.24, reproduced here as 3.31 involves a fact noun – **fact** – with a fact clause as Qualifier. A fact clause is also a projection, but an impersonal one in which the clause cannot be traced back to a projecting Process (mental or verbal); the fact clause is “ready made” (Halliday & Matthiessen, 2004: 470). Halliday and Matthiessen identify four classes of fact nouns: ‘cases’, which relate to non-modalised propositions; ‘chances’, which relate to modalised propositions; ‘proofs’, which relate to caused modalised propositions; and ‘needs’, which relate to proposals. While in 3.31 the fact clause construes Qualifier to a fact noun, **fact**, a fact clause may also stand alone as Thing in a nominal group. Whether standing alone as Thing or functioning as Qualifier for a fact noun, as in 3.31, a fact clause is always rankshifted, i.e. embedded. Table 3.7 shows Halliday and Matthiessen’s (2004) classification of projection and fact nouns. It

is to be emphasised that all of these nouns can be used as shell nouns, and all are grammatical metaphors.

			Projection nouns	Fact nouns
Propositions	stating	locutions	statement, report, news, rumour	(1) 'cases' (nouns of simple fact): fact case, point, rule (2) 'chances' (nouns of modalization): chance, possibility, certainty (3) 'proofs' (nouns of indication – caused modalization): proof, indication, implication, confirmation, evidence
		ideas	thought, view, belief, knowledge, feeling	
	questioning	locutions	question, query, argument, dispute	
		ideas	doubt, question	
Proposals	offering	locutions	offer, suggestion, proposal, threat	
		ideas	intention, desire, hope, inclination	
	commanding	locutions	order, command, instruction	(4) 'needs' (nouns of modulation): requirement, need, rule, obligation, duty
		ideas	wish, desire, hope, fear	

Table 3.7: Nouns of projection and nouns of fact (adapted from Halliday and Matthiessen, 2004: 469)

3.32 But it is **a good idea** to stop and think about it.

Example 3.25, reproduced here as 3.32, also involves a fact clause – *to stop and think about it*. In this case the fact clause construes Carrier in an impersonal attributive intensive Process clause, in which the Subject is post-posed, and Attribute is realised by the nominal group *a good idea*. Halliday and Matthiessen (2004: 472-475) discuss the nouns that typically favour this lexicogrammatical environment. While fact nouns can and do appear in this environment, evaluation nouns such as *pity*, *tragedy* or *shame* are particularly prevalent although this latter group is less likely to function as Thing with a fact clause as Qualifier. The interpersonal and evaluative nature of this lexicogrammatical environment is evident from the kinds of nouns that function in it, e.g. possibility, pity, surprise, and highlighted by the inclusion of Epithets with the fact nouns in Halliday and Matthiessen's original table. The interpersonal aspect corresponds to Schmid's (2000) characterising function. All of the evaluation nouns listed belong to the semiotic domain of attribution (Halliday & Matthiessen, 2004: 223). Like projection nouns and fact nouns, all of the evaluation nouns discussed in Halliday and Matthiessen can be used as shell nouns. These nouns are listed in table

3.8. It does not purport to be an exhaustive list, and relevant examples of nouns that function as shell nouns have been added to categories that were blank in the original table.

		Noun
Proposition	cognition	<i>doubt</i>
	probability	possibility, likelihood, certainty, coincidence
	usuality	<i>trend</i> <i>secret</i>
	emotion	pity, shame, relief, tragedy, surprise, regret nuisance, <i>mistake</i> , <i>inconvenience</i>
Proposal	desire	requirement
	obligation	necessity, rule, principle, law

Table 3.8: Evaluation nouns appearing in *it is* Attribute: *noun that* (adapted from Halliday & Matthiessen (2004: 475), added nouns in italics)

The relation of the nouns in table 3.8 to shell nouns is grounded in the fact that they are all grammatical metaphors of qualities of projection. Halliday and Matthiessen (1999: 209-212) propose a distinction between qualities of projection and qualities of expansion, where qualities of projection are agnate with figures of sensing (i.e. mental Process clauses) and qualities of expansion are agnate with figures of being (i.e. relational Process clauses). Thus, qualities of projection can be assigned to metathings while qualities of expansion normally cannot. In addition to evaluative qualities such as *important* or *tragic*, qualities of projection include modalised qualities such as *possible* or *likely* and modulated qualities such as *necessary* or *desirable*. When these qualities undergo the semogenic process of grammatical metaphor and are nominalised, the result is a shell noun.

The last three examples of post-nominal clauses involve the logico-semantic relation of expansion, and not projection. They are reproduced here for convenience as 3.33 - 3.35.

3.33 Britain is **a great place** to live and work in.

3.34 The lists of Indonesian communists were compiled by the CIA and State Department over two years, at **a time** when Washington believed that there was a real threat that southeast Asia would fall under communist rule.

3.35 Is there **any place** where you can go and play snooker or anything like that?

Halliday and Matthiessen (2004: 435-437) discuss expansion nouns in which the circumstantial relation is construed in the head noun, rather than in the enhancing embedded clause. They class them into two groups: those that can take a finite or non-

finite embedded clause and those that can take only non-finite clauses, as shown in table 3.9.

Category of expansion		finite [[why, etc./ for, etc which/(that)]]	non-finite [[of doing]]
enhancement	time	time, day, occasion	
	place	place	
	manner	way	
	cause: reason	reason	
	cause: purpose		purpose, point, aim
	cause: result		result

Table 3.9: Nouns of expansion (Halliday & Matthiessen, 2004:435)

The table could be extended to include some other nominalisations of qualities of expansion (Halliday & Matthiessen, 1999: 211), which can also function as shell nouns. However, instead of enhancing, these nouns extend or elaborate. These nouns are listed in table 3.10.

Category of expansion		
elaboration	identity	similarity, analogy, difference
extension		alternative, contrast

Table 3.10 Other nouns of expansion that function as shell nouns

As already explained, Schmid considers circumstantial shell nouns to be marginal due to their relation with relative clauses. Halliday and Matthiessen also note the similarity between enhancing embedded clauses with a noun of expansion as Head and elaborating embedded clauses (2004: 435-436) while noting their similarity to the logico-semantic relation of projection (Halliday & Matthiessen, 1999: 295). Moreover, they point out that expansion nouns often appear as Value in identifying clauses. One of the examples they give is 3.36, below, and this example illustrates the second lexicogrammatical pattern that Schmid uses to distinguish the function of shell nouns: intensive identifying relational clauses with shell noun as Value and an embedded fact clause as Token.

3.4.2. Shell nouns as Value in intensive identifying relational Process clauses

3.36 Another **reason** is that the quantity of the literature is not overwhelming yet. (Halliday & Matthiessen, 2004: 437)

In example 3.36 the expansion noun construes Value and the Token is realised by a fact clause. In the pattern identified by Schmid (2000), *Th-N be cl*, the relational

clause is encoding; Value, realised by a shell noun, is encoded by reference to the fact clause that construes Token, which is also the lexicalisation of the shell noun (figure 3.5). Halliday and Matthiessen argue that such a structure enables information to be distributed so as to create a textual impact, and this can be strategically used by writers. Schmid (2000) makes a similar point, calling it topicalising (c.f. section 3.1.4).

Another reason is that the quantity of the literature is not overwhelming yet		
Value / Identified	Pro. Rel	Token / Identifier
Subject		Complement
Shell noun		Shell content or lexicalisation

Figure 3.6 Shell noun as Value / Identified in an encoding identifying clause

Expansion nouns are obviously not the only nouns that can construe Value in encoding intensive identifying relational Process clauses. Clearly projection nouns, fact nouns and evaluation nouns can do so as well. Nor is the status of the embedded clause realising the Token restricted to fact clauses: act clauses (cf. Halliday & Matthiessen, 2004: 204, 436-449) can also appear in this grammatical environment, as shown in example 3.37, in which the shell noun is in bold and the lexicalisation underlined.

3.37 The **challenge** is (for him) to revise the manuscript. (Halliday & Matthiessen, 2004: 440)

In this example ‘for him to revise the manuscript’ is an act clause, i.e. a clause that represents not a meta-phenomenon but a macro-phenomenon, which denotes the name of an action, event or other phenomenon and is realised by a non-finite clause. Halliday and Matthiessen identify some nouns that favour the grammatical environment of Value, with an act clause as Token. The Value can be realised by a noun of expansion or a noun such as *challenge*, *difficulty* or *task*. Once again, all these nouns can function as shell nouns, and in general they correspond to Schmid’s (2000) category of eventive shell nouns, the logico-semantic relation of the embedded act clause being that of expansion and not projection.

Although the lexicogrammatical environments of projection and expansion differ, the semantic distinctions become blurred under certain conditions. In other

words, expansion nouns with a non-finite Qualifier approach the borderline of projection (Halliday & Matthiessen, 1999: 295), and projection nouns and fact nouns with Qualifiers can enter into Processes other than mental or verbal Processes, e.g. material Processes, where a fact clause could not. This is due to the spillovers of grammatical metaphor. Halliday and Matthiessen (2004: 441) put it thus: “[Expansion and projection] come together under conditions of nominalization, where there is metaphor in the grammar and many of the semantic distinctions expressed in the clause tend to be neutralized”.

But what of shell nouns that enact Subject in identifying clauses where the relational Process is not realised by the verb ‘be’? As noted in section 2.8, verbs such as *show*, *prove*, *indicate*, and *suggest* are an important resource in academic writing as they contribute to structuring the argument of the writer. An example is given in 3.38.

3.38 The **evidence** that I have seen from laboratory studies indicates that liquid sulphuring acid particles will not provide such an efficient surface for heterogeneous chemistry, partly because the rate of reaction proceeds more slowly compared to that with ice crystals, and partly because the typical density of the sulphuric acid aerosols is less than that for ice crystals over Antarctica. (Halliday & Matthiessen, 2004:235)

As in 3.36, in 3.37 the nominal group with the shell noun **evidence** as Head enacts Subject, and Complement is realised by an embedded projected clause. However, this is where the similarity ends. While in 3.36 the identifying clause was encoding and the Token represented the lexicalisation of the shell noun, in 3.38 the identifying clause is decoding and the embedded projected clause does not represent the lexicalisation of the shell noun. Rather, the information that is encapsulated in ‘The evidence that I have seen from laboratory studies’ must be retrieved from elsewhere in the text. The nominal group with shell noun as Head construes the Identified and the projected clause construes the Identifier (figure 3.6). Consequently, relational Process clauses of this type – decoding – are not examples of Schmid’s second pattern for identifying shell nouns.

The **evidence** that I have seen from laboratory studies indicates that liquid sulphuring acid particles will not provide such an efficient surface for heterogeneous chemistry, partly because the rate of reaction proceeds more slowly compared to that with ice crystals, and partly because the typical density of the sulphuric acid aerosols is less than that for ice crystals over Antarctica.

Token / Identified	Relational Process	Value / Identifier
Subject		Complement
Shell noun		

Figure 3.7 Shell noun as Token / Identified in a decoding identifying clause

It is, nonetheless, possible to rewrite the clause, decomposing the Process and introducing a shell noun as Head of an embedded clause in the Complement (3.38'). This construal can be seen as "a metaphorical expression of the Process of a 'mental' clause" (Halliday & Matthiessen, 2004: 226). Notwithstanding, in this case the clause becomes attributive, not identifying.

3.38' The **evidence** that I have seen from laboratory studies is an **indication** that liquid sulphuring acid particles will not provide such an efficient surface for heterogeneous chemistry, partly because the rate of reaction proceeds more slowly compared to that with ice crystals, and partly because the typical density of the sulphuric acid aerosols is less than that for ice crystals over Antarctica.

The previous two points have shown how shell nouns can be explained at the stratum of the lexicogrammar using the logical semantic relations of projection and expansion. While projection nouns and fact nouns account for the majority of shell nouns due to their ability to be modified by a projected clause, evaluation nouns and expansion nouns can also construe Value/Identified in decoding intensive identifying relational clauses with an embedded clause as Token/Identifier, and in doing so, function as shell nouns. Furthermore, nouns of expansion and some other nouns can function as shell nouns in similar grammatical environments with act clauses. The examples of expansion nouns which function as shell nouns given above can be considered at the border of projection. The other two grammatical environments that are favoured by shell nouns involve the resources of cohesion. This will be dealt with in the following sections.

3.4.3. Shell nouns and reference

The use of nouns of projection, fact nouns and expansion nouns, i.e. shell nouns, in a nominal group to refer anaphorically to a clause is clearly related to the

textual metafunction and demonstrative reference. An example is given in 3.39. The shell noun is in bold and its lexicalisation is underlined.

3.39 The Labour Party opposed Thor missiles, because, he said, they were out of date and vulnerable and would attract enemy action. **That argument** did not apply to the Polaris submarine (Halliday & Matthiessen, 2004:468)

In 3.39 the nominal group with shell noun as Head construes the Carrier in an attributive relational Process clause, and it both composes Theme and enacts Subject. *That argument* refers to the preceding projected clause nexus *they [the Thor missiles] were out of date and vulnerable and would attract enemy action*. Halliday and Matthiessen note the similarity of the cohesive effect between this use of projection nouns and text reference by means of *this*, *that* or *it* (Halliday & Matthiessen, 2004: 468), and Martin and Rose (2003: 164,165) even use a shell noun which refers anaphorically to several clauses an example of text reference. In addition to its cohesive function, the noun concomitantly functions to classify the projected information (Halliday & Matthiessen, 1999: 203). It is this classification of the projected clause that can be seen as contributing to the instantial ideational system of the text (Matthiessen, 1995), and it corresponds to Schmid's (2000) semantic, characterising function of shell nouns.

According to Frances (1994), an anaphorically-referring shell noun is always presented as given information. Hence, part of the means of recovering the information that the shell noun encapsulates – i.e. the lexicalisation – is realised through the Deictic. The system for deixis within the nominal group is shown in figure 3.7, below.

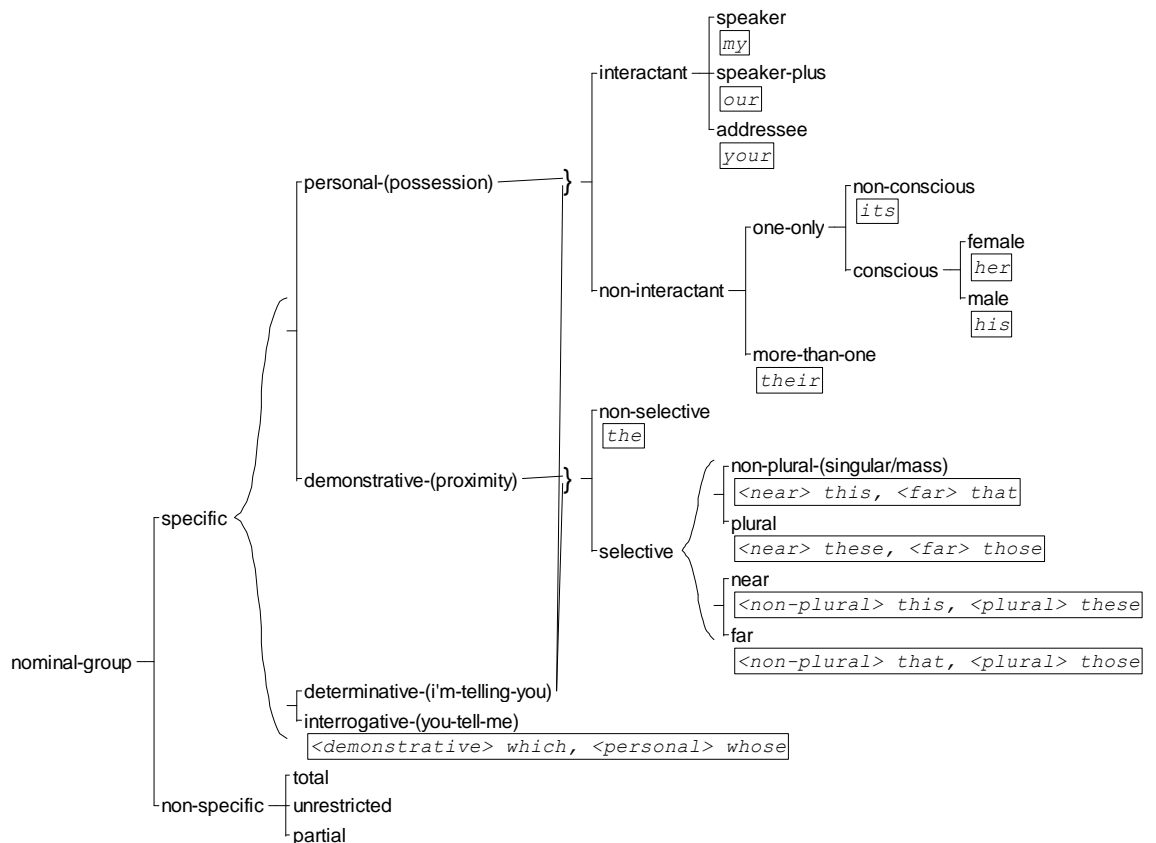


Figure 3.8: The system of DETERMINATION (Halliday & Matthiessen, 2004: 313)

Of relevance here are the demonstratives of proximity, in particular the selective Deictics - *this*, *these*, *that*, *those*. *This* and *that* can also function as demonstrative reference items, construing Head/Thing in a nominal group. The function of the specific Deictic in the nominal group is to identify the Thing as information known or treated as recoverable for the listener or reader. As argued in section 3.1.1, the nominal group with the shell noun as Head might be seen to relate back to the referent by means of identity of reference, in this case lexical cohesion (Halliday & Hasan, 1976:63). Lexical cohesion works by means of *reiteration*. Reiteration can be interpreted as relexicalising the referent along a cline from a) the same word, through b) a synonym or near-synonym, c) a superordinate, to d) a general word (Halliday & Hasan, 1976:279). However, it is hard to place shell nouns on such a cline. While they are inherently unspecific due the semantic gap identified by Schmid (2000), they do carry experiential meaning that interprets the presumed clause. Although reference is considered to be “overwhelmingly nominal in character” (Halliday & Hasan, 1976: 43), Martin (1992: 309) provides a test for determining when

non-nominal lexical items such as Processes or qualities are cohesive⁹, and it might be applied to shell nouns as well. He claims that lexical items “can be treated as such [cohesive] where one item codes sufficient experiential meaning to be presumed by another, either directly, or indirectly through bridging”. Bridging depends on “experiential connections between presuming and presumed which facilitate the recovery of an implied identity” (Martin, 1992:124), for instance when the identity of something must be established with respect to one of its parts through the relation of meronymy or when a hyponym is used to presume a superordinate. Such experiential connections are not restricted to relations between nominal groups. This argument comes close to Schmid’s relation of “experiential identity”, as detailed in section 3.1.1.

Martin (1992), himself, does not treat shell nouns at the stratum of the lexicogrammar. Rather, he argues that shell nouns are metaphorical realisations of internal conjunctive relations or they function as meta-text (Martin, 1992: 409, 416-417). In fact, this is the typical treatment given to shell nouns, and it will be discussed in 3.5.1. However, Thompson and Zhou (2004: 123) argue that there is no absolute division between reference and conjunction given that “on the one hand the relationship between an unspecific noun and its specific lexicalization is in the broadest sense one of repetition [i.e. reference], and on the other hand repetition may in itself indicate the type of conjunction.”

The use of shell nouns to refer anaphorically to a preceding clause or stretch of discourse enables the writer to move the argument on (Halliday & Matthiessen, 1999; 2004). Repackaging the information previously realised as a clause in a nominal group with an anaphoric shell noun as Head allows the writer to take full advantage of the textual resources of the Theme to build up instantial ideational systems (Matthiessen, 1995), thus contributing to the logogenesis of the text. Furthermore, as a nominal group, the information can construe transitivity roles that would not be possible otherwise, and the opportunity for modification is opened up.

⁹ Martin suggests nominalising the processes or qualities to see whether they are experientially close.

3.4.4. Shell nouns and text reference

The final lexicogrammatical pattern that Schmid (2000) identifies as being common for shell nouns is exemplified in 3.40.

3.40 The talks lasted for three hours. This was a **surprise**, for they had only been scheduled to last two hours. (Halliday & Matthiessen, 2004: 468)

In this pattern, the shell noun realises Attribute in an intensive relational Process clause and the Carrier is realised by text reference, a demonstrative pronoun. The shelled content or lexicalisation includes both the text reference, *this*, and the preceding clause to which *this* refers.

In text reference the demonstrative pronoun refers to information in the semiotic domain; i.e. it refers to text as fact or report (Halliday & Hasan, 1976: 52). Although the demonstrative pronoun *this* typically refers to the preceding clause, it may also refer to extended passages of text. Moreover, text reference can be cataphoric, as shown by 3.41, though such use is rarer than anaphoric reference.

3.41 In brief, **the soon widely held assumption** was this: man could understand the universe because it was natural and he was rational. Moreover, he might be able to control, even reorder his environment, once he had knowledge of it. (Halliday & Matthiessen, 2004:552)

Once again, the relation between the shell noun and text reference is construed by means of an intensive relational Process, but in this case, it is identifying and not attributive. The shell noun **assumption** construes Value, and the demonstrative text reference, *this*, Token in a Thematic equative clause (Halliday & Matthiessen, 2004: 68-71). The text reference points cataphorically to the following clauses, which become highlighted as more prominent in newsworthiness. This particular example can be seen as related to the second lexicogrammatical pattern described by Schmid (cf. section 3.4.2), where the shell noun construes Value in a decoding identifying relational Process clause and the Token is realised by an embedded fact clause. The relation becomes apparent in a rewrite of 3.41; however, it is less clear whether the stretch of discourse to be encapsulated in the shell noun extends beyond the embedded clauses or not.

3.41' In brief, **the soon widely held assumption** was that man could understand the universe because it was natural and he was rational. Moreover, he might be able to control, even reorder his environment, once he had knowledge of it.

One of the advantages, then, of using text reference in an intensive relational Process clause with a shell noun is that the text reference can be used to refer to extended passages of discourse, and the relation between the shell noun and the text reference can be construed succinctly by means of an intensive relational Process.

The two examples discussed here, 3.40 and 3.41, illustrate the textual benefits of these structures: the writer is able to select which information is to receive prominence. In 3.40 the evaluation noun **surprise** composes New, so it is this evaluative aspect that is treated as newsworthy. In 3.41 it is the clauses making up the content of the shell noun **assumption** that receive greater prominence, and the nominal group with shell noun as Head provides the point of departure. As Theme and Given, it is less likely to be disputed. Furthermore, as noted by Francis (1994), the modification *soon widely held* is slipped in, and similarly is less likely to be questioned.

3.4.5. Summary of shell nouns at the stratum of the lexicogrammar

This section has reviewed some typical environments in which shell nouns function in the lexicogrammar. It has been found that projection nouns, fact nouns, expansion nouns and evaluation nouns can function as shell nouns. As semiotic abstractions, they are able to be qualified with projections and/or enter as Value into decoding intensive relational Process clauses with fact clauses in the semiotic domain as Token, or in some cases, with act clauses as Token. These two features are what unite the nouns as a class. The ability of shell nouns/semiotic abstractions to feature in these lexicogrammatical environments derives from their status of grammatical metaphor, nominalisations of projecting Processes, projecting qualities, facts or some enhancing relations.

In intensive identifying Process clauses the nominal group with shell noun as Head can construe Token or Value, while in attributive clauses, it can construe Carrier or Attribute. When nominal groups with a shell noun as Head construe Value or Attribute, their semantic function is to characterise (Schmid, 2000) the shell content. Notwithstanding their preference for relational Process clauses, shell nouns can also

enter into other Process types, which a fact clause without a noun as Head could not do.

The textual role of shell nouns (Schmid's textual function) is apparent in their propensity to compose Theme, particularly in anaphorically referring nominal groups. When shell nouns are used to refer anaphorically or cataphorically to other information, they construe Participants in the discourse. This reification – reconstrual of the message realised as a clause as a Participant – can be seen as similar to Schmid's (2000) cognitive function. As a Participant, the compacted information is then able to be tracked through the discourse, thus, anaphorically referring shell nominal groups constitute an important resource for the thematic progression of the text. Moreover, as a Participant, the nominal group with shell noun as Head benefits from the possibility for modification. The coupling of nominal groups with shell noun as Head and text reference enables the writer create textual effects to focus on what he/she considers newsworthy. In brief, because they can refer anaphorically and cataphorically to a clause, shell nouns contribute to the logogenesis of a text by means of their interplay with Theme and Information structure.

The next section will explore how shell nouns are covered in an SFL perspective at the stratum of discourse semantics.

3.5. Shell nouns in the discourse semantics stratum

Analysis at the discourse semantics stratum explores meaning making resources beyond the stratum of the lexicogrammar. The contribution of shell nouns to textual, interpersonal, and ideational meaning will be discussed in the following sections.

3.5.1. Shell nouns and textual patterning

In abstract discourse such as that found in academic writing, positions of prominence are associated with macro- and hyper-Themes, and macro- and hyper-News (Martin, 1992), which establish a hierarchical structure for the text. Shell nouns play an important role in hyper-Themes (Martin, 1992; 2008; Ravelli, 2004) and in hyper-News (Martin, 1992). Martin, for example, notes the significance of grammatical

metaphor and Winter’s (1977) vocabulary 3 nouns in hyper-Themes and hyper-News, and Martin and Rose (2003) highlight the importance of referring to discourse as Thing to frame the beginning and ends of a story embedded in another text. They argue that it is the act of naming that helps manage the transition between the two genres by providing scaffolding, and they refer to such nouns as metadiscourse (Martin & Rose, 2003: 188; 2008: 119, 254-255). However, as discussed in section 3.3.1., not all “names” are shell nouns. That is to say, not all “names” are semiotic abstractions; some may be semiotic objects (e.g. letter, extract) without the ability to project although they may construe the Sayer with a verbal Process. This distinction does not, however, undermine the argument that shell nouns are a useful resource in hyper- and macro-Themes and in hyper- and macro-News.

Ravelli (2004) argues that the core function of semiotic abstractions and grammatical metaphor in hyper-Themes is to introduce a shift from familiar to more abstract and analytical levels of discussion, thus foregrounding the “processes of technicalizing and rationalizing” (Ravelli, 2004: 124). Semiotic abstractions/shell nouns are a valuable resource for hyper-Themes because they can point both backwards, by distilling information, and forwards, by predicting future development (Ravelli, 2004; Francis, 1994), but their presence alone is insufficient to compose a successful hyper-Theme; they need to appear in certain colligational patterns, in particular relational Process clauses, which Hoey (1998, cited in Ravelli, 2004) refers to as “micro-wording choices”. This dual function is illustrated in figure 3.7. While the lexicalisation of the shell noun is not shown in the example, the shell noun functions cataphorically.

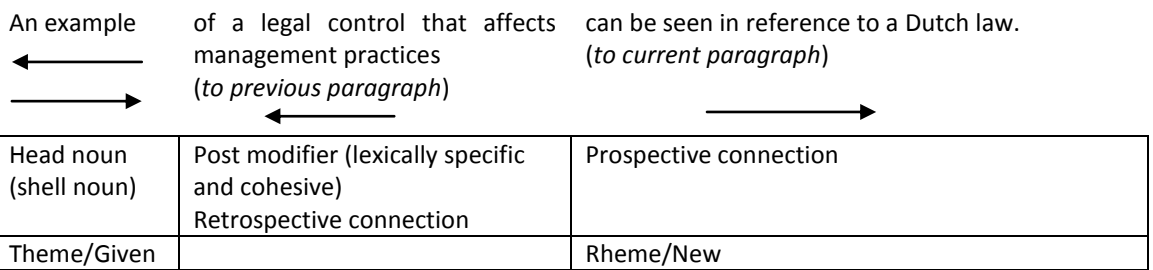


Figure 3.9 Dual facing of semiotic abstraction/shell noun in hyper-Theme (adapted from Ravelli, 2004: 120)

A clear example of a shell noun functioning in a hyper-Theme can be found in Hood (2010) in her work on the genre of research article introductions (3.42). The shell noun is in bold and its lexicalisation or meaning is underlined.

3.42 His methodology showed certain other **refinements**. First, he excluded overseas students. Such students tend to be older than average and also to fare worse academically (Woodley 1979), thus influencing any age/performance relationship. Secondly, he used two measures of performance; the proportion leaving without obtaining a degree and the degree results of those taking final examinations. Finally, he weighted the degree class obtained according to its rarity value in each faculty. (Hood, 2010: 144)

In 3.42 the shell noun **refinements** cataphorically predicts its lexicalisations, which are clearly signalled by the conjunctive Adjuncts *First*, *Secondly* and *Finally*. The example also illustrates Martin and Rose's (2003) claim that hyper-Themes often involve evaluation. The contribution of shell nouns to evaluation from a discourse semantics perspective will be developed in section 3.5.2.

In Ravelli's (2004) analysis of the hyper-Themes in history and management essays, she found that both disciplines used semiotic abstractions to signal the move to more abstract levels of discussion, but that history essays favoured 'clear' grammatical metaphor, i.e. material rather than semiotic abstraction, in hyper-Themes of enhancing paragraphs while management essays favoured semiotic abstractions in hyper-Themes of extending paragraphs. Semiotic abstractions in history often served as an intermediary to introduce nominalisations of Processes whose congruent realisation had appeared in a previous paragraph. An example of such a sequence from Ravelli's (2004: 122) analysis is given in 3.43.

3.43 They [the Dutch] chose to rule through the local regent (§ 2)
this notion of indirect rule (§3).

As Ravelli rightly states, the nominal group with the semiotic abstraction as Head is analogous with, rather than parallel to the antecedent clause. In fact, *this notion of indirect rule* does shift the grammatical metaphor *indirect rule* to a more abstract level. This becomes apparent if the antecedent clause is rewritten as an embedded clause rather than a prepositional phrase, either with directed reference in 3.43' or undirected reference in 3.43''. In both cases the meaning is more specific than the original wording *this notion of indirect rule*.

3.43' this notion that the Dutch chose to rule through the local regent

3.43'' the notion that the Dutch chose to rule through the local regent

Halliday (1988/1993; 1989/1993) notes how a grammatical metaphor realised as a nominal group loses some of the detail of a congruent reading. In this case, presenting the information as Given, the demonstrative *this* as Deictic signals to the reader that he/she must retrieve information from the co-text, and the modifier *Dutch* is inferred rather than explicitly realised. A more appropriate gloss might be “this notion of indirect rule, such as that in which the Dutch chose to rule through the local agent”, or “this notion of indirect rule, i.e. that the Dutch chose to rule through the local agent”, where the fact clause is an appositive elaboration. Following Schmid (2000), it can be argued that in combination with the Deictic *this*, the use of **notion** in 3.43 creates a temporary concept around the antecedent clause and is therefore functioning as a shell noun, the grammatical metaphor in the Qualifier re-lexicalising the congruent realisation of the antecedent clause and specifying the ideational meaning of the shell noun, i.e. filling the semantic gap of the shell noun (figure 3.8). The use of shell nouns in this way – i.e. specified by prepositional phrases with a nominalisation – for temporary concept formation has been found by Aktas and Cortes (2008). It also suggests that the ideational meaning of a Qualifier to a shell noun may contribute to field.

		They [the Dutch]	chose to rule	through the local regent (§ 2)
		Actor	mat. Process	Circumstance
this	notion	of indirect rule (§3)		
Deictic	Head	Qualifier		

Figure 3.10 Reconstituted clause as Qualifier to shell noun

Hood (2008) makes similar claims with regard to concept formation in her analysis of the process of re-instantiation from a source text to notes to a summary. She argues that shell nouns are used to “despecify” meanings while still committing the nominal group to a general ideational meaning such as circumstantial, factual or linguistic. This is in keeping with Halliday’s (1998) claim that assigning an entity to a class is a theoretical operation that involves abstraction.

As noted in section 3.4.3, the discourse organising function of shell nouns has been related to internal conjunction because the shell nouns are organising text, not field (Martin, 1992: 416). Martin develops a system for internal conjunctive relations across clause complexes. The main nodes of the system network are comparative relations, additive relations, consequential relations, and temporal relations. Martin

(1992) notes that textual metaphors – shell nouns – express a meta-message relation. Their contribution to internal conjunction is shown in his analysis of a revised short text from social studies (figure 3.9). The shell nouns are highlighted in bold.

I think Governments are necessary at different levels for	a number of reasons
	INTERNAL CONJUNCTION
For example,	they make laws, without which people would be killing themselves
INTERNAL CONJUNCTION	
	and help keep our economic system in order.
	Let me begin by pointing out that the Federal government fixes up problems that occur in the community.
Another example	is that State Government looks after schools;
INTERNAL CONJUNCTION	
	this prevents vandalism and fighting.
As a final point	the Local Government is important to look after rubbish:
INTERNAL CONJUNCTION	
	otherwise everyone would have diseases.
As a result of these factors ,	Governments at several administrative levels are necessary.
INTERNAL CONJUNCTION	

Figure 3.11 Shell nouns and internal conjunction (adapted from Martin, 1992: 416-417)

In this text *another example* and *As a final point* would be classed as temporal relations, successive ordering and terminating respectively, while *a number of reasons* and *As a result of these factors* would be classed as consequential relations. **Reasons** in the first clause, which functions in a hyper-Theme, points forwards to and is lexicalised by the following five clause complexes. Two other shell nouns – **example** and **point** – appear in textual Themes and contribute to the thematic development of the paragraph. The last shell noun – **factors** – refers anaphorically to the reasons presented in the previous clause complexes, distilling them while reclassing *reasons* as *factors*. As Martin notes, there is no hyper-New.

The above overview shows that shell nouns contribute to textual patterning in the instantiation of a text, both at a more local level and a higher structural level. At a local level, the resources of reference – combining a specific Deictic with a shell noun – move the argument forward. At higher levels, shell nouns – often in relational Process clauses – compose hyper-Themes, which may then be expanded with or without shell nouns by means of the resources of internal conjunction. Similarly, they may be used to signal hyper-News, bringing information to together from one or more clauses.

Although the function of shell nouns has been noted in internal conjunction, they also crop up in external conjunction, but their contribution to the text is either unexplored or dismissed. In fact, some conjunctive Adjuncts include a semiotic abstraction that could be seen to be functioning as a shell noun, e.g. as a **result**, as a **consequence**, for that **purpose**, for that **reason**, in that **event**, in that **case**, at this **moment**. Consider examples 3.44 and 3.45.

3.44 S02: Do they tend to pay; how do they – S04: Per isse. – S02: Per issue. Well in that **case** do they pay after the issues come out? – S04: I think so. (Halliday and Matthiessen, 2004: 360)

3.45 Now prices have sunk for secondary schools and experienced secondary inspectors are shifting into primary and special schools with minimal training. As a **result**, primary schools and teachers are being judged ‘failing’ by inspectors who have never taught younger children, but only watched a couple of lessons on video during their training! (Halliday and Matthiessen, 2004: 547)

It could be argued that in 3.44 and 3.45, **case** and **result** exemplify Schmid’s (2000) characterising function: in 3.44, the noun characterises the preceding information [*they pay*] *per issue* as a condition upon which the following clause is dependent, while in 3.45, the noun labels and signposts the following clause complex as an outcome of changes in the education system. For Halliday and Hasan (1976: 256-257), the use of a prepositional phrase containing a noun such as *reason*, *purpose* or *result* makes it possible to distinguish among the specific kinds of causal conjunctive relations, which is not possible with the general *so*, the simplest form of expression of a causal conjunction. The link between these causal conjunctive Adjuncts and lexicogrammatical patterns for shell nouns identified by Schmid (2000) as key is apparent in the rewrites of 3.44 and 3.45 below.

3.44’ S02: Do they tend to pay; how do they – S04: Per isse. – S02: Per issue. Well, if the **case** is that they pay per issue, do they pay after the issues come out? – S04: I think so.

3.45’ Now prices have sunk for secondary schools and experienced secondary inspectors are shifting into primary and special schools with minimal training. The **result** of this situation is that primary schools and teachers are being judged ‘failing’ by inspectors who have never taught younger children, but only watched a couple of lessons on video during their training!

Despite the apparent link between the causal conjunctive Adjuncts discussed above, it is not particularly helpful to consider them as examples of shell nouns. Moreover, Schmid (2000: 25) notes that the link between a nominal group with a shell noun as Head and its lexicalisation is weaker in indefinite nominal groups than definite nominal groups.

In their discussion of conjunctive or discourse Adjuncts, Halliday and Hasan (1976: 230-233) note that some prepositional expressions may require a reference item, e.g. *in spite of*, while for others such a reference item is optional, e.g. *as a result*. They argue that only when the reference item – either demonstrative *this*, *that* or noun as Head – is anaphoric, is the Adjunct cohesive by means of conjunction. They provide the following example (Halliday & Hasan, 1976: 231).

3.46 The captain had steered a course close in to the shore.

As a result,	}	they avoided the worst of the storm.
As a result of this,		
As a result of this move ,		
As a result of his caution,		

In this example, the shell noun **move** refers anaphorically to the preceding clause. Halliday and Hasan (1976: 232) argue that the reference item, often a general noun such as **move**, “does no more than make explicit the anaphoric function of the whole phrase” and because it is the conjunctive relation that provides the link and not the reference item, any referring item in a prepositional expression is simply included in the conjunctive Adjunct. By contrast, the conjunctive Adjunct *As a result of his caution* is not cohesive by means of conjunction because the pronominal Deictic *his* is cohesive by means of identity of reference.

In fact, Halliday and Matthiessen (2004: 79) argue that discourse/conjunctive Adjuncts construe no experiential role in the clause; they are merely textual, often appearing as textual Theme. Yet, at the stratum of discourse semantics, the textual role of a conjunctive Adjunct qualified by a nominal group with shell noun as Head is clear, and the choice of shell noun may add interpersonal meaning. This becomes apparent in a rewrite of the example 3.46.

- 3.46' The captain had steered a course close in to the shore.
As a result of this **mistake**, they ran aground on a reef.

Like **move**, the shell noun **mistake** encapsulates the preceding clause, but this time the captain's steering close to the shore is assessed negatively. The choice of shell noun clearly opens up a wide range of possibilities for introducing interpersonal meaning in the conjunctive Adjunct. Previous work that broaches on their contribution to interpersonal meaning at the stratum of discourse semantics will be reviewed in the following section.

3.5.2. Shell nouns and interpersonal meaning

Shell nouns can carry interpersonal meaning, that is, they may be attitudinally inscribed or may take on attitudinal meanings by means of prosodic spread. An example of an attitudinally inscribed shell noun whose positive interpersonal meaning spreads to the ensuing text is discussed in Hood (2010). Although Hood does not identify it as a shell noun, understanding it as one is helpful in explaining the prosodic spread of meaning. In the example the shell noun is in bold and its lexicalisation is underlined.

- 3.47 His methodology showed certain other **refinements**. First, he excluded overseas students. Such students tend to be older than average and also to fare worse academically (Woodley 1979), thus influencing any age/performance relationship. Secondly, he used two measures of performance; the proportion leaving without obtaining a degree and the degree results of those taking final examinations. Finally, he weighted the degree class obtained according to its rarity value in each faculty. (Hood, 2010: 144)

Previously I have explained the first clause in the example as a hyper-Theme. Hood (2010: 144) explains the radiating prosody as spreading from the coupling of *methodology* and the positively inscribed **refinements**, to the Processes in the ensuing clauses – *excluded*, *used* and *weighted* – interpreted as activity sequences making up component parts of the field *methodology*. She also shows how flipping the positive appreciation to negative – replacing **refinements** with **problems**, which can also function as a shell noun – results in a spread of negative value, while removing the shell noun completely results in no implication of value. Because shell nouns encapsulate the meanings of entire clauses, according to Winter (1977), when this

information does not precede the shell noun, the reader looks for the information prospectively. This is the case in example 3.47; the positive attitude of **refinements** spreads to its lexicalisations in the following clauses. The ability of the shell noun to enter into identifying relational Process clauses with a fact clause as Token makes both the assessment and the connection between the shell noun and its lexicalisations explicit as the rewrite in 3.47' shows. In the rewrite, the instances of conjunctive Adjuncts are reconstrued as Value in a relational Process clause, and the clauses realising the lexicalisation in the original have been rank-shifted to function as Token. The nominal groups with shell noun as Head that compose Theme can also be interpreted as instances of internal conjunction.

(3.48') His methodology showed certain other **refinements**. **The first refinement** was that he excluded overseas students. Such students tend to be older than average and also to fare worse academically (Woodley 1979), thus influencing any age/performance relationship. **The second refinement** was that he used two measures of performance; the proportion leaving without obtaining a degree and the degree results of those taking final examinations. **The final refinement** was that he weighted the degree class obtained according to its rarity value in each faculty.

In academic discourse shell nouns are also highly likely to enter into the system of ENGAGEMENT for as Halliday and Matthiessen (2004: 468) note: “[In scientific discourse] proposals and demands are opposed, renounced and rejected.” Nominalisations of verbal and mental Processes can project the writer’s or other researchers’ ideas (Martin & White, 2005: 111), enabling the writer to position him/herself with regard to other texts, either aligning with or distancing him/herself from them. A couple of examples will be given as illustration. In 3.48, from Martin and White (2005), shell nouns are in bold, lexicalisations are underlined, and contraction is boxed.

3.48 Sir, Your report ('Anthrax vaccine refused by half Gulf personnel', February 12), recorded comments by Paul Keetch MP who claimed that the Ministry of Defence was 'sowing confusion' among troops by making this programme voluntary and that by doing so it was abdicating leadership. May I repeat my assurances that this is not the case. Anthrax represents a real threat to our armed forces and we seek to protect our troops through detection systems, individual physical protection and medical countermeasures (immunisation and antibiotics). But the best single protection against anthrax is immunisation. [*The Times*, letters to the editor, 21/02/03, from Lewis Moonie, MP, Parliamentary

Under-secretary of State for Defence and Minister for Veterans Affairs] (Martin and White, 2005: 119)

The text is heteroglossic but contractive; the writer is clearly disaligning himself from the views of Paul Keetch (*claimed* - Attribute: distance) and invites the reader to do the same. There are two shell nouns in this example: **assurances** and **case**. The writer uses text reference to refer anaphorically to Paul Keetch's argument, and, in an embedded relational clause, identifies the argument as incorrect with a shell noun (*not the case* – Disclaim: deny). The negation of Paul Keetch's view is presented as New, and the structure sets up the context in which his own view, presented following the denial, *is* implicitly seen to be the actual case. There is, however, more interpersonal work going on. The denial of Paul Keetch's view is strengthened with two projecting motifs: *May I repeat* (Proclaim: pronounce) and *my assurances that* (Proclaim: pronounce), both of which serve to suppress alternative views. The nominal group - *my assurances* – has a shell noun as Head, and it adds an additional layer of assessment to the embedded proposition, which constitutes the lexicalisation of the shell noun, **assurances**. The writer's view could have been realised as 3.48', 3.48'' or 3.48''', none of which involves the use of a second shell noun to frame the proposition, although they all retain the anaphorically referring shell noun **case**.

3.49' This is **not the case**.

3.49'' May I repeat that this is **not the case**.

3.49''' I assure you that this is **not the case**.

A second example 3.49 illustrates how shell nouns can be used to close the dialogic space between the speaker and listener. Again the resources of contraction are boxed, the shell noun is in bold, and the lexicalisation is underlined.

3.49 Now it is time to take longer strides – time for a great new American enterprise – time for this nation to take a clearly leading role in space achievement, which in many ways may hold the key to our future on earth. I believe we possess all the resources and talents necessary. **But the facts of the matter are that we have never made the national decisions or marshalled the national resources required for such leadership.** (John F. Kennedy) (Martin & White, 2005: 127)

In this extract from a speech by J.F. Kennedy, Kennedy closes the dialogic space between himself and his audience by using a shell noun, **facts**, to embed the propositions that 'we have never made the national decisions or marshalled the

national resources required for such leadership'. As Martin and White (2005: 128) explain: "Such insisting or emphasising imply the presence of some resistance, some contrary pressure of doubt or challenge against which the authorial voice asserts itself." The overt, but objective sounding, *the facts of the matter are that* (Proclaim: pronounce) makes it harder to challenge the embedded propositions. The shell noun provides an explicit and objective matrix for the proposition at stake (Martin & White, 2005: 131). At the stratum of the lexicogrammar, once again the identifying clause with the shell nominal group as Theme sets up the embedded clauses as New for the listener to pay attention to. The discourse semantics approach here differs from Schmid (2001), who discusses such uses as topicalising (c.f. section 3.1.4 pp 43-44), which limits the interpretation to the stratum of the lexicogrammar.

In addition to these examples of shell nouns that function to contract the dialogic space between speaker and listener, shell nouns can be used to expand the dialogic space by attributing arguments to others, whether they are explicitly named or not.

3.50 ... debate about Maastricht. The rest of Europe, they tell us, has not matched up to us and would have avoided many of their own domestic political embarrassments had they done so. There may once have been a certain justice in **this claim**. (Schmid, 2000: 162)

In 3.50, **this claim** (Attribute: distance), the shell nominal group distances the writer from view expressed by the un-named 'they' in the preceding clause that 'The rest of Europe has not matched up to us and would have avoided many of their own domestic political embarrassments had they done so'; however, the writer does not suppress that view. The distancing couples with resources from GRADUATION (*a certain justice* – Focus: valeur: specificity) to disalign the writer further with the lexicalised proposition of the shell noun **claim**.

While it seems easy to relate mental and linguistic shell nouns to the system of ENGAGEMENT, factual, modal, eventive and circumstantial shell nouns may not be so straightforward. Hood (2010: 180) argues that 'fact' clauses (Halliday & Matthiessen, 2004) such as those projected by shell nouns - e.g. the **fact** that - "contribute to the heteroglossic or multi-voiced nature of engagement in the discourse" because it is

possible to identify the source to whom “knowledge of the fact is attributed”. The source may be general or specific; the writer/speaker, as in Kennedy’s speech, which was discussed earlier in example 3.49, or someone else. The previous discussion, however, also suggests that shell nouns may compose a matrix that contributes to enacting the dialogic position of the writer. If this is the case, then logically it might be possible to classify all shell nouns as functioning in the ENGAGEMENT system. One of the aims of the analysis of the corpus at the discourse semantics stratum, therefore, is to explore to what extent this surmise is valid.

Hood (2010) also provides examples of how shell nouns can contribute to the system of GRADUATION. The system of GRADUATION, which functions to grade attitudinal and experiential meanings along a scale of more or less, has been extended by Hood (2010: 103-105), who offers an alternative approach to projecting Processes and nouns to Martin and White’s system of ENGAGEMENT. Hood argues that in academic discourse projecting Processes such as ‘suggest’ or ‘indicate’ and their nominalisations contrast with Processes such as ‘show’ or ‘demonstrate’, and it is possible to grade them on a scale as more or less fulfilled in terms of completion or irrealis. Thus, by means of a shell noun, a projected proposition may be graded as more or less fulfilled, and this can function to invoke an attitudinal reading. Shell nouns such as **attempt**; which operate in the system of CONATION, evaluate the projected proposition as incomplete, in contrast with a shell noun such as **achievement**. Similarly, shell nouns such as **possibility** or **probability** can be seen as not fully actualised.

3.5.3. Shell nouns and ideational meaning

At the stratum of discourse semantics ideational meaning is realised through activity and implication sequences, in which the role of entities and Processes are classified according to taxonomic relations (Martin & Tose, 2003). Through the activity and implication sequences it is possible to determine the field of discourse. While the literature review has shown that most shell nouns are not field-specific, in academic discourse a concentration of shell nouns related to the field of research has been noted (e.g. Swales, 2005). This suggests that some shell nouns such as **result** or **finding** may contribute to field. Hood (2008) notes how shell nouns can be useful in writing

summaries as they commit less ideational meaning to the text than the source wording, but they retain the meaning of some kind of abstract relationship. The abstraction may then be respecified by means of a Qualifier. Schmid (2000) accounts for this kind of meaning in his explanation that shell nouns are made up of two kinds of meaning: one stable, and the other that must be retrieved from the co-text. Given the lack of discussion on the relation of shell nouns to ideation at the stratum of discourse semantics, one of the aims of this study is to further understanding of the matter.

3.5.4. Summary of shell nouns at the discourse semantics stratum

This section has provided an overview of how shell nouns contribute to interpersonal meaning, textual and ideational meaning at the discourse semantics stratum. First, they enter into the resources of the systems of APPRAISAL. Some shell nouns are inherently attitudinal, e.g. **benefit, refinement, problem**, and can set a positive or negative prosody that spreads to the surrounding discourse, in particular to the lexicalisations of the shell noun. Others may render an attitudinal reading by invocation. Shell nouns also enter into the system of ENGAGEMENT where they serve to position the writer with regard to propositions and proposals. They may open up the dialogic space (e.g. **claim, hypothesis**) or close it down (e.g. **fact, proof**), thus contributing to the writer's stance. Second, shell nouns contribute to the periodicity of a text through composing macro- and hyper-Themes and macro- and hyper-News. As abstract nouns that require their meaning to be retrieved from the surrounding discourse, they provide an efficient means of announcing or summarising their co-text, and by compacting prior or up-coming text in a nominal group, they contribute to the development of the text. Finally, shell nouns seem to play a limited function in construing field, with a notable exception, that of the field of research.

The discussion in this section suggests that shell nouns could play a significant role in research articles given that the writer is engaging in a community where there are conflicting views, and one of the means by which positions of power within the community are negotiated is the research article. However, the full extent of their functions at the stratum of discourse semantics is not yet completely understood. One of the aims of this study is to redress this gap. Chapter 4 will address the methodology

used to carry out the study of shell nouns in the research articles that comprise the corpus of this thesis.

Chapter 4. Methodology

4.1 Research questions

The study of shell nouns in this thesis is approached using the principles of Systemic Functional Linguistics. The analysis seeks to answer the following general questions:

1. In what ways do shell nouns contribute to ideational, interpersonal and textual meaning?
2. Are there significant differences in their use between published articles and article submissions by Portuguese academics in the field of economics?

In order to answer these questions two types of analysis will be carried out. One is qualitative and it will focus on the ways in which shell nouns make meaning in a Systemic Functional Linguistics perspective. The qualitative analysis will examine the ideational, interpersonal and textual functions of shell nouns at the stratum of the lexicogrammar and the stratum of discourse semantics. The former is based on the principles of Halliday and Matthiessen (2004) and the latter on Martin and Rose's (2003) approach to ideation, Martin's (1992) textual systems of hyper-Themes and hyper-News, and Martin and White's (2005) system of APPRAISAL. This qualitative analysis seeks to answer the questions

- i. What textual, ideational and interpersonal functions do shell nouns carry out in the lexicogrammar?
- ii. What textual, ideational and interpersonal functions do shell nouns carry out in the discourse semantics?

In addition to the qualitative analysis, I undertake a quantitative analysis of the differences in function of the shell nouns between the sub-corpus made up of published articles and the sub-corpus made up of article submissions. This quantitative analysis is made possible by the statistical tools incorporated in the software used to tag the corpus: Corpus Tool (O'Donnell, 2008).

Having outlined the research questions for the study of shell nouns, the rest of the methodology chapter is structured as follows. First I shall describe the criteria for

selection of the corpus and the use of Corpus Tool, then I shall describe the systems developed to undertake the analysis in order to answer the research questions for the analyses at the strata of the lexicogrammar and discourse semantics.

4.2. Selection of the corpus

Corpus studies are well established in SFL, in which the description of grammar is corpus driven, i.e. it is informed by theory – the system – and by the data – the instance (Halliday & Matthiessen, 2004). A corpus study, therefore, is empirical as it uses authentic texts, which may be spoken or written (Biber & Conrad, 2009; Halliday & Matthiessen, 2004; Leech, 1991), and which may be ‘interrogated’ in order to understand the system better (Halliday, 1991).

The construction of a corpus is governed by the research questions of the study (Engwall, 1991; Gellerstorm, 1991; Rogers, 2000; Teubert & Čermáková, 2004/2007). While computer techniques have made it possible to work with large corpora of 1 million words or more, such as a reference corpus (Teubert & Čermáková, 2004/2007) or text bank (Leech, 1991), smaller and specialised corpora of 20,000 – 250,000 words can provide valuable insights that would not come to light in larger corpora due to search restrictions (Flowerdew, 2004). One advantage is that they enable qualitative, contextually informed analyses that use custom-built tagging systems that can tag, for example, larger units of text such as rhetorical moves (Flowerdew, 2004). They are also suitable for studies of lexical phenomena (Engwall, 1991). Small corpora – ranging from about 10,000 words to 250,000 words, have been used to study various lexicogrammatical features such as Theme in peer reviews of scientific article submissions (Godsen, 2009) and across textual categories (Gómez-González (2001); the extent to which economic agents in finance articles are construed congruently or incongruently as Senses (Chevalier & Hudson, 2001); and the distribution of anaphoric items in written and spoken texts (Biber, 1991). However, studies using specialised corpora or small corpora have raised concerns about the representativeness of the corpus and the generalizability of the findings (Mair, 1991). Flowerdew (2004) argues that the latter drawback can be mitigated by comparing findings with those from a general corpus. At any rate the selection of texts to compile the corpus should be principled (Biber & Conrad, 2009) and based on clearly stated criteria (Halliday &

Matthiessen, 2004: 36). Flowerdew (2004: 21) provides criteria for determining whether a corpus is specialised (table 4.1). The parameters are not discrete, but involve some overlapping.

Parameters	Details/Examples
Specific purpose for compilation:	To investigate particular grammatical, lexical, lexicogrammatical, discoursal or rhetorical features
Contextualisation:	Setting (e.g. lecture hall) Participants (role of speaker/ listener; reader/ writer)
Size:	
whole corpus	1-5 million words
subcorpus or small-scale corpus	20,000 – 250,000 words
Genre:	Promotional (grant proposals, sales letters)
Type of text/ discourse	Biology textbooks / casual conversations
Subject matter / topic	Economics, the weather
Variety of English	Learner, non-standard (e.g. Indian, Singaporean)

Table 4.1 Parameters for defining a corpus as specialised (Flowerdew, 2004: 21)

The study of shell nouns in this thesis uses a specialised (Flowerdew, 2004), parallel corpus (Teubert & Čermáková, 2004/2007) comprising five published research articles and five article submissions by Portuguese academics working in the Economics department of ISEG. The principle behind a parallel corpus is that by means of comparison, differences between the two sub-corpora will become apparent, even if the corpus is small (Sinclair, 2001 in Flowerdew, 2004: 17). Relevant studies to this thesis that use specialised parallel corpora have been reviewed in chapter 3. They include Flowerdew (2003a; 2003b) on shell nouns in spoken lectures and textbooks; Charles (2003; 2007) on shell nouns in two contrasting disciplines; Moreno (2004) on metatext, including shell nouns, in Spanish and English research articles; Fontaine and Kodratof (2003) on Thematic progression in English-medium scientific articles by French and English researchers; Corbett (2009) on differences in Theme between academic articles and their popularisations; and Flowerdew (2003c) on differences in interpersonal selections for the problem-solution pattern between student and professional writers.

Parameters	Subcorpus: Published articles	Subcorpus: article submissions
Specific purpose for compilation:	To investigate the functions of shell nouns in the lexicogrammar and discourse semantics and their contribution to the logogenesis of the text published articles	nouns in the lexicogrammar and discourse semantics and their contribution to the logogenesis of the text article submissions
Contextualisation: Setting	Member of discourse community	Member of discourse community
Participants: Writer	Member of discourse community	Reviewer/Member of discourse community
Reader		
Size:	39,216 words	35,908 words
Genre:	Multiple genres	Multiple genres
Type of text/ discourse	Research article	Research article
Subject matter / topic	Economics	Economics
Variety of English	published	EAL; Portuguese L1

Table 4.2 Parameters of the corpus for the study of shell nouns

The parameters of the corpus for this study are summarised in table 4.2. The two sub-corpora were chosen to reflect the context in which Portuguese academics work, i.e. to be representative of an international specialised discourse community of academics undertaking research in the field of economics. As discussed in the introduction, it has become increasingly important for career purposes for these academics to publish in peer-review international journals. For the academics working in ISEG, the economics journals indexed on the Institute for Scientific Information's (ISI) Journal of Citation Reports (JCR) operate as a kind of benchmark, and publishing in these journals constitutes a goal. In light of the normative views concerning language and the language policies practised by these journals (Henshall, 2012), articles published in them can be taken as examples of acceptable standards of language for the discourse community regardless of whether they were written by a native speaker of English or not. As pointed out by Fontaine and Kodratoff (2003), the fact that the article has been published is a sign that it has been accepted by the community, both in terms of quality of research and in terms of acceptable language standards.

The Portuguese academics wishing to participate actively in the community need to have their submissions accepted, and they may not have the luxury of having them copy-edited by a language specialist before submission (Barros, 2014) as many journals recommend (Henshall, 2012). That is why I decided to use article submissions that had not been edited by a language specialist. In addition, to ensure that the discourse was representative of L1 speakers of Portuguese, co-authored papers with L1 speakers of languages other than Portuguese were excluded. This decision stems, in part, from a curiosity to see, albeit indirectly, if the criticism levelled at southern

European writers about their use of 'vague' nouns (Lillis & Curry, 2010: 150) is in any way valid.

Having explained the criteria for the selection of the corpus, I shall now describe the procedure for compiling it. In May 2010 an email was sent to all the professors in the Economics department at ISEG requesting their cooperation to build the corpus. In the email the purposes of the study were explained, and the researchers involved were assured that their contributions would remain anonymous. I asked them to provide me with research papers that they had written and wished to submit to a referenced journal for publication. I also requested that the submission had not been edited for language by a language expert. 14 professors responded that they would be interested in participating, and 9 of these sent a total of 15 research papers they had written. Of these papers, one was excluded from the study because it was a working paper, not an article submission, and a further two were excluded because they were joint papers with one or more authors whose L1 was not Portuguese. Another paper was excluded because it was not clear whether it complied with the condition that it had not been revised by a language expert. Of the remaining 11 papers, nine had already been accepted for publication or had, in fact, already been published although the authors of the papers assured me that what they had sent me was what they submitted. Because two papers would be insufficient for the purposes of the analysis of shell nouns, three other papers were selected to make up the corpus for article submissions. Care was taken to ensure that the topics broached in the articles were varied. As a result of this procedure, the final corpus for article submissions comprised five papers.

The parallel corpus was built using Google Scholar. Using the keywords from each article submission, a search was entered on Google Scholar to find a study that covered similar ground to that of each article submission. On the basis of the abstract and a read of the article, five published articles that were in journals indexed on the ISI JCR were selected.

Each text from the two sub-corpora was saved in a Word document, then cleared of tables, figures, footnotes, references, and information on the authors. The

cleaned text was then saved in plain text format so that it could be uploaded onto Corpus Tool. Information on the two sub-corpora are summarised in Table 4.3, and the corpus is available in Annex B.

Published articles				
Title	Author	Journal	Words	Corpus Code
Paths to stability for matching markets with couples	Bettina Klaus, & Flip Klijn	Games & Economic Behavior, 58, 2007: 154-171	6,711	P_Rand
The Lifecycle of a metropolitan business network: Liverpool 1750-1810	John Haggerty & Sherrylene Haggerty	Explorations in Economic History, 48 (2) 2011: 189-206	10,918	P_Net
The Probabilistic Counter-Revolution, or How Stochastic Concepts came to Neoclassical Economic Theory	Philip Mirowski	Oxford Economic Papers, New Series, 41 (1), History and Methodology of Econometrics, 1989: 217-235	8,080	P_Econ
Bayesian Factor analysis for spatially correlated data, with application to summarizing area-level material deprivation from census data	Joseph W. Hogan & Rusty Tchernis	Journal of the American Statistical Association, 99 (466) 2004: 314-324	6,355	P_Pov
Managerial pay and governance in American nonprofits	Kevin F. Hallock	Industrial Relations, 41 (3) 2002: 377-406	7,152	P_CEO
Total words			39,216	
Article submissions				
Random matching in the college admissions problem			10,677	S_Rand
Social capital and economic performance: did networks really work?			8,729	S_Net
<i>The Fellowship of Econometrics</i> - Selection and Diverging Views in the Province of Mathematical Economics, from the 1930s to the 1950s			6,705	S_Econ
Deprivation analysis based on Bayesian Latent Class Models			5,302	S_Pov
Governance and CEO Pay and Performance in Non-Profit Organizations			4,495	S_CEO
Total words			35,908	

Table 4.3 Summary of characteristics of the two sub-corpora

4.3. Corpus Tool

The corpus, which comprises two sub-corpora – published research articles and Portuguese article submissions – is analysed using UAM Corpus Tool (O’Donnell, 2008). The main advantage of this software is that it enables specific segments of the corpus to be tagged and coded according to a system developed by the researcher. The tagged items can then be retrieved easily. The software also incorporates statistical tools that make it possible to compare data sets with regard to their features and to compute the statistical significance of differences in the relative frequencies of features. The software runs two commonly used tests: T statistics and Chi Square.

4.4. Identifying and tagging the shell nouns

The identification of shell nouns is based on the criteria developed by Schmid (2000), as detailed in chapter 3. Hence, a shell noun is identified by its function: the capacity of the noun to project in the pattern *N-cl* or appear in an identifying relational clause such as *N-be-cl* and the environment in which it occurs, i.e. one of the patterns identified by Schmid (2000) (see table 4.4). From a Systemic Functional Linguistics perspective, shell nouns are **grammatical metaphors**, corresponding to semiotic abstractions, as discussed in chapter 3, sections 3.3.1 and 3.4.1.

Shell noun Pattern	Example
<i>N-cl</i>	<u>The fact that the head of one of Nigeria’s northern-based regiments, based in Kaduna, has already pledged support for President Babangida</u> reinforces the view that the coup is ethnically-based. (Schmid, 2000: 362)
<i>N-be-cl</i>	<u>The fact is that neither of these institutions have significant social or political research arms.</u> (Schmid, 2000: 97)
<i>th-be-N</i>	<u>(I won the freshmen’s cross-country. – Mm) That was a great achievement</u> wasn’t it? (Schmid, 2000: 22)
<i>th-N</i>	Yet another theory suggests that <u>the goal of the welfare state, within a society in which economic competition under capitalism dominates, must be to effect gradual reform.</u> This Fabian approach argues for ... (Schmid, 2000: 344)

Table 4.4: Shell noun patterns (Schmid, 2000)

While it is easy to identify nouns functioning as shell nouns in the first three patterns exemplified in table 4.4, not all anaphorically referring abstract nouns – although potential shell nouns – are in fact functioning as one. Whether an anaphorically referring noun is functioning as a shell noun or not depends on its relation to the text that functions as its lexicalisation. For Schmid (2000: 4), the

lexicalisation should be ‘complex, proposition-like pieces of information’. As revealed in chapter 3, in some studies (e.g. Atkas & Cortes, 2008; Gray & Cortes, 2011) nouns referring to nominalisations have been counted as shell nouns, and indeed, Schmid (2000) includes examples where the lexicalisation of the shell noun is realised as a prepositional phrase qualifying the noun. However, because the criteria for defining shell nouns can become somewhat blurred once it is accepted that the lexicalisation of a shell noun can be realised by a nominalisation, I have chosen to limit the interpretation to lexicalisations that are realised by a clause or another nominal group with a shell noun as Head. As a result, examples such as 4.1 are excluded from the analysis. In 4.1 the potential shell noun *stipulation* refers anaphorically to an idea that is realised as a nominalisation only while the postmodifying ‘that’ clause is a defining relative clause, i.e. an elaborating embedded clause, not a projection.

4.1 Furthermore, it goes against the stipulation of the European Commission that stresses the importance of having a non-executive Chairman of the Board (Berglof, 1997). This **stipulation** is mirrored ... (S_CEO)

Some anaphoric cases only superficially look like shell nouns, as in 4.2. In this case ‘governance principles’ represents a step up in generality of the ‘codes of practice’ and ‘lack’ is a nominalised Process, but not a shell noun given that it cannot accept a projection. Such nominalisations are not deemed to be functioning as shell nouns.

4.2 Hence we can assert that Portugal lacks necessary codes of practice, there being no code whatever for non-profit organizations, which may affect negatively their efficiency. ... Since there is a recognized positive relationship between best-governance practices and efficiency (Ames, 2003), **this putative lack of governance principles** is a cause of inefficiency in the non-profit activity.” (S_CEO).

Once identified, the shell nouns can be classified according to the six main classes proposed by Schmid (2000): factual, linguistic, mental, modal, eventive, or circumstantial; and their lexicogrammatical and discourse semantic functions are tagged using the Corpus Tool software.

4.5. Tagging the lexicogrammatical stratum

The aim of the analysis of shell nouns at the stratum of the lexicogrammar is to identify the functions carried out by the shell nouns with regard to the three metafunctions: ideational, interpersonal and textual. In addition the study seeks to explore the kind of clauses in which shell nouns appear. The system used for the analysis, therefore, distinguishes between the metafunctions and the clausal environment. This initial distinction is shown in figure 4.1. The systems for tagging the metafunctions of shell nouns at the stratum of the lexicogrammar are based mainly on Halliday and Matthiessen (2004). In this section I shall first describe the ideational, textual and interpersonal systems that are used to tag the corpus, then the system that classifies the type of clause in which the shell noun appears. In keeping with SFL practice, system names are in upper case throughout the description.

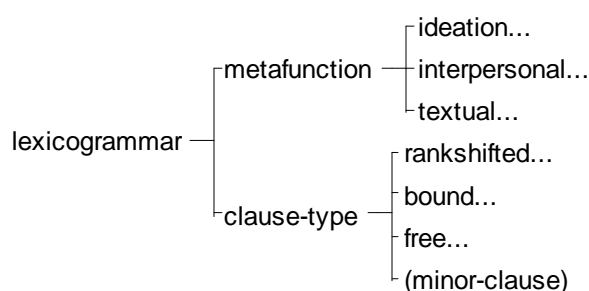


Figure 4.1 Primary systems for analysis at the stratum of the lexicogrammar

4.5.1. Tagging ideation at the lexicogrammatical stratum

The scheme developed for tagging the ideational metafunction of the corpus at the stratum of the lexicogrammar aims to identify the ideational functions – both experiential and logical – that shell nouns construe. Given that the base of the lexicogrammar is the clause and shell nouns construe Head in a nominal group, this task might assume that the shell noun is functioning either as a Participant or an indirect Participant, i.e. within a prepositional phrase. Yet Schmid's criteria for identifying shell nouns include cases where the shell noun appears within a complex conjunctive preposition (e.g. in the **event** that) or within a conjunctive Adjunct (e.g. As a result of this **move**), the latter not being considered an element in the transitivity of the clause (Halliday & Matthiessen, 2004). In order to include these uses, an initial distinction is made between shell nouns with an ideational function at the rank of the

clause and those without. The latter node – NO IDEATIONAL FUNCTION – is used to tag shell nouns appearing within conjunctive Adjuncts. Notwithstanding, only conjunctive Adjuncts that contained anaphoric reference to prior text are tagged (e.g. With this **purpose** in mind, In view of this **result**); general conjunctive Adjuncts such as As a **consequence**, In **conclusion** or As a **result** are not tagged. The motivation for excluding the latter is that the link between the potential shell noun and the following text is more tenuous, as detailed in section 3.5.1. The node NO IDEATIONAL FUNCTION is also used for minor clauses, which include headings, given that they do not play a role in the transitivity of a clause (Halliday & Matthiessen, 2004: 100).

The system IDEATIONAL FUNCTION is split to reflect the difference between the experiential and logical functions within the clause. In Halliday and Matthiessen (2004: 373) the logical function includes the systems of TAXIS, LOGICO-SEMANTIC RELATIONS and the option of ‘going round again’, and the network gives the resources for specifying the relationship between clauses. As such, it pertains to information beyond the clause, so I have not included these relationships here. Nevertheless, shell nouns may function within binders, which construe the logical function of introducing a hypotactic clause. For the purposes of the analysis, I have included a LOGICAL system that consists of one node: Relator (Halliday & Matthiessen, 1999) to capture shell nouns with this function. These binders are all nominal conjunctions (Halliday & Matthiessen, 2004: 419). The other basic system of IDEATIONAL FUNCTION is EXPERIENTIAL FUNCTION. This system distinguishes between ranking at the group and clause levels. The systems network to this level is given in figure 4.2. I shall first describe the system developed for tagging the elements at group level and then do the same for the system to tag ranking elements in the clause.

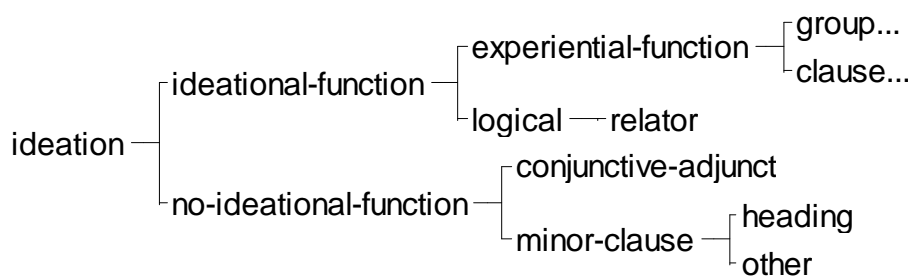


Figure 4.2 Basic systems for the analysis of the ideational metafunction at the lexicogrammar stratum

Since the focus of this thesis is the shell noun and given the potential for modification within a nominal group, the experiential system at group level was developed to explore the lexicogrammatical environments of shell nouns in more detail. There are three subsystems. First, the Deictic opens up to the system of DEIXIS, which is taken from Halliday and Matthiessen (2004). Second, Thing is extended in delicacy to capture the kind of shell noun involved: factual, linguistic, mental, modal, eventive, or circumstantial (Schmid, 2000). Finally, the node for Qualifier is extended in delicacy. One of the defining patterns for shell nouns is that the noun can accept a projection or an act clause as Qualifier. Therefore an initial distinction between post-modifying clauses that are projected and those that are expanding is made. One of the purposes of this distinction is to be able to identify nouns that are functioning as shell nouns by virtue of the kind of post-modifying clause, in other words, to be able to separate shell nouns whose lexicalisations are the following post-modifying clause from those whose lexicalisations lie elsewhere. Although strictly speaking an act clause is in the realm of expansion, as explained in section 3.4.1, such act clauses are counted as projected, given their proximity to the logico-semantic relation of projection (Halliday & Matthiessen, 2004: 441) and the fact that they would occur with eventive shell nouns. This distinction between projected and expanding Qualifiers has other benefits. It makes it possible to systematically describe the ways in which shell nouns are modified: whether it is by a prepositional phrase or a clause, and, if a clause, the type of clause: finite/non-finite: perfective/imperfective/neutral. It also makes it possible to identify when a shell noun's experiential function does not extend beyond the group. For example a shell noun may appear within a prepositional phrase that functions as Qualifier to another noun (figure 4.3). The shell noun is in bold.

What	are	the	implications	of these results?	(P_CEO)
Token	Pro: Rel	Value			
		Deictic	Thing	Qualifier	

Figure 4.3 Shell noun in Qualifier realised as prepositional phrase

The network for the experiential function at group rank is illustrated in figure 4.4



Figure 4.4 System network for experiential function at group level

The system designed for tagging the functions of elements containing shell nouns in the transitivity of the clause is not restricted to this purpose. In addition I tag the Process of the clause. The motivation behind this decision is to understand better the lexicogrammatical environments in which shell nouns appear. The basic systems for tagging elements at clause rank are shown in figure 4.5.

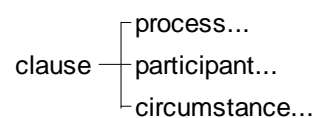


Figure 4.5 Primary systems for tagging experiential elements of the clause

At the rank of the clause, the network makes it possible to tag the transitivity function of the element within which the shell noun appears, i.e. Participant, Attribute or Circumstance. Although Attributes are not considered Participants because they cannot enact the Subject (Martin, 1992), for the purposes of this study they are tagged as such so that it is possible to compare the propensity of shell nouns to construe Attribute as opposed to another Participant. The role of PARTICIPANT is extended in delicacy to include the type of Participant (Actor, Goal, Token, etc) (figure 4.6) and the selection of Circumstance is refined in the system of CIRCUMSTANCE, which distinguishes between elaborating, extending, enhancing, and projecting Circumstances (figure 4.7). These systems follow Halliday and Matthiessen (2004).

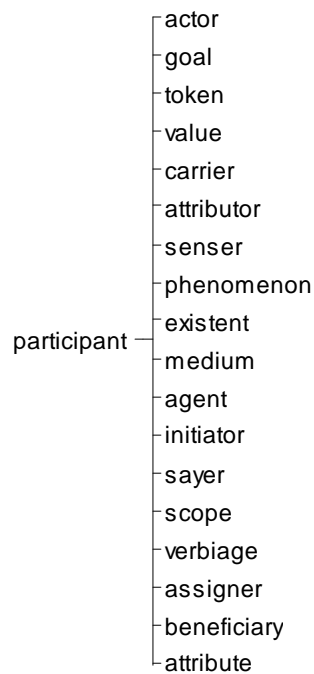


Figure 4.6 System network for PARTICIPANT

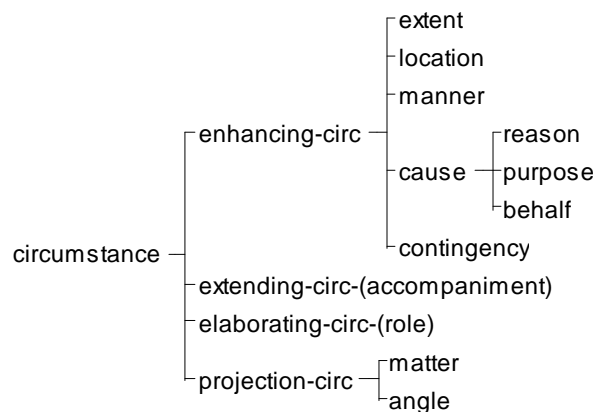


Figure 4.7 System network for CIRCUMSTANCE

In addition to parsing the shell noun's role in the transitivity of the clause, I tag the Process of the clause in which it appears. This procedure includes when shell nouns appear in embedded clauses or hypotactic bound clauses as shown in figures 4.8 and 4.9 respectively, but not when they appear within prepositional phrases functioning as Qualifiers as in figure 4.10. The aim, again, is to obtain a deeper understanding of the environments in which shell nouns appear.

Thus the structure of the group provides further information on the principals and respective ego-networks to be considered in the model which

	tests	the	second	hypothesis (P_Econ)
	Pro: mat	Goal		
		Deictic	Num	Thing: mental shell noun

Figure 4.8 Transitivity analysis showing tagging of elements in an embedded clause containing a shell noun

Precisely because this **claim** may appear controversial, I shall not attempt a full elaboration here

	Carrier	Pro: Rel: Attributive
	Deictic	Thing: mental shell noun

Figure 4.9 Transitivity analysis showing tagging of elements in a hypotactic bound clause containing a shell noun (P_Econ)

In this context, restricting each agent to hold the potential partner that is higher on some fixed preference ordering sustains the validity of the **results** of the preceding section (S_Rand)

	Qualifier		
	Deictic	Thing: factual shell noun	Qualifier: prepositional phrase

Figure 4.10 Transitivity analysis showing tagging of elements in a Qualifier containing a shell noun

In cases where the shell noun appears within a complex conjunction introducing a bound hypotactic clause (Halliday and Matthiessen, 2004: 419; Schmid, 2000: 291), the Process of the primary or bound clause is not tagged unless there is a shell noun functioning within the TRANSITIVITY system of the clause. As explained above, shell nouns appearing within binders are included not within the system of EXPERIENTIAL FUNCTION but within that of the LOGICAL. The Processes of clauses introduced by a conjunctive Adjunct that contains a shell noun are also not tagged because the conjunctive Adjunct is considered to lie outside the transitivity of the clause. Examples of tagging the shell nouns appearing within a conjunction and a conjunctive Adjunct are given in figures 4.11 and 4.12 respectively.

This model was suggested on the **grounds** that conditional spatial correlation $\text{corr}(z_i, z_j) = \frac{1}{k} \sum_{R_i} R_j$ is not dependent on neighbourhood size (P_Pov)

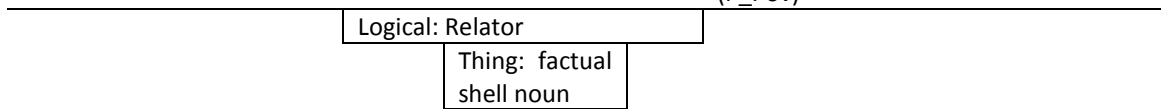


Figure 4.11 Transitivity analysis showing tagging of elements in a conjunction

In view of this **result**, the class of weakly responsive preferences is a natural starting point for the study of decentralized decision making in couples markets. (P_Rand)

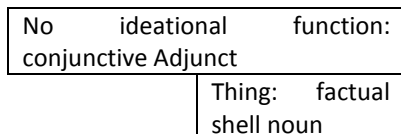


Figure 4.12 Transitivity analysis showing tagging of elements in a conjunctive Adjunct

Because it is to be expected that the shell nouns will appear most often in relational clauses, a system with greater delicacy for relational processes is desirable. Following Martin (1992) and Halliday and Matthiessen (2004), relational Processes are considered to include a neutral form (*be*) as well as marked variants (e.g. *turn*, *become*, *prove*). As detailed in section 2.8, what appear to be verbal or mental Processes may be functioning as relational Processes, such as when the Sayer of a verbal Process is construed by a semiotic abstraction or semiotic object. I have chosen to include this distinction as a systemic choice operating simultaneously with the systems of MODE OF RELATION and TYPE OF RELATION rather than as a subsystem of intensive attributive Processes (neutral/phased) (Halliday & Matthiessen, 2004) or a subsystem of intensive identifying Processes (unmarked/marked) (Martin, 1992: 281). The downside of this choice is that the distinction between neutral and phased in attributive Processes is lost. However, as well as being more economical (Martin, 1992: 284), it makes it possible to call up concordances of each type as well as compare frequencies using the statistical tools incorporated in Corpus Tool.

Cross-classification of the systems for MARKEDNESS, MODE OF RELATION and TYPE OF RELATION gives rise to combinations such as attributive intensive or identifying intensive Processes, which, as discussed in section 2.8, constitute an important environment for shell nouns. Corpus Tool can accommodate these combinations and the network can be extended in delicacy to account for different types of identifying intensive Process clauses etc. The subsystems for identifying intensive Processes are derived from

Halliday (1988/1993), Martin (1992) and Halliday and Matthiessen (1999; 2004); however, building the network was not a straightforward operation due to divergences in interpretations of certain kinds of relational Process clauses. A brief discussion of the differences precedes an explanation of the systems network used to tag relational Process clauses in this thesis.

Within identifying intensive relational processes, Halliday and Matthiessen (2004) discriminate between general verbs and several subtypes. Most of these subtypes correspond to Martin's (1992) taxonomy although some subtypes are conflated (table 4.5).

Halliday & Matthiessen (2004:234-5; 238)	Martin (1992)
Neutral ↘ <i>be</i> Exemplification ↘ <i>exemplify, illustrate</i> Equation ↘ <i>equal add up to make</i>	Single agency: Exemplifying ↘ <i>exemplify, illustrate, instance</i> Exhausting: equality: equation ↘ <i>equal add up to, make, amount to</i>
Equivalence ↘ <i>correspond to</i> Symbolization: Symbol ↘ <i>express, signify, realize, spell, stand for, mean</i> Symbolization: Significance ↘ <i>represent, constitute, form</i> Symbolization: Sign ↘ <i>mean, indicate, suggest, imply, show, betoken, mark, reflect</i>	Exhausting: equality: wording ↘ <i>translate</i> Exhausting: signification: realise ↘ <i>express, signify, mean, realise, write, etc</i>
Role-play ↘ <i>play, act as, function as, serve as</i> Naming ↘ <i>be</i> Definition ↘ <i>be</i> Kind/part ↘ <i>comprise, feature, include</i> Demonstration ↘ <i>show, suggest, indicate</i> Assignment: neutral ↘ <i>make</i> Assignment: elaborating ↘ <i>elect, choose (as), dub, name, pronounce etc</i> Assignment: projection ↘ <i>think, consider, prove, call declare</i>	Exhausting: signification: manifest ↘ <i>stand for, symbolise, imply, reflect, etc</i> Exhausting: signification: connote ↘ <i>indicate, suggest, evoke, reveal, betray, etc</i> Exhausting: role ↘ <i>play, act, portray etc</i>
	Double agency: Assigning ↘ <i>make, name, call, etc</i> Proving ↘ <i>prove, attest, confirm etc</i>

Table 4.5 Verbs serving as intensive identifying relational Processes, adapted from Halliday and Matthiessen (2004) and Martin (1992)

There is a certain amount of ambiguity and overlapping in Halliday and Matthiessen's taxonomy, partly the result of combining the subtypes listed and exemplified on pages 234-235 and the tabulated classification on page 238. For example the subclass 'demonstration' does not appear in the table on page 238, so *indicate* appears both as 'demonstration' and 'sign'. The class boundaries are certainly not hermetic, and verbs can function in more than one way, therefore appearing in more than one class. Nonetheless, the distinction between 'demonstration' (*show,*

suggest, indicate) and 'sign' (*mean, indicate, suggest, imply, show, betoken, mark, reflect*) seems somewhat tenuous. I shall propose an alternative view below.

There is also some ambiguity about exemplifying identifying intensive Process clauses. Martin (1992: 280) defines them as Processes in which the Value resembles the nominal Attribute of an attributive clause, i.e. it is typically an indefinite nominal group, but the Process is reversible. He gives the following example (4.3).

4.3 Cricket's a good sport to watch.

A good sport to watch is cricket. (Martin, 1992: 280)

4.4 Those missiles constitute a threat to our security. (Halliday and Matthiessen, 2004: 237)

Example 4.4 is from Halliday and Matthiessen, and it is clearly reversible by means of the passive. Yet, elsewhere in Halliday and Matthiessen, an example very similar to 4.3 – involving two shell nouns – is interpreted as attributive (4.5), not identifying.

4.5 The **fact** [[that Lear never even alluded to that at the end]] is a **sign** [[that he didn't learn very much through the course of the play]]. (Halliday & Matthiessen, 2004: 226)

4.5' A **sign** [[that he didn't learn very much through the course of the play]] is the **fact** [[that Lear never even alluded to that at the end]].

It is to be noted that 4.5 and 4.5' do not mean exactly the same thing, and this may be behind the decision to classify 4.5 as attributive rather than identifying. As an intensive identifying clause, the difference in meaning stems from the relation between Value and Token. The two are stratally distinct; Token represents "the lower 'expression' and Value the higher 'content'" (Halliday & Matthiessen, 2004:230).

A further difference as to whether verbs are functioning as attributive or identifying Processes is found for the verb *confirm*, which Halliday and Matthiessen consider to be attributive, but Martin, identifying. As previously noted, it is, of course, possible for a verb to be at home in more than one category (Halliday & Matthiessen, 1999; 2004). For example, *think, consider* and *prove* can function as an attributive or identifying Process (Halliday & Matthiessen, 2004: 238).

Notwithstanding these differences in classifications, both Halliday and Matthiessen, and Martin classify *prove* and *confirm* as a relational Process with assignment/double agency only. Consider example 4.6 taken from Martin (1992: 227), where *confirm* is a metaphorical realisation of internal cause, ‘x so I think/say y’ (Halliday, 1988/1993).

- 4.6 - Let’s go out to dinner.
 - That confirms you finally got paid.

Martin (1992: 227,228) unpacks the structure of *That confirms you finally got paid* as ‘*the fact that you are inviting me out to dinner confirms the fact that you got paid to be a fact*’ (figure 4.13).

The fact that you are inviting me out	confirms	the fact that you got paid	to be	a fact
Assigner	Process	Token	Process	Value

Figure 4.13 Identifying intensive clause with agency

In addition to *confirm* and *prove*, other verbs such as *demonstrate*, *imply*, *show*, *suggest*, and *indicate* can also realise internal cause metaphorically (Halliday, 1988/1933). I would argue that such verbs can be seen on a cline of caused modality with maximum proof at one end and minimum proof at the other (figure 4.14), and that due to the absence of a third Participant, they can all function with single agency.

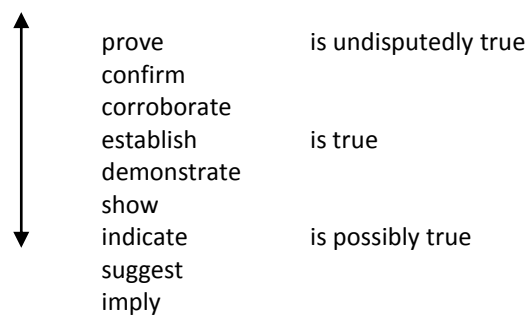


Figure 4.14 Verbs realising intensive identifying relational processes graded by degree of commitment to proof

Consider the following variations to an example of an identifying intensive clause with single agency taken from Halliday and Matthiessen (2004: 227).

- 4.7 [Token] Hence evidence from numerous studies above clearly **shows** [Value] that the Genetic explanations cannot be true, that differences in IQ among races and groups is because of genetic differences.

- 4.7' [Token] Hence evidence from numerous studies above clearly **proves** [Value] that the Genetic explanations cannot be true, that differences in IQ among races and groups is because of genetic differences.
- 4.7'' [Token] Hence evidence from numerous studies above clearly **suggests** [Value] that the Genetic explanations cannot be true, that differences in IQ among races and groups is because of genetic differences.

While 4.7' increases the commitment to caused modality, 4.7'' tones it down. The view that relational clauses can operate along a cline similar to that of modalisation is in keeping with Hood (2010: 104), who discusses mental and verbal projecting processes at the stratum of discourse semantics. She proposes that they enter into the system of GRADUATION: FOCUS: FULFILMENT: actualisation by virtue of the fact that they contrast with completion (see figure 2.5, section 2.3.2, p 16). Furthermore, this cline of caused modality can be seen as agnate to the fact nouns of the 'evidence' type, the nominalised process representing the final step in the evolution of grammatical metaphor (see figure 2.14, section 2.8, p 29). Table 4.6 illustrates this relation. Note, however, that construing the clause with a fact noun may result in an attributive clause rather than identifying.


	Verbs functioning as identifying intensive Processes	Fact nouns (Evidence type)
	proves	is proof that
	confirms	is confirmation that
	corroborates	is corroboration that
	demonstrates	is a demonstration that
	shows	is a sign that
	indicates	is an indication that
	suggests	is a suggestion that
	implies	is an implication that

Table 4.6 Cline of caused modality in relational Processes and fact nouns of the 'evidence' type

The network system for relational clauses is developed in light of the previous discussion. Accordingly, for the purposes of the analysis of shell nouns, the system for relational processes reflects the caused modality continuum (figure 4.14) in the subsystem SIGN. Three levels of commitment – high, median, low – are included. The system network for tagging Processes is shown in figure 4.15.

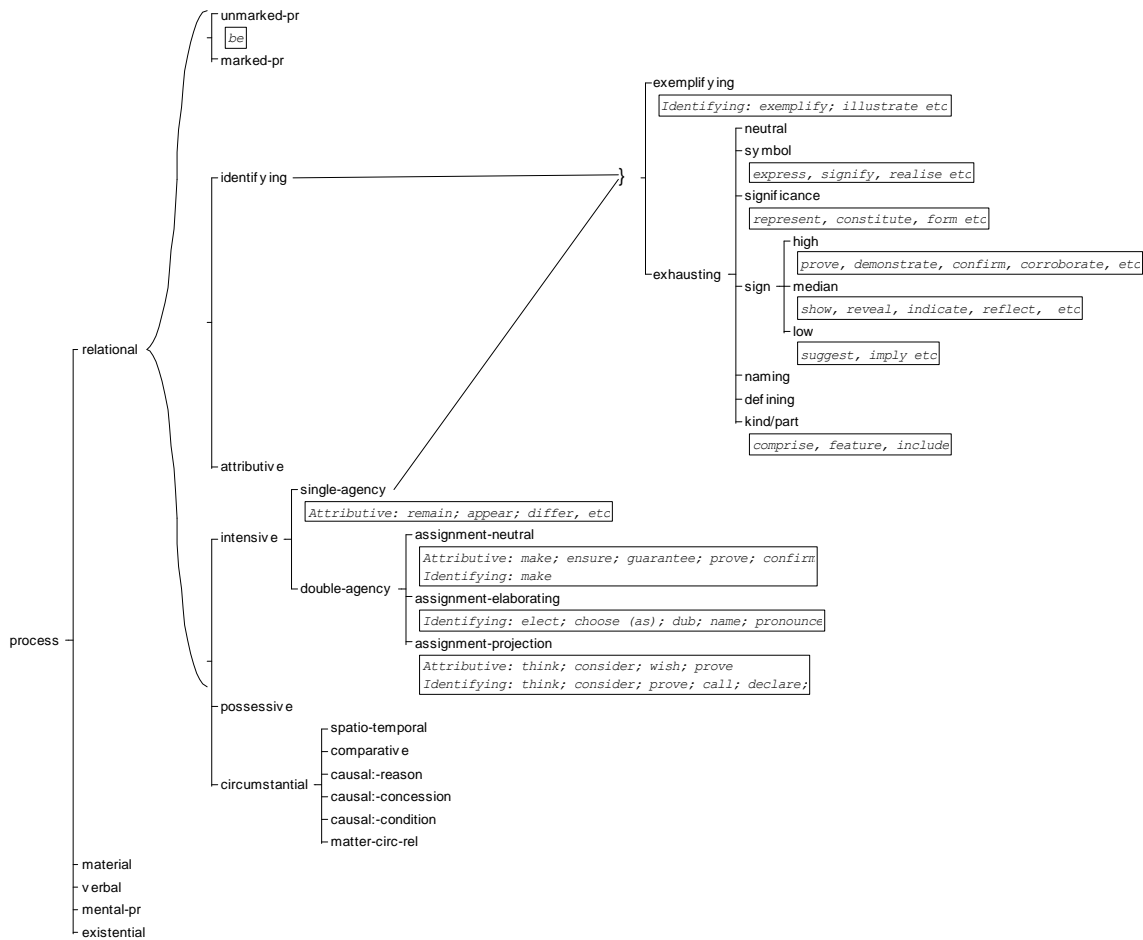


Figure 4.15 System network for tagging PROCESS

A second methodological decision concerns intensive relational Process clauses with an indefinite nominal Subject or Complement that are reversible. As noted above, they have been interpreted as attributive (ex 4.5) or as identifying of the exemplifying subtype (ex 4.3). In order to avoid arbitrary decisions about the interpretation, clauses with the above features are classed as identifying. Figures 4.16 and 4.17 illustrate the analysis of both unmarked and marked intensive relational clauses.

Thus	one	objective	of this area of research	is	to develop indices with high construct validity	as it relates to use of services. (P_Pov)
	Value				Pro: Rel: Token (not tagged)	
					Unmarked	
					Identifying	
					Intensive:	
					Exemplifying:	
					Single	
					agency	
	Deictic	Thing: mental [shell noun]	Qualifier			

Figure 4.16 Transitivity analysis showing unmarked identifying intensive relational Process clause

The Cluster analysis demonstrates			that these two decades were characterised by much cross- institutional subnet membership (P_Net)
Token			Pro: Rel: Marked Identifying Intensive: Exhausting: Sign: high single agency
Deictic	Classifier	Thing: mental [shell noun]	Value (not tagged)

Figure 4.17 Transitivity analysis showing marked identifying intensive Process relational clause

It must be remembered, however, in the example given in figure 4.17 the shell noun, **analysis**, which construes the Token, is not lexicalised by the fact clause that construes Value; the meaning encapsulated by the shell noun must be sought elsewhere in the text, as argued in chapter 3, section 3.4.2.

4.5.2. Tagging the interpersonal metafunction in the lexicogrammar

The interpersonal analysis at the stratum of the lexicogrammar aims to identify the interpersonal functions that shell nouns enact at the stratum of the lexicogrammar. The systems network for this analysis is a straight-forward application of Halliday and Matthiessen (2004) with a couple of modifications to ensure that all shell nouns are tagged and accounted for. Because some shell nouns do not enact an interpersonal function at clause rank or lie outside the system of MOOD, an entry system for NO INTERPERSONAL FUNCTION accounts for this possibility. The systems network is shown in figure 4.18.

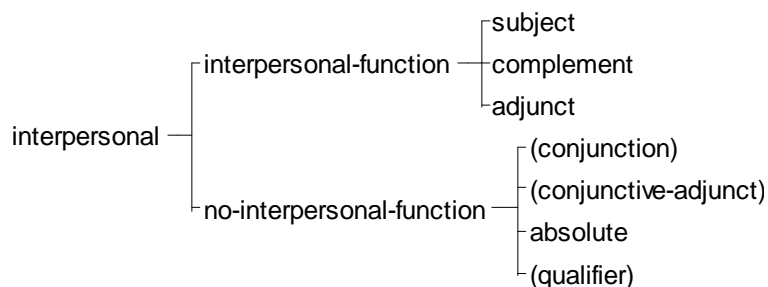


Figure 4.18 System for tagging the interpersonal metafunction in the lexicogrammar

Using this system network, elements in which the shell noun enacts an interpersonal function at the rank of clause are tagged as Subject, Complement or Adjunct. All other elements containing shell nouns are tagged NO INTERPERSONAL FUNCTION and then classified according to the subsystems. For example, at group rank, shell nouns appearing in prepositional phrases as Qualifiers are considered to have an ideational function only. These shell nouns are captured in the analysis of the ideational metafunction, previously detailed. At clause rank, conjunctive Adjuncts do not enact an interpersonal function (Halliday and Matthiessen, 2004: 132). These shell nouns are captured in the textual metafunction analysis as textual Themes, as are conjunctions. Minor clauses, which have no mood structure and consist of headings, are tagged Absolute, in accordance with Halliday and Matthiessen (2004: 154). Table 4.7 summarises the situations in which an element containing a shell noun is tagged NO INTERPERSONAL FUNCTION, and examples of analyses of the interpersonal metafunction at the stratum of the lexicogrammar follow (figures 4.19 – 4.21).

	Example	Coding
Conjunction	with the purpose of	NO INTERPERSONAL FUNCTION: (conjunction)
Conjunctive Adjunct	In view of this result	NO INTERPERSONAL FUNCTION: (conjunctive Adjunct)
Heading	Proof	NO INTERPERSONAL FUNCTION: Absolute
Qualifier	of these results	NO INTERPERSONAL FUNCTION: (Qualifier)

Table 4.7 Elements tagged NO INTERPERSONAL FUNCTION

An advantage of allowing different weights for each variable	is that similar information from manifest variables does not contribute more than once to the index.
Subject	

Figure 4.19 Interpersonal analysis showing shell noun as Subject (P_Pov)

A classic example of how difficult it is to measure performance for managers in the non-profit sector is	the case of a manager of a nursing home	(Weisbrod and Schlesinger (1986) (P_CEO)
	Complement	

Figure 4.20 Interpersonal analysis showing shell noun as Complement

For a variety of reasons	non-profits provide an interesting context in which to study the pay of managers in this period. (P_CEO)
Adjunct	

Figure 4.21 Interpersonal analysis showing shell noun as Adjunct

4.5.3. Tagging the textual metafunction in the lexicogrammar

The system developed to tag the textual metafunction seeks to identify the textual functions that shell nouns compose at the stratum of the lexicogrammar. The network for tagging the textual metafunction is based on Halliday and Matthiessen (2004) and Fries (1994). An initial distinction between shell nouns composing a textual function and those without is made. Elements are considered not to have a textual function at the rank of the clause if they are functioning either above the clause, as a heading, or below the clause, as Qualifier within a group. Ranking elements composing a textual function are further classified as contributing to Theme selection, N-Rheme or other. The network to this level of delicacy is shown in figure 4.22

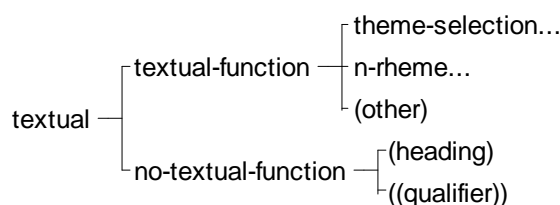


Figure 4.22 Basic systems for the analysis of the textual metafunction at the stratum of the lexicogrammar.

Following Halliday and Matthiessen (2004), the system network for THEME SELECTION allows for the analysis of Theme in two domains: that of the clause and that of the clause complex. Halliday and Matthiessen (2004: 392-394) discuss the implications to the logogenesis of a text stemming from systemic selections of the sequence of the clauses in a hypotactic clause nexus. When the hypotactically bound clause realises the first clause in the clause nexus, there are two thematic domains to consider: that of the clause nexus and that of each clause within the clause nexus. I shall first describe the system used to tag Themes operating at the clause rank, then the system for tagging clauses within the domain of the clause nexus.

The first node, *clause theme selection*, is used to tag Themes operating at clause rank. It allows for multiple Themes, not all of which are obligatory elements of the clause. The subsystem CLAUSE THEME SELECTION classifies a Theme as topical Theme, textual Theme or predicated Theme. Within topical Theme, the system distinguishes between unmarked Theme, i.e. conflated with Subject (figure 4.23), and marked topical Themes, e.g. Adjunct (figure 4.24).

His only alternative	is to remain unmatched. (S_Rand)
CLAUSE THEME SELECTION: TOPICAL THEME: unmarked	

Figure 4.23 Textual analysis showing elements tagged unmarked Theme

For a variety of reasons	non-profits provide an interesting context in which to study the pay of managers in this period. (P_CEO)
CLAUSE THEME SELECTION: TOPICAL THEME: marked	

Figure 4.24 Textual analysis showing elements tagged marked Theme

In a topical Theme, a shell noun can construe Head in the nominal group composing Theme or it can appear within a Qualifier. Because a shell noun in a Qualifier is functioning at group rank, it is not tagged as composing a textual function at clause rank. Rather it is tagged NO TEXTUAL FUNCTION: ((Qualifier)), as illustrated in figure 4.25.

The proof of the above result	reveals that	[[a sufficient condition for a Nash equilibrium in the game (P, ? \sim , P)	is in fact being a Nash equilibrium in every game (P, ?, P).]]	(S-Rand)
CLAUSE THEME SELECTION: TOPICAL THEME: unmarked				
	NO TEXTUAL FUNCTION ((Qualifier))			
		CLAUSE THEME SELECTION: TOPICAL THEME: unmarked:		

Figure 4.25 Textual analysis showing tagging of shell noun in Theme and in Qualifier

Textual Themes are realised by either Conjunctions, such as *in the **event** that*, which bind a clause in a hypotactic clause nexus, or conjunctive Adjuncts that contain an anaphorically referring shell noun, such as *With this **purpose** in mind*. As stated earlier, conjunctive Adjuncts such as *As a **consequence*** are not considered to be realised with shell nouns, so they are not tagged. Conjunctions are termed structural Themes (Halliday & Matthiessen, 2004: 99), but they are also textual, as exemplified by Halliday and Matthiessen's (2004: 101-105) analysis of a biographical text. The system network allows for the tagging of textual Themes as Conjunction or conjunctive Adjunct. One other node in the system of CLAUSE THEME SELECTION enables the tagging of predicated Themes. The system network for THEME SELECTION is shown in figure 4.26.

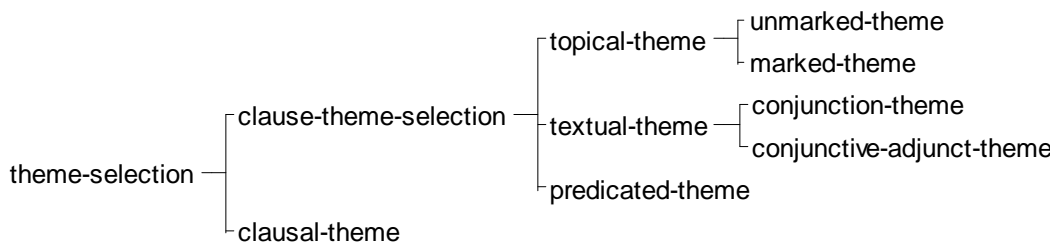


Figure 4.26 System network for THEME SELECTION

At the domain of the clause nexus a hypotactically bound clause in initial position in a clause nexus can be analysed as the Theme of the clause nexus, the free clause composing the Rheme. In this thesis, when a hypotactically bound clause containing a shell noun realises the first clause of the clause nexus, this clause is tagged with the second feature of THEME SELECTION: clausal Theme, exemplified in figure 4.27.

Since one of the clearest **lessons** from the study of deterministic procedures is that understanding such incentives is crucial to understand the behaviour of the market, the paper is devoted to equilibrium analysis. (S_Rand)

Theme selection: clausal Theme

Figure 4.27 Textual analysis showing clausal Theme

The Process of a clause composing a clausal Theme can be finite or non-finite. Finite clausal Themes have been called ‘marked clause-complex Themes’ (Cummings, 2005: 133) while non-finite clausal Themes have been called ‘modifying clausal Themes’ (Francis, 1990 in Fries, 1995a) or ‘marked Sentence Themes’ (Cummings, 2005:151). Furthermore, the function within the clausal Theme of the element containing the shell noun varies. At the domain of the clause it could compose a textual Theme, i.e. appear within a Conjunction (figure 4.28) or topical Theme (figure 4.23 above), or it could appear within the Rheme of the hypotactically bound clause. Examples of analyses of clause nexuses containing clausal Themes are given in figures 4.28 and 4.29.

With the **aim** of understanding the effect of weights, we compared point estimates from weighted and non-weighted data (S_Pov)

THEME SELECTION: clausal Theme
LOGICAL: Relator

Figure 4.28 Textual analysis showing shell noun in Conjunction in clausal Theme in clause nexus

Based on a growing **concern** to induce other perspectives of realities that are directly connected with a possible social disadvantage of the households,

the following well being dimensions are considered to be relevant: (S_Pov)

THEME SELECTION: clausal Theme

Fig 4.29 Textual analysis showing non-finite clausal Theme in clause nexus

With the system TEXTUAL FUNCTION, non-thematic elements containing shell nouns are tagged either N-Rheme or (other). I have chosen not to analyse shell nouns in terms of the system of INFORMATION STRUCTURE. This decision stems from Halliday and Matthiessen’s (2004) view that information structure is more readily identified in spoken discourse because it is a property not of the clause, but of the information unit, and the two do not necessarily cover the same information. As such, written discourse does not lend itself to a clear analysis of what information is to be considered New. Instead, Fries’ (1994) concept of N-Rheme seems more appropriate.

Fries (1994: 234) defines the N-Rheme as “the last constituent of the clause [...], the newsworthy part of the clause, that is, the part of the clause that the writer wants the reader to remember.” Clearly a shell noun appearing as the final experiential element in a clause cannot be Theme, so it is possible to analyse the two functions within the one textual system. In fact, this approach has been taken by Cummings (2005) and Moore (2011). But a clause may be part of a clause complex, so the network for N-Rheme is extended in delicacy to distinguish among the N-Rhemes of a clause simplex or those that are the last experiential element of the clause complex and N-Rhemes that appear within a clause that is not the final clause in the clause complex. The system for tagging N-RHEME is given in figure 4.30. Because a shell noun could appear within an element that is neither Theme nor N-Rheme, the node ‘other’, opening directly off TEXTUAL FUNCTION, accounts for this possibility.

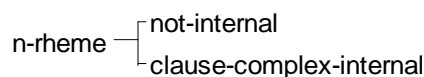


Figure 4.30 System network for tagging N- RHEME

Examples for tagging N-Rheme and ‘other’ are given in figures 4.31 – 4.33.

However, certain unstable matchings can also arrive in this way. (S_Rand)

N- RHEME: not internal

Figure 4.31 Textual analysis showing elements tagged N- RHEME in clause simplex

In particular, let the **probability** that w is matched to a when the profile ^Q is used in
 ~?w[^Q](UPw(v)) be partner at least as good as v (P, ?~, P). (S_Rand)

N- RHEME: clause complex internal

Figure 4.32 Textual analysis showing elements tagged N- RHEME, clause complex internal

It follows from the above **definition** that Q is an ordinal Nash equilibrium when... (S_Rand)

Textual function: other

Figure 4.33 Textual analysis showing elements tagged (other)

As already noted, shell nouns functioning in Qualifiers in an element composing Theme are not tagged within the system THEME SELECTION. Instead they are tagged as having no textual function. Minor clauses, which in this study typically are headings, are similarly tagged as having no textual function, in accordance with Halliday and Matthiessen (2004: 100). The system NO TEXTUAL FUNCTION distinguishes between these two realisations.

The final system for the analysis of shell nouns at the stratum of the lexicogrammar aims to tag the clausal environment in which the shell nouns occur. While many analyses do not extend to embedded clauses (e.g. Fries 1994; 1995a; 1995b; Moore, 2011), this analysis does, so it is necessary to account for both rankshifted clauses as well as ranking ones. The system is shown in figure 4.34.

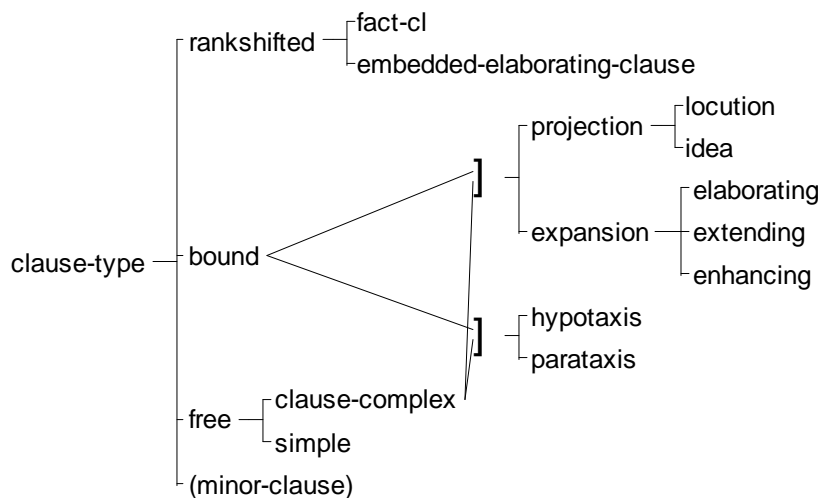


Figure 4.34 System for tagging the clausal environment in which a shell noun appears

This system was developed specifically to describe the clausal environments in which shell nouns appear. It makes it possible to tag when a shell noun appears in a rankshifted clause, either in a fact clause functioning as Token in a relational Process clause, or in an elaborating embedded clause functioning as Qualifier. It also distinguishes between bound and free clauses. A free clause may, of course, be simple or one of a clause complex. The system accounts for this choice. Bound clauses and clause complexes are extended in delicacy to tag both the logico-semantic relationship of the clause nexus and the type of taxis: paratactic or hypotactic. When parataxis and hypotaxis are combined with nesting (Halliday & Matthiessen, 2004: 376), tagging the hypotactic clausal environment takes precedence over tagging the paratactic. This means that a the clausal environment of a shell noun appearing in the dominant clause of a hypotactic clause nexus which is in fact a sub-complex of a paratactic clause nexus will be tagged as `CLAUSE TYPE: FREE: clause complex: TAXIS: hypotaxis` as well as the logico-semantic relationship of projection or expansion (figure 4.35).

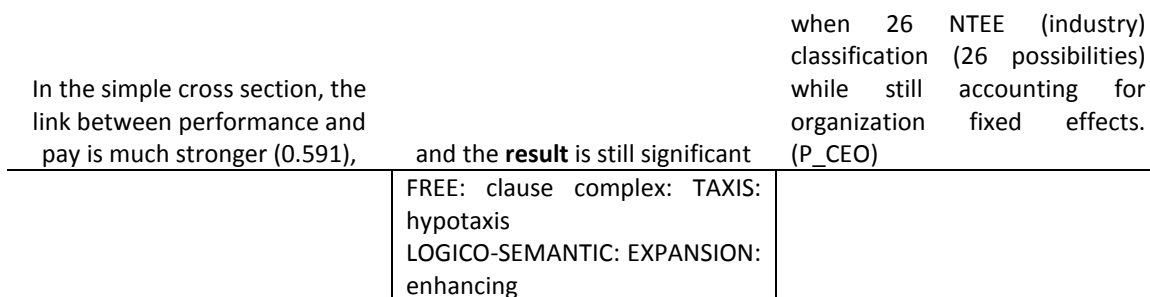


Figure 4.35 Analysis showing tagging of dominant clause in hypotactic sub-complex nested in paratactic clause nexus

This section has described the systems used to tag the corpus at the rank of the lexicogrammar. The next section will describe the systems used to tag the corpus at the rank of discourse semantics.

4.6. Tagging the corpus at the discourse semantics stratum

The qualitative analysis of the functions of shell nouns at the stratum of discourse semantics seeks to identify the ways in which shell nouns contribute to ideational, interpersonal and textual meaning at this stratum. The discourse semantic systems used to tag the corpus are developed with that purpose in mind. They are described in the following section, and examples are given to illustrate the tagging procedures.

4.5.1. Basic systems

Similar to the method used to tag the functions of shell nouns at the stratum of the lexicogrammar, each nominal group with shell noun as Head is tagged separately for the ideational, interpersonal and textual metafunctions. In addition, I propose a further system, SEGMENT, which enables the tagging of the shell noun or a lexicalisation. This system does not play a function in the grammar. Rather its purpose is to enable a more finely tuned functional analysis; cross-classification of the functional elements at the discourse semantics stratum with those of the system SEGMENT make it possible to explore the propensity of shell nouns and their lexicalisations to construe particular ideational functions. The system network to this level is given in figure 4.36.

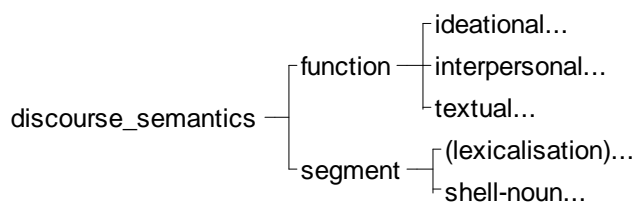


Figure 4.36 Primary systems for tagging at the discourse semantics stratum

Although not shown here, the node shell noun is extended in delicacy to the six classes of shell nouns identified by Schmid (2000): factual, linguistic, mental, modal, eventive, and circumstantial. The system LEXICALISATION is extended to tag the realisation type of the lexicalisation, i.e. whether it is a nominal group with shell noun as Head, a projected clause as Qualifier, a fact clause, etc. (figure 4.37). Although this information is also collected under the textual system IDENTIFICATION, the motivation behind tagging it in the segment is to enrich the ideational analysis. The procedure makes it possible to use the statistical tools of Corpus Tool to analyse the propensity of realisation types to construe a particular field. An example of the tagging procedures for the shell noun and its lexicalisation in the system SEGMENT is given in 4.38 below.

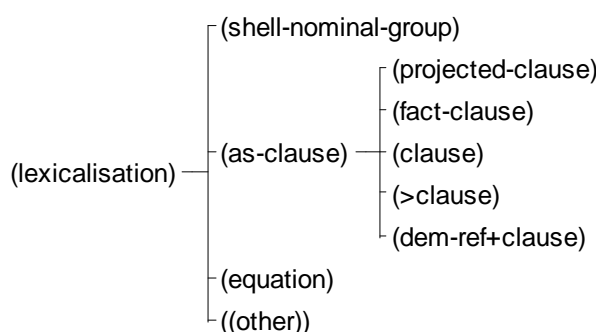


Figure 4.37 Systems network for tagging the lexicalisation of a shell noun

There is a	tendency	to assume that just because actors are linked they must form a cohesive network.	However, this	is not necessarily the	case.(P_Net)
	SHELL NOUN: modal	LEXICALISATION: projected clause			
		LEXICALISATION: dem ref + clause			SHELL NOUN: factual

Figure 4.38 Tagging the elements in the system SEGMENT

Note that in figure 4.38 there are two shell nouns and two lexicalisations. The lexicalisation of **tendency** does not coincide completely with that of **case** although there is some overlapping.

4.5.2. Tagging the ideational function in discourse semantics

In Systemic Functional Linguistics ideation at the discourse semantics stratum can be experiential or logical in meaning. Experiential meaning is realised through activity and implication sequences, where the role of entities and Processes are classified according to taxonomic relations (Martin & Rose, 2003), which construe the field of discourse. Shell nouns, however, have been shown in Chapters 2 and 3 to be not necessarily field dependent. Martin (1992:377-378; 417) sees many shell nouns such as **reason** or **factor** as textual metaphors, which construe a “meta-message relation”, and because they draw on the resources of Deixis and lexical relations, they can be seen as lying at the intersection of textual and ideational systems, but are not tied to logical meaning. This view suggests that they can contribute to both the experiential and textual functions of the discourse semantics systems, and should be tagged in both. The aim of the analysis of experiential ideational meaning in this study is to explore the relation of shell nouns and their lexicalisations to field. On the other hand, logical meaning – logico-semantic relations – is construed through the system of EXTERNAL CONJUNCTION, which relates activity or implication sequences as elaborating, extending or enhancing. As noted in Chapter 3, section 3.5.1, anaphorically referring shell nouns have been found to appear in conjunctive Adjuncts.

The systems network for the ideational function, therefore, is split to identify experiential and logical meanings (figure 4.39). The LOGICAL system opens onto external conjunction. Although external conjunction is a system in its own right, it has not been extended in delicacy in this study.

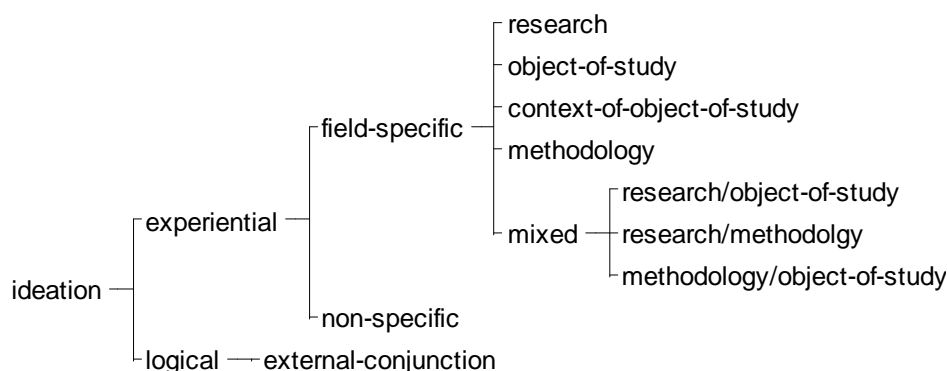


Figure 4.39 Systems network for tagging ideational function in discourse semantics

Figure 4.40 shows tagging of a shell noun within an element of external conjunction. The shell noun is in bold and its lexicalisation underlined.

The founders took opposite camps: Roos and Fisher argued for an open society whereas Schumpeter, Frisch and Bowley preferred a closed centre of excellence with mathematically trained scientists.

As a consequence of his different **view** of the nature of the society, Schumpeter opposed a number of

IDEATION: LOGICAL: external conjunction

names proposed by Fisher. (S_Econ)

Figure 4.40 Analysis showing tagging of elements construing external conjunction

The EXPERIENTIAL system within ideation is designed to elucidate how shell nouns and their lexicalisations contribute to field. The system distinguishes between items that contribute to construing a particular field and those that do not (figure 4.39 above). Hood (2010) has shown that in research article introductions two fields predominate: that of research and that of the object of study. She proposes that the relation between them be seen as one of projection, where the field of research metaphorically projects that of the object of study. She also posits that in other sections of the research article other fields will come to the fore. The subsystem to specify field was based on her work. The subsystem 'MIXED' accounts for sections of texts where more than one field is present in Hood's (2010) relation of metaphorical projection. In each instance, the shell noun and its lexicalisation are tagged for their ideational functions separately. Figure 4.41 illustrates tagging for the ideational function and segment. The shell noun is in bold and the lexicalisation is underlined.

Results indicate a low value for clustering coefficient, meaning a low

average **probability**

IDEATION: EXPERIENTIAL: FIELD SPECIFIC: research	IDEATION: EXPERIENTIAL:non specific
	SEGMENT: SHELL NOUN: modal

of two actors having a common neighbour being themselves connected.

(S_Net)

IDEATION: EXPERIENTIAL: FIELD SPECIFIC: object of study
SEGMENT: LEXICALISATION: projected clause

Figure 4.41 Analysis showing tagging of elements in SEGMENT and IDEATION

In addition to each shell noun and its lexicalisation, modifiers to the shell noun – e.g. Classifiers, Qualifiers – are also tagged if they construe a specific field. Tagging of a field-specific Classifier is shown in figure 4.41 above and tagging of a non-projected Qualifier is shown in figure 4.42. The shell noun is in bold and its lexicalisation is underlined.

A final	limitation	of empirical indices	is <u>that only data from the specific area are used.</u> (P_Pov)
	IDEATION: EXPERIENTIAL:: not field specific	IDEATION: EXPERIENTIAL: FIELD SPECIFIC: methodology	IDEATION: EXPERIENTIAL: FIELD SPECIFIC: methodology

Figure 4.42 Analysis showing tagging of elements at and below clause rank in IDEATION

4.5.3. Tagging the textual function in discourse semantics

The review of the literature has established that shell nouns construe a textual function in three domains. First, they can contribute to the periodic waves of textual prominence (Martin, 1992) in macro-Themes, hyper-Themes or hyper-News; second they can contribute the organisation of the discourse in the system of INTERNAL CONJUNCTION; and third, as nominal groups, they enter into the system of IDENTIFICATION. The systems for tagging the textual function were developed in light of these three domains, which constitute the initial nodes in the textual system, as exemplified in figure 4.43.

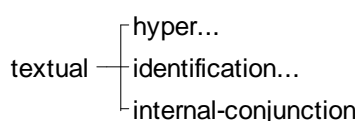


Figure 4.43 Primary systems network for the textual function

The system for tagging the macro-Themes and hyper-Themes/News is based on Martin (1992). It distinguishes between macro-Themes, which are realised as headings, hyper-Themes, which are realised as clauses, and hyper- News (figure 4.44).

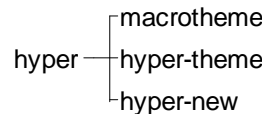


Figure 4.44 Systems network for macro-Themes and hyper-Themes/News

Figures 4.45 and 4.46 illustrate the tagging of a shell noun in a macro-Theme and a hyper-Theme respectively, while figure 4.47 illustrates the tagging of a shell noun in a hyper-New. Where present, the lexicalisation of the shell noun is underlined, and the shell noun is in bold.

Reasons why pay levels for managers of non-profits may be different (P_CEO)
 TEXTUAL: HYPER: macroTheme

Figure 4.45 Analysis showing tagging of macro-Theme

The motivation for the present research stems from several **issues**. First the sole code of
 TEXTUAL: HYPER: hyperTheme
governance in Portugal was established by the Stock Exchange ... (S_CEO)

Figure 4.46 Analysis showing tagging of hyper-Theme

Hence agency problems can be critical in periods of transition from traditional systems of trust to market-based systems of law because they reduce incentives to invest in reputation mechanisms and raise transaction costs.

If the system under analysis does not collapse, some devices must have been promoted to cope with this **paradox**. (S_Net)
 TEXTUAL: HYPER: hyperNew

Figure 4.47 Analysis showing tagging of hyper-New

Shell nouns construing a textual function within the system of INTERNAL CONJUNCTION are tagged as in following example (figure 4.38), where the shell noun **advantage** refers anaphorically.

... Besides these **advantages**, residing agents often overlapped trade activities with
 INTERNAL CONJUNCTION
 administrative tasks, which added political capital to principal-agent relationships. (S_Net)

Figure 4.48 Analysis showing tagging of shell noun in INTERNAL CONJUNCTION

The third domain of the textual function is IDENTIFICATION. The function of the system of IDENTIFICATION is to track Participants in the discourse, and it is somewhat more complex than the other two domains. Martin (1992: Chapter 3) develops an exhaustive system of IDENTIFICATION that cross classifies comparative reference with

generic/specific reference and presenting/presuming reference. It was not deemed necessary to enter into such detail in this thesis because the definition of shell nouns given at the beginning of this chapter – that they are unspecific until the rest of their meaning is retrieved from their co-text – precludes instances of generic reference. Notwithstanding, Martin’s distinction between neutralised reference for headings and effected reference in clauses is relevant given that shell nouns appear in both headings and clauses. In addition, his distinction between presuming and presenting reference is equally relevant as it has been shown that cataphorically referring shell nouns often are composed with a non-specific Deictic. It has also been shown that Post Deictics such as Numeratives, which frequently co-occur with shell nouns, have a textual function (Martin; 1992; Francis, 1994). Martin accounts for these by means of a subsystem, dependent on undirected reference, which identifies the Thing as a member of a superset. However, there are examples in the corpus of possessive Deictics co-occurring with ordinal Post Deictics composing a textual function, so a system of +/- superset operates simultaneously with presuming reference. The system for tagging REFERENCE as a subsystem of IDENTIFICATION is given in figure 4.49.

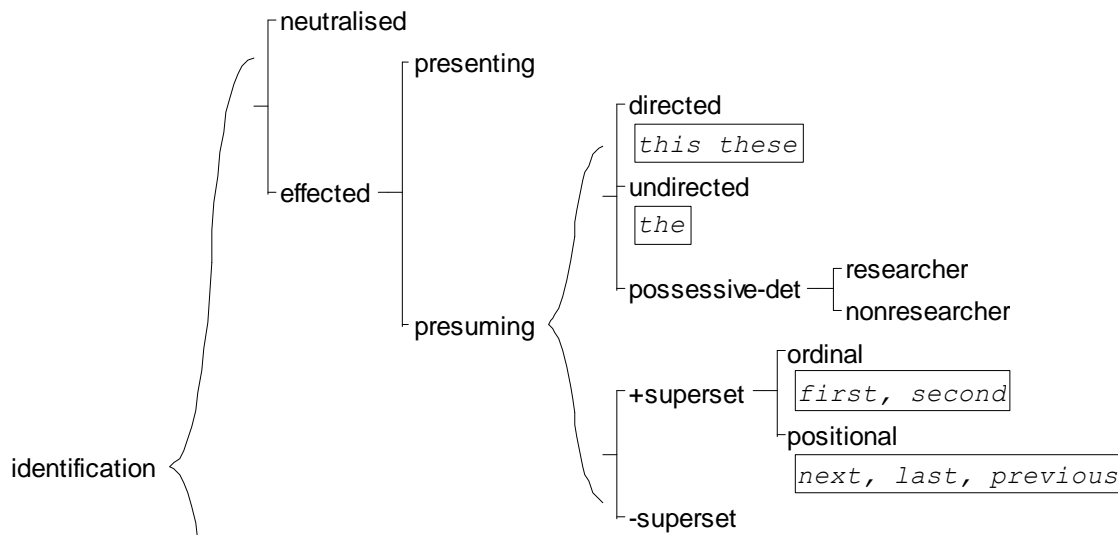


Figure 4.49 Systems network for analysis of IDENTIFICATION

As previously mentioned, to be functioning as a shell noun, the noun must refer to a stretch of discourse realised elsewhere. In order to understand better the nature of the textual relation between the shell noun and its lexicalisation, I use a system that operates simultaneously with that of REFERENCE. This system – LEXICALISATION – is derived, in part, from the work of Biber (1991), who analyses the distribution of anaphoric

forms in a spoken and written corpus, and Biber *et al.* (1998; 1999). The notions of distance between the referent and referring expression, and form of expression of the antecedent are relevant for an analysis of shell nouns. Biber *et al.* measure distance by the number of nominal groups between the reference item and its antecedent. In the case of shell nouns, however, it is more insightful to measure distance in terms of clauses because the antecedent is much more likely to be a clause than a nominal group.

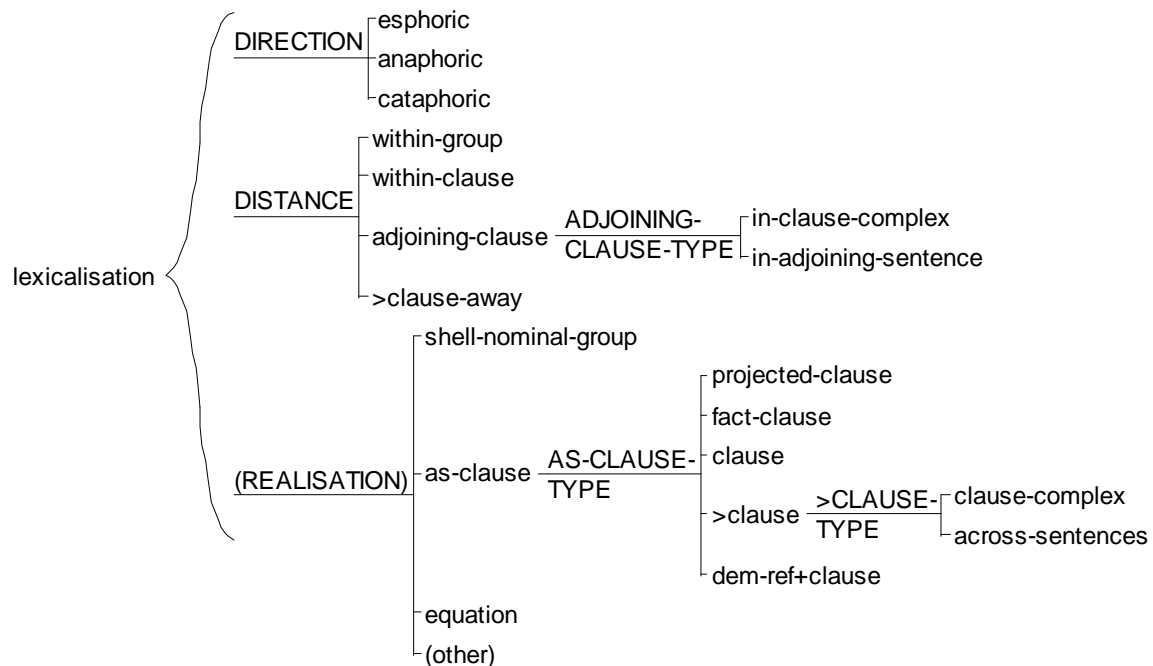


Figure 4.50 Systems network for tagging the relation of the shell noun to its lexicalisation

In her work on the relation between sentence initial extended text reference or a nominal group and its antecedent, Gray (2010) proposes the following categories for the form of the antecedent: local discourse and global discourse. Local discourse is subdivided into simple and complex nominal groups and clause/clause complex. Global discourse is considered to span sentence boundaries and is subdivided into extended preceding discourse and overall text. Although her study was restricted to anaphorically referring items, it suggests a useful category of REALISATION for the stretch of text that realises the lexicalisation of the shell noun. My study further includes the notion of DIRECTION because it encompasses shell nouns whose lexicalisation appears before as well as after the noun. Consequently, the system operating simultaneously with REFERENCE is built to explore the textual relation of the shell noun to its

lexicalisation in terms of DIRECTION, DISTANCE and REALISATION. This system for LEXICALISATION is given in figure 4.50 above.

The system of DIRECTION distinguishes between realisations that are esphoric, i.e. the lexicalisation immediately follows the shell noun in the same clause either as Qualifier or fact clause (Martin, 1992), anaphoric, where the lexicalisation precedes the shell noun in a different clause, and cataphoric, where the lexicalisation follows the shell noun but is in a different clause (Halliday & Hasan, 1976). The categories in the system of DISTANCE are based on elements in the lexicogrammar, ranging from ‘within group’ for when the lexicalisation construes a Qualifier to the shell noun to more than one clause away, ‘>clause’, for when there is a clause or more intervening between the clause containing the shell noun and its lexicalisation. The category of ‘adjoining clause’ is expanded to distinguish between lexicalisations that appear within a clause complex and those that are in a separate sentence. The final system within LEXICALISATION tags the realisation of the lexicalisation in terms of the lexicogrammar: whether it is a nominal group with shell noun as Head, a clause or more, mathematical equation, or another realisation that is not foreseen by the grammar. The latter category includes the few grammatical infelicities found in the Submissions sub-corpus. Simultaneously tagging the nominal group with the shell noun as Head with the systems of IDENTIFICATION and LEXICALISATION makes it possible to identify the typical shell noun patterns posited by Schmid (2000) as well as explore combinations between the textual function and interpersonal function. Figures 4.51 and 4.52 illustrate the tagging of shell nouns for the textual function of IDENTIFICATION. Again the lexicalisation is underlined and the shell noun is in bold.

Finally, once computed, the index lacks a measure of uncertainty. For policy making decisions in particular, **this last feature** may be problematic (P_Pov)

REFERENCE: effected: presuming: directed SUPERSET: +superset: positional LEXICALISATION: DIRECTION: anaphoric DISTANCE: adjoining clause: in adjoining sentence REALISATION: >clause: clause complex
--

Figure 4.51 Analysis showing tagging of element containing anaphorically referring shell noun in IDENTIFICATION

An **advantage** of allowing different weights for each variable

is that the similar information from different manifest variables does not contribute more than once to the index. (P_Pov)

REFERENCE: effected: presenting
 SUPERSET:-superset
 LEXICALISATION: DIRECTION: esphoric
 DISTANCE: within clause
 REALISATION: as clause: fact clause

Figure 4.52 Analysis showing tagging of element containing esphorically referring shell noun in IDENTIFICATION

4.6.4. Tagging the interpersonal function in discourse semantics

The literature review shows that shell nouns often contribute interpersonal meaning to the discourse. Schmid (2000) describes this in terms of a ‘characterising function’, while Francis (1994) notes the role of interpersonally oriented modification to the shell noun, and Charles (2003; 2007) discusses it in relation to the evaluative stance of the writer. In a Systemic Functional Linguistics perspective, interpersonal meaning at the discourse semantics stratum is enacted through the system of APPRAISAL (Martin & White, 2005; Hood, 2010). The codes used to tag the interpersonal function of shell nouns are a straightforward application of Appraisal theory. Because my interest is in the contribution of shell nouns to interpersonal meaning, I restrict tagging to elements that are realised within the nominal group with shell noun as Head. I also include the systemic possibility that the nominal group with shell noun does not carry interpersonal meaning. Hence, there is an initial distinction between the system of APPRAISAL and elements that lie outside or are ‘unaccounted’ for in Appraisal theory (figure 4.53)

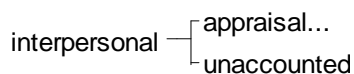


Figure 4.53 Primary systems network for the interpersonal function in discourse semantics

I include this possibility despite Hood’s (2010: 180) argument that ‘fact’ clauses (Halliday & Matthiessen, 2004) such as those projected by shell nouns “contribute to the heteroglossic or multi-voiced nature of engagement in the discourse” because it is possible to identify the source to whom “knowledge of the fact is attributed”. According to Hood, the source may be general or specific, and she gives the following example (4.8)

4.8 The **fact** that writing science is especially difficult for students is widely appreciated (Hood, 2010: 181)

This argument suggests that shell nouns – at least those with an esphorically realised lexicalisation, i.e. the lexicalisation construes Qualifier – may well be inherently heteroglossic. In Hood’s example, the Process ‘is widely appreciated’ clearly attributes the idea to unnamed sources. However, the case is less clear with some anaphorically referring shell nouns such as in the following example (4.9).

4.9 As to the respondents’ general characteristics, they were male (98.4%), with a college degree, an average age of 42, and an average monthly income of 1,396.64 euros. This **profile** leads to an overall definition of the responding administrators as middle-aged, middle class Portuguese males. (S_CEO)

‘This profile’ refers to the profile that the respondents were male, and had a college degree, an average age of 42 and an average monthly income of 1,396.64 euros. While this might be considered a view put forward by the writer, the text appears more monoglossic than heteroglossic as there is no other voice acknowledged. Thus, in this instance, the nominal group with the shell noun **profile** as Head seems to serve merely a textual function, and does not specifically contribute interpersonal meaning. Such instances of shell nouns were tagged INTERPERSONAL: unaccounted. This procedure does not exclude the possibility that anaphorically referring shell nouns can be tagged with an interpersonal function. Indeed, Martin and White (2005: 194) include such an example from Doris Lessing as an instance of CONTRACTION (4.10). In the example the resources of CONTRACTION are boxed.

4.10 Everything is taken to extremes. We all know this, but the fact is seldom taken into account when we try to understand what is going on.

In example 4.10 ‘the fact’ is functioning as a shell noun and its lexicalisation is the preceding clause, which itself contains an instance of text reference; ‘the fact’ refers anaphorically to the proposition that ‘We all know that everything is taken to extremes’, and it functions as an overt intervention of the writer to preclude other views, hence is contractive.

Shell nouns, therefore, can contribute to the dialogic nature of the text by means of the system of ENGAGEMENT, which was presented in figure 2.6, chapter 2, p.17.

An example of tagging a nominal group with shell noun as Head within the ENGAGEMENT system is given in figure 4.54. The shell noun is in bold and its lexicalisation underlined.

The extensive examination of the survey’s validity, reliability and generalizability leads to that there is nothing in the evaluation to suggest that it is invalid or unreliable. (S_CEO)

the **inference**

ENGAGEMENT: HETERGLOSS: Expand: entertain

Figure 4.54 Analysis showing tagging of element with shell noun in ENGAGEMENT

While linguistic shell nouns such as **argument** or **proposal** and mental shell nouns such as **knowledge** or **belief**, which correspond to nominalisations of linguistic or mental Processes, clearly contribute to the heteroglossic nature of the text and are coded within the ENGAGEMENT system, a consequence of this practice is that it does not distinguish systemically between voices within the academic community such as those of the writer and other academics, and Participants who construe the field of discourse. In the following example the lexicalisation ‘that the residual money cannot be distributed to those in control’ can be attributed by means of the shell noun, **knowledge**, to the people who can afford to contribute more money to pay for higher priced seats rather than to the writer or another academic. Hence, its contribution to heterglossia is questionable.

4.11 Museums and operas are similar in that they practise “voluntary price discrimination” (Hansmann 1980) by offering low-priced seats, but those who can afford to contribute do so with the **knowledge** that the residual money cannot be distributed to those in control. (P_CEO)

However, it is expected that the number of cases of this type will be low, so they can be set aside and tagged INTERPERSONAL: unaccounted rather than within the ENGAGEMENT system.

In some other cases mental shell nouns such as **idea** are not tagged in the ENGAGEMENT system. This procedure follows Martin and White (2005: 249), whose treatment of modulation of obligation (e.g. be a good **idea** to), codes such instances as attitudinal rather than dialogic. When shell nouns are appraised either by inscription or invocation, they are tagged with the system of ATTITUDE, which was given in figure 2.3 (Chapter 2, p 14). Examples of tagging ATTITUDE are given in figures 4.55 and 4.56. In

4.55 the shell noun **result** is appraised as ‘uninformative’, and in 4.56 the evaluative element is infused in the shell noun **desire**. In the examples the shell noun is in bold and the lexicalisation underlined.

Concluding, $1 = \text{“? F [Q](U?P f (SG))} > \text{“? F [Q](U?P f (SG))$ and Q is not an ON equilibrium in (P, “?, P).
 is as uninformative as large the set of individually rational matchings may be. (S_Rand)

The above **result**

ATTITUDE: APPRECIATION: valuation EXPLICITNESS: inscribe: isolated ATTITUDE POLARITY: negative
--

Figure 4.55 Analysis showing tagging of shell noun as APPRECIATION

Our choice was shaped by **the desire** to use a range of institutions which

ATTITUDE: AFFECT: EXPLICITNESS: inscribe: infused ATTITUDE POLARITY: positive

would cover merchants in different spheres (P_Net)

Figure 4.56 Analysis showing tagging of shell noun as AFFECT

In some cases there may be some ambiguity in the evaluative resource. For example Martin and White (2005: 67-68) explain how resources from APPRECIATION can be used to invoke a judgement. In other words an assessment of an entity, in particular nominalised Processes, can imply a token of judgement of the person who carried out the action. Similarly, assessment of a person using resources from the system of JUDGEMENT can imply APPRECIATION of the accomplishment. Martin and White use double coding – one inscribed, the other invoked – to cover such situations. Hood (2010) also finds tokens of Attitude in academic writing when the prosodic value of inscribed Attitude spreads to the surrounding discourse. Following Hood and Martin and White, I recognise invoked attitudinal assessments. An example of coding for invoked Attitude is given in 4.57. The negative value spreads from the inscribed negative value in ‘criticism’, and it is propagated by the resources from ENGAGEMENT that render the discourse contractive (not ... rather).

The concept rose criticism not for what it intends to apprehend – the connection between institutions, shared believes and economic performance – rather for its promoters’ **intention** to include

ATTITUDE: APPRECIATION: reaction EXPLICITNESS: invoked ATTITUDE POLARITY: negative
--

the issue in the range of “capital” theoretical, implications given the similarities with other sort of capital, namely human capital. (S_Net)

Figure 4.57 Analysis showing tagging of element with shell noun as invoked ATTITUDE

The third system to contribute interpersonal meaning at the discourse semantics stratum is GRADUATION. As mentioned in Chapter 2, the system of GRADUATION makes it possible to grade attitudinal meanings as more or less. Following Hood (2010) shell nouns such as **attempt** or **success** can be seen to grade meanings of completion, and **possibility** and **probability** can grade meanings as being more or less actualised. Examples 4.58 and 4.59 illustrate instances of the tagging of shell nouns operating within the system of GRADUATION.

We assume that each individual has preferences over the other side of the market and the **prospect** of being unmatched (S_Rand)

GRADUATION: FOCUS: FULFILMENT: actualisation downscale
--

Figure 4.58 Analysis showing tagging of element with shell noun as downscaled actualisation

Such a criticism, however, still instigates **intellectual efforts** to submit the institutional

GRADUATION: FOCUS: FULFILMENT: completion downscale

level of an economic system to quantification (S_Net)

Figure 4.59 Analysis showing tagging of element with shell noun as downscaled completion

In addition to shell nouns that grade meanings of focus, other elements operating within the system of GRADUATION that could function to invoke an attitudinal reading were tagged only if they appeared within the nominal group with a shell noun as Head. This procedure is exemplified in 4.60 and 4.61.

A classic **example** of how difficult it is to measure performance for managers in

GRADUATION: FOCUS: VALEUR: authenticity upscale

the nonprofit sector is the case of a manager of a nursing home. A manager could be paid ... (P_CEO)

Figure 4.60 Analysis showing tagging of element within nominal group with shell noun as Head as

GRADUATION: FOCUS

There is a host of other Potential **ways to measure**

GRADUATION: FORCE: QUANTIFYING: an entity: mass/presence upscale	GRADUATION: FOCUS: FULFILMENT: actualisation downscale
--	---

performance in nonprofits. (P_CEO)

Figure 4.61 Analysis showing tagging of element within nominal group with shell noun as Head as

GRADUATION

4.7. Summary of Chapter 4

This chapter has described the methodology used to carry out the study of shell nouns in a Systemic Functional Linguistics perspective. The parallel corpus comprises 5 published articles and 5 article submissions by Portuguese academics writing in the medium of English in the field of Economics. The study uses a strict definition of shell nouns (Schmid, 2000), and analyses them functionally using the theory of Systemic Functional Linguistics. The corpus is tagged for the functions of shell nouns using Corpus Tool (O'Donnell, 2008). While in the main the lexicogrammatical systems for tagging are based on Halliday and Matthiessen (2004), the ideational system also draws on the work of Halliday and Matthiessen (1999) and Martin (1992), and the textual system on Fries (1994). At the stratum of discourse semantics, the main sources for the systems used to tag the corpus are Martin (1992) and Hood (2010) for the ideational function, Martin, (1992) for the textual function, and Martin and White (2005) and Hood (2010) for the interpersonal function. The results for the analysis of shell nouns at the stratum of the lexico grammar are presented and discussed in the next chapter, and those of the discourse semantics stratum will be addressed in Chapter 6.

Chapter 5. Results and discussion for shell nouns at the stratum of the lexicogrammar

This thesis uses Systemic Functional Linguistics to study the functions of shell nouns in a parallel corpus comprising published articles in the field of economics and article submissions written by Portuguese academics in the same field. It aims to answer two general questions:

1. In what ways do shell nouns contribute to ideational, interpersonal and textual meaning?
2. Are there significant differences in their use between published articles and article submissions by Portuguese academics in the field of economics?

The analysis is carried out at two strata: that of the lexicogrammar and that of discourse semantics. Using Corpus Tool (O'Donnell, 2008), the shell nouns were tagged and classified according to six basic categories: factual, linguistic, mental, modal, eventive, or circumstantial (Schmid, 2000). In addition, the lexicogrammatical functions of the shell nouns were tagged at clause and group levels such that a qualitative study of the differences between the two sub-corpora could be complemented by a quantitative analysis. This constitutes the basis for the analysis at the lexicogrammatical stratum. The aim of this chapter is to present and discuss the results of these analyses. First I shall present the results to a more global quantitative analysis and then a more qualitative approach that examines the ideational, textual and interpersonal functions of shell nouns at the stratum of the lexicogrammar, comparing differences between the published and submissions sub-corpora. The qualitative analysis is organised around the ideational functions realised by the shell nouns, incorporating the interpersonal and textual functions. It is followed by a summary and discussion of the findings.

5.1. A brief global comparison

Before proceeding to the analysis of the functions of shell nouns in the lexicogrammar, I shall present a brief comparison of the frequencies of shell nouns in the two sub-corpora. The results for the frequency of shell nouns in the two sub-

corpora show that the writers of the published articles use comparatively fewer shell nouns than the Portuguese-speaking writers of the article submissions (table 5.1).

Shell nouns	Corpus 1 (published articles)	Corpus 2 (article submissions)
Total nº	284	403
Normed to 1000 words	7.24	11.22

Table 5.1 Frequency of shell nouns in Published and Submissions

When these results are broken down by type of shell noun, some differences in the frequencies of each type are found (table 5.2).

Feature	Published		Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
SHELL-NOUN-TYPE	N=284		N=403					
factual	36.3	103	27.3	110	2.51	+++	6.27	+++
linguistic	15.1	43	19.4	78	1.43		2.04	
mental	24.6	70	24.8	100	0.05		0.00	
modal	7.4	21	13.9	56	2.67	+++	7.08	+++
eventive	3.2	9	4.5	18	0.86		0.74	
circumstantial-shell-noun	13.4	38	10.2	41	1.30		1.68	

Table 5.2 Distribution of shell noun types in Published and Submissions

The results show that there is a similar use of linguistic, mental, eventive, and circumstantial shell nouns in the two sub-corpora; however, there are differences in the frequency of factual and modal shell nouns. These differences are significant to the 98% level according to the statistical tests of T statistics and Chi Square tests. The published articles are found to have comparatively more factual shell nouns and fewer modal shell nouns than the article submissions. The most commonly used factual shell nouns in the Published sub-corpus are **fact** (20), **result** (8), **results** (8), **case** (7), and **problem** (5) while in the Submissions sub-corpus they are **results** (16), **result** (11), **example** (9), **fact** (8) and **problem** (8). For linguistic shell nouns they are **remark** (6), **metaphor** (3), **claim** (3), and **question** (3) in the Published, and **proposition** (22), **proposal** (9), **suggestion** (7), **remarks** (5), and **conclusion** (4) in Submissions. For mental shell nouns they are **analysis** (6), **idea** (6), and **view** (3) in Published, and **hypothesis** (24), **strategy** (10), **aim** (7), and **solution** (6) in Submissions. For modal shell nouns they are **trend** (5), **probability** (3) and **ability** (3) in Published, and **probability** (16), **ability** (4) and **opportunity** (4) in Submissions. For eventive shell nouns they are **attempt** (2) in the Published and **efforts** (4) and **attempts** (2) in Submissions. Finally,

for circumstantial shell nouns the most frequent ones are **approach** (4), **criteria** (4) and **ways** (4) in Published and **case** (8), **criteria** (7), **approach** (6), and **ways** (5) in Submissions. It is to be noted that in the case of factual, modal, eventive and circumstantial shell nouns, there is considerable overlap between Published and Submissions in the most frequently used nouns.

5.2. Shell nouns at the stratum of the lexicogrammar

In order to shed light on the basic frequency analysis above, a qualitative analysis of the functions of shell nouns is undertaken. The qualitative analysis is based on the ideational, interpersonal and textual metafunctions of Systemic Functional Linguistics (Halliday & Matthiessen, 2004), and the system networks and procedures used to tag the corpus were explained in Chapter 4. Where appropriate, quantitative comparisons will be established between the two sub-corpora over the course of the chapter; a complete breakdown of the frequency of all functions tagged in the corpus and a comparison between the Published and Submissions sub-corpora is given in Appendix A.

5.2.1. Basic functions of shell nouns at the stratum of the lexicogrammar

A broad answer to the questions - What ideational functions do shell nouns construe in the published articles and in the article submissions? What interpersonal functions do they enact? and What textual functions do they compose? - finds that, overall, there is little difference between the distribution of the functions of shell nouns in the two sub-corpora in the experiential and textual metafunctions, but there is a significant difference for the interpersonal metafunction. The distribution of experiential, interpersonal and textual functions of shell nouns at the rank of the clause is given in table 5.3.

		Published		Submissions	
		N	%	N	%
Ideational function	Participant	211	84.8	273	79.4
	Circumstance	38	15.2	65	18.9
	Relator	1	0.4	6	1.7
	Total	250		344	
Interpersonal function	Subject	117	47.0	128	38.2
	Complement	84	33.7	134	40.0
	Adjunct	48	19.3	73	21.8
	Total	249		335	
Textual function	Topical Theme	121	48.0	151	43.8
	Textual Theme	3	1.2	11	3.9
	Predicated Theme	0	0	2	0.6
	N-Rheme	103	40.9	150	43.5
	Other	25	9.9	31	9.0
	Total	252		345	

Table 5.3. Ideational, interpersonal and textual functions of shell nouns at clause rank

The relative share of ideational functions construed by shell nouns is quite similar across the two sub-corpora. Participants – making up over three quarters of the instances - account for the greatest share, followed by Circumstances making up under a fifth. Relators, by contrast, account for less than 2% of the instances (Appendices A.1; A.2). In the interpersonal metafunction, there are differences in the relative shares of the functions between the Published and Submissions sub-corpora. In Published, shell nouns enact Subject, Complement and Adjunct in decreasing order of frequency, but the trend is reversed for Complement and Subject in Submissions. Thus, the likelihood that a shell noun enacts Subject is more marked in the Published sub-corpus than in the Submissions, a difference that attains medium statistical significance. (See Appendix A.3.) In the Published sub-corpus, Subjects account for almost half the instances, Complements around a third and Adjuncts under a fifth. In the Submissions sub-corpus, shell nouns are more evenly distributed across the interpersonal functions, with Subjects and Complements each making up just under two fifths of the instances, and Adjuncts just over a fifth. In the textual metafunction, shell nouns are quite equally distributed between the Theme and Rheme. Topical Themes containing shell nouns account for around half the instances while Textual Themes make up just over 1%. Predicated Themes appear only in the Submissions sub-corpus. Around two fifths of the instances are clause final, composing an N-Rheme, and the remaining instances – less than one tenth – compose neither Theme nor N-Rheme. In other words, although they compose part of the Rheme, they are not the final element with a transitivity function in the clause. Yet, even a slightly deeper analysis shows that in the

Submissions sub-corpus the relative share of marked Themes realised with shell nouns is higher, significant at a high level (Appendix A.4). Other differences between the two sub-corpora are brought to light by means of a more detailed qualitative analysis.

The following description of the functions carried out by shell nouns in the two sub-corpora is organised around the ideational metafunction. I shall first describe the results for shell nouns that construe Participants, then those for shell nouns appearing within circumstantial elements, and finally shell nouns that construe Relators. The analysis of the interpersonal and textual metafunctions carried out by shell nouns will be integrated within the results for the experiential metafunction. Results for the modification of shell nouns – i.e. transitivity functions at group rank rather than clause – will be presented after those of each function: Participants and Circumstances, respectively. To conclude the chapter there is a summary of the results and a discussion.

5.3. Participants realised by shell nouns

Analysis of the transitivity of the clause in which a shell noun construes a Participant shows that with the exception of mental and existential clauses, the distribution of Processes is similar in both the Published and Submissions sub-corpora (table 5.4). Existential Process clauses are found more frequently in the Published corpus while mental Process clauses are more frequent in the Submissions sub-corpus. These results attain weak statistical significance. As expected, most shell nouns are found in relational Process clauses, which account for more than half of the instances. The second most common environment for shell noun Participants is material Process clauses, which make up between a quarter and a third of the instances, while verbal, mental and existential Process clauses account for only a small proportion. The following analysis is organised by Participant functions in the clause by Process type.

PROCESS-TYPE	Published		Submissions		T Stat	Signif.	Chi square	Signif.
	N=226		N=299					
	%	N	%	N				
relational	58.0	131	53.5	160	1.02		1.03	
material	27.4	62	32.8	98	1.32		1.73	
verbal	8.0	18	5.7	17	1.04		1.07	
mental-pr	3.1	7	6.7	20	1.85	+	3.40	+
existential	3.5	8	1.3	4	1.67	+	2.79	+

Table 5.4 Distribution of Process types in clauses in which shell noun construes a Participant

5.3.1. Shell noun as Participants in material Process clauses

In material Process clauses the shell noun is found to construe Actor (5.1), Goal (5.2) or, more rarely, Scope (5.3). In the examples, following Halliday and Matthiessen (2004), the Participant function is given before the Participant containing a shell noun, and, in line with the previous chapters, the shell noun is highlighted in bold and its lexicalisation is underlined. Note that the lexicalisation is not part of the focus of the transitivity analysis at this point.

- 5.1 I also dropped organizations that have negative values for management and general expenses, fund raising expenses, or payment to affiliates. [Actor] The financial **criteria** reduce the sample by 788 observations (P_CEO)
- 5.2 This supports [Goal] our **argument** that Liverpool's metropolitan business networks, as represented here, were more permeable with a bridging quality (Burt, 2004) at the beginning of our period, but coalesced towards more bonding or closed networks towards the end. (P_Net)
- 5.3 we make [Scope] the customary **assumption** that ± » N.0; 1=mi/. (P_Pov)

Examples 5.1- 5.3 show shell nouns functioning as Participants in a ranking clause simplex or clause complex. In 5.1 the shell noun **criteria** compacts prior information and the nominal group construes Actor, composes Theme and enacts Subject. In 5.2 and 5.3, the shell nouns **argument**, which construes Goal, and **assumption**, which construes Scope, also encapsulate the information realised in the lexicalisation, but instead of ranking at a clause, the information is rankshifted to function as Qualifier.

Shell nouns are also found to construe Actor or Goal in clauses that have been rankshifted. For example, they can construe Actor in an embedded fact clause (5.4),

Goal in a clause downranked to construe Qualifier to another shell noun (5.5), or Goal in an embedded elaborating clause (5.6).

- 5.4 After noting quite, correctly, that the statistical requirement of a large number of observations would dictate the collection of data over time, and deducing that it was inevitable [[that [Actor] this **practice** would violate the orthodox ceteris paribus conditions and the static framework of neoclassical price theory]],... (P_Econ)
- 5.5 An effective way [[to address [Goal] these **concerns**]]... (P_Pov)
- 5.6 Times of social change and of market growth promote this kind of structures which are losing effective prior and self-sustaining mechanisms whereas a third party [[that supplies [Goal] **incentives for obligations and expectations being accomplished**]] is not effective yet. (S_Net)

Like 5.1, 5.4 the shell noun **practice**, which construes Actor, enacts Subject and composes Theme, involves anaphoric reference, but in 5.4 the clause – a fact clause – is rankshifted to construe Carrier in a relational Process clause, which is projected by the mental Process *deducing*. Despite these differences in rank, the shell noun in both examples enables the discourse to move forward (Halliday and Matthiessen, 2004; Francis, 1994; Schmid, 2000) by compacting prior information in the semiotic abstraction. Schmid (2000) argues that this use demonstrates the textual linking function of shell nouns. In 5.5 the shell noun **concerns** construes the Goal in an embedded projected clause that functions as Qualifier to a shell noun construing Value in a relational Process clause. The shell noun **concerns** refers anaphorically, and it provides continuity of Theme, serving as an anchor to ground New information being presented to the reader as Theme. This could be considered another manifestation of the linking function described by Schmid (2000). In 5.6 the shell noun **incentives** enables information to be inserted concisely in an embedded elaborating clause, contributing to restrict the meaning of *a third party*. Despite some of the grammatical infelicities in the clause complex, the density of information in the Subject and Theme *a third party* does not make it top heavy or unwieldy. This could well be due to the ability of the shell noun to project. Schmid (2000) refers to this use as a cognitive function of ‘temporary concept formation’ (see Chapter 3, section 3.1.1). These examples illustrate the flexibility of shell nouns to construe Participants in material Process clauses that enter into different ranks of the discourse.

5.7 [Actor] The **fact** that there can be a coincidence between sender and receiver gave the diagonal of the social-matrixes under study a particular meaning. (S_Net)

A second example of temporary concept formation is given in 5.7, which illustrates how the shell noun can be used to condense information that could be realised at clause rank. The rankshifted Qualifier to the factual shell noun **fact** might otherwise have been expressed as a clause in its own right, but such an option would have made the text more longwinded, less concise. Example 5.2 provided another example of temporary concept formation, whereby the clause that lexicalises the shell noun is rankshifted to construe Qualifier. The shell noun **argument** construes Goal, enacts Complement and composes N-Rheme. The effect is to highlight the prominence of the information in the embedded projection.

The results of the analysis show that when the nominal group with a shell noun construes Actor, the nominal group always conflates with Subject and Theme (Appendices B.1; B.2). Notwithstanding this result, the shell noun is more likely to construe Goal than Actor in both sub-corpora. The frequency of shell nouns as Goal is approximately triple that of Actor in the Published corpus and nine-fold in the Submissions corpus (Appendix A.5). As Goal, the nominal group with shell noun as Head typically conflates with Complement and N-Rheme (Appendices B.3; B.4). While both sub-corpora include shell nouns construing Actor and Goal in ranking and rank-shifted elements of the clause and there is little difference between the Published and Submissions in the frequency with which the shell noun construes Actor, significant differences are found for Goal. The distribution of shell noun types construing Actor and Goal are given in tables 5.5 and 5.6.

	Actor in Published		Actor in Submissions					
	%	N	%	N	T Stat	Signif.	ChiSqu	Signif.
SHELL-NOUN-TYPE	N=8		N=5					
factual	0.0	0	20.0	1	0.00		1.73	
linguistic	12.5	1	40.0	2	1.11		1.31	
mental	12.5	1	20.0	1	0.34		0.13	
modal	0.0	0	0.0	0	0.00		0.00	
eventive	12.5	1	0.0	0	0.00		0.68	
circumstantial	62.5	5	20.0	1	1.51		2.24	

Table 5.5 Distribution of shell noun types construing Actor in Published and Submissions

	Goal in Published		Goal in Submissions					
	%	N	%	N	T Stat	Signif.	ChiSqu	Signif.
SHELL-NOUN-TYPE	N=28		N=50					
factual	7.1	2	18.0	9	1.32		1.75	
linguistic	28.6	8	14.0	7	1.57		2.45	
mental	28.6	8	32.0	16	0.31		0.10	
modal	7.1	2	24.0	12	1.88	+	3.46	+
eventive	3.6	1	6.0	3	0.46		0.22	
circumstantial-shell-noun	25.0	7	6.0	3	2.47	+++	5.80	+++

Table 5.6 Distribution of shell noun types construing Goal in Published and Submissions

None of the differences in the findings for Actor attain statistical significance. Analysis of the instances shows that in the Published corpus, the higher proportion of circumstantial shell nouns derives mostly from repeated use of the shell noun **criteria** in P_CEO. All the others represent a single instance in each text. There are no instances of modal shell nouns construing Actor. A full list of the shell nouns that construe Actor is given in Appendices C.1 and C.2.

Analysis of the instances of shell nouns that construe Goal shows that the Published sub-corpus has a higher proportion of circumstantial and linguistic shell nouns and a lower proportion of modal shell nouns than the Submissions sub-corpus (table 5.6). The first two findings attain high and weak statistical significance respectively and the latter weak. The distribution of factual, mental and eventive shell nouns is similar in the two sub-corpora.

In general the higher proportions can be attributed to repeated uses in mainly one text. This is not surprising given the limited number of texts used in the study and the fact that the number of instances is relatively few. For example, in the Published sub-corpus, the linguistic shell nouns appear in three texts (P_Rand, P_CEO and P_Net), and P_Net accounts for five of the eight instances. Four of these are one shell noun: **argument**. The circumstantial shell nouns come mainly from one text in the Published sub-corpus (P_CEO), which employs repeated uses of **shortcoming(s)** and **approach(es)**, but three in the Submissions sub-corpus (S_CEO, S_Econ and S_Pov). Similarly, the higher proportion of modal shell nouns in the Submissions sub-corpus comes mostly from one text (S_Net), and they are made up of repeated uses of the shell nouns **hypothesis**, **incentive(s)**, and **solution(s)**. These findings suggest that the

choice of shell noun may be constrained by personal preferences of the writer and/or the field. A full list of shell nouns construing Goal is given in Appendices C.3 and C.4

In addition to Actor and Goal, in material Process clauses there are four instances in which shell nouns construe Scope: two in the Published corpus and two in the Submissions corpus (table 5.7). Three of the four are modal shell nouns: **trend** (follow the expected **trend** (P_Net)), **challenge** (took up the challenge (S_Econ) and **task** (took its first **task** (S_Econ)); and the other is mental: **assumption** (make the customary **assumption** (P_Pov)). The distribution among the texts for shell nouns construing Scope is given in Appendices C.5; C.6.

	Scope in Published		Scope in Submissions					
	%	N	%	N	T Stat	Signif.	ChiSqu	Signif.
SHELL-NOUN-TYPE	N=2		N=2					
factual	0.0	0	0.0	0	0.00		0.00	
linguistic	0.0	0	0.0	0	0.00		0.00	
mental	50.0	1	0.0	0	0.00		1.33	
modal	50.0	1	100.0	2	1.00		1.33	
eventive	0.0	0	0.0	0	0.00		0.00	
circumstantial-shell-noun	0.0	0	0.0	0	0.00		0.00	

Table 5.7 Distribution of shell noun types construing Scope in Published and Submissions

5.3.2. Shell nouns as Participants in verbal Process clauses

In verbal Process clauses shell nouns appear only at clause rank; there are no examples of shell nouns appearing within rankshifted verbal Process clauses. The results show that the shell noun is more likely to construe Verbiage than Sayer in both the Published and Submissions sub-corpora. The number of instances of Verbiage is approximately ten-fold that of Sayer. Appendices C.7, C.8 and C.9 show the instances of shell nouns across Participant functions in verbal Processes for the two sub-corpora. Examples of a shell noun construing Verbiage are given in 5.8 - 5.11. Again, the Participant function precedes the Participant, the shell noun is in bold, and the lexicalisation, which is not the objective of the transitivity analysis here, is underlined.

5.8 To describe [Verbiage] this **process** of satisfying blocking coalitions, we use the following terminology. (P_Pov)

5.9 Hansman outlines [Verbiage] several **reasons**. The first is that ... (P_CEO)

5.10 In section 3 we state and prove [Verbiage] our **result** that, for weakly responsive preferences and starting from an arbitrary matching, a stable

matching can always be obtained through at least one path of blocking coalitions (P_Rand)

- 5.11 Yermack (1986) demonstrates that larger boards are detrimental to performance, although Bhagat and Black (1999) fail to confirm [Verbiage] the **results**. (S_CEO)

In general, as Verbiage, the nominal group enacts Complement and composes N-Rheme (see Appendices B.5; B.6). The lexicalisation of the shell noun is realised either esphorically, i.e. within the same clause, in this case as Qualifier to the shell noun (5.8, 5.10), or cataphorically in a different clause (5.9). There is, however, one instance in which the Verbiage composes N-Rheme, but the lexicalisation of the shell noun is realised in a preceding clause (5.11). In this example, the Verbiage could have been realised by demonstrative reference *this*. What the shell noun **results** does is to characterise the information as the outcome of an academic study, rather than leave it unspecified. Thus, it can be seen as contributing to the field. Such a use can be considered an example of Schmid's (2000) semantic function (c.f. chapter 3, section 3.1.1.)

The results for the distribution of shell nouns construing Verbiage across the Published and Submissions sub-corpora are given in table 5.8. Factual and linguistic shell nouns make up most of the instances, and it is found that circumstantial shell nouns appear only in the Published sub-corpus. None of the differences attains statistical significance. Interestingly, but not surprisingly, the factual shell nouns are common to both sub-corpora: **result**, **reason**, and **finding**. The circumstantial shell nouns – **way**, **approach**, **process** – all refer to methods of carrying out a research process and the lexicalisation is realised as a projected non-finite clause functioning as Qualifier. There is more variety in the linguistic shell nouns in the Submissions sub-corpus: **conclusions**, **argument**, **suggestion**, **proposal**. The latter three all come from S_Econ. The prevalence of linguistic shell nouns in this text is not surprising as it discusses the history of the evolution of Econometrics as a legitimate field of study in Economics and recounts the debates and correspondence of the economists involved.

Analysis of the Sayer in the verbal Process clauses in which the shell noun construes Verbiage shows that in most cases the Sayer is either the writer(s) of the paper, another researcher, or the product of research, i.e. a semiotic object such as a

particular paper or research in general. It is also found that in S_Econ the Sayer is an economist from the first half of the 20th century, which, as previously noted given the subject matter of the text, is unsurprising.

	Verbiage in Published		Verbiage in Submissions					
	%	N	%	N	T Stat	Signif.	ChiSqu	Signif.
SHELL-NOUN-TYPE	N=10		N=10					
factual	40.0	4	40.0	4	0.00		0.00	
linguistic	20.0	2	40.0	4	0.95		0.95	
mental	10.0	1	20.0	2	0.60		0.39	
modal	10.0	1	0.0	0	0.00		1.05	
eventive	0.0	0	0.0	0	0.00		0.00	
circumstantial	20.0	2	0.0	0	0.00		2.22	

Table 5.8 Distribution of shell noun types construing Verbiage in Published and Submissions

There is only one example of a shell noun construing Sayer in each corpus. (5.12, 5.13), and in both cases it is a factual shell noun whose lexicalisation is realised in a different clause. In all the other instances of verbal Process clauses the nominal group with shell noun as Head construes Verbiage.

5.12 In fact, the set of ordinal Nash equilibria is larger here, as [Sayer] the following **example** demonstrates. (S_Rand)

5.13 Hopefully [Sayer] these basic **findings** can help inform the issues (P_CEO)

5.3.3. Shell nouns as Participants in mental Process clauses

In mental Process clauses the results show that there are no instances of shell nouns construing Sayer in the Published corpus and only four in the Submissions corpus. An example of a shell noun construing Sayer is given in 5.14.

5.14 [Sayer] The **hypothesis** considers that cohesion is a form of social capital. (S_Net)

The paucity of instances in which a shell noun construes Sayer is to be expected given the slippage between relational and verbal and mental Processes (Christie & Cléirigh, 2008), as discussed in chapter 4. It is possible, for example, to interpret 5.14 as a relational Process clause rather than a mental Process clause. However, the shell nouns construing Sayer are either linguistic or mental so it is always possible to trace the shell noun back to a person who is either the writer of the paper or another researcher. For instance, in 5.14, the hypothesis in question is one

put forward by the writer of the paper, and it could have been realised with a possessive Deictic. As a result, the Senser can be seen as a kind of extension of the person who produced the locution or idea, as argued by Halliday and Matthiessen (2004). In addition, the lexicalisation of the shell noun does not construe Phenomenon in the clause, but is realised anaphorically. A paraphrase of the meaning is “the hypothesis that was previously put forward is based on the assumption that cohesion is a form of social capital”. The shell nouns found to realise Senser are **hypothesis**, **proposal** and **interpretation** (Appendix C.10). At any rate, there are more instances of the shell noun construing Phenomenon than Senser.

Examples of mental Process clauses containing shell nouns construing Phenomenon are given in 5.15 – 5.19. The transitivity function is given before the Participant in which the shell noun construes Head of the nominal group, the shell noun is in bold, and the lexicalisation, which is not the objective of the transitivity analysis at this point, is underlined.

- 5.15 Mobility in long-distance trade management has not been correlated to a negative expectation about others’ behaviour. This paper considered [Phenomenon] that **hypothesis** (S_Net)
- 5.16 In this approach it is also possible [[to estimate [Phenomenon] the **probabilities of a household (with a specific well-being profile) belonging to different deprivations groups**]]. (S_Pov)
- 5.17 In the 1920s people at NBER who looked at distributions of price changes noted that they had “tails” that were much too fat to qualify as Gaussian; Mills (1927, p. 336) was perceptive enough to notice that this phenomenon might imply “infinite probable error”. [Phenomenon] These **observations** were persistently ignored by neoclassical critics (P_Econ)
- 5.18 Yet he [Tinbergen] revealed [Phenomenon] an early **disinclination to defend the neoclassical theory of pricing and allocation** (P_Econ)
- 5.19 The problem with this interpretation is [[that it wilfully disregards [Phenomenon] the **fact that the main point of contention between the disputants was the validity of neoclassical economic theory, and not all theory tout court**]]. (P_Econ)

Unlike the verbal Process clauses, but like the material Process clauses, there is variation in the rank of the mental Process clause in which the shell noun construes Phenomenon. In both sub-corpora there are instances of the mental Process clause being rank-shifted to construe Carrier (5.16) or Token (5.19) in a relational Process

clause. Notwithstanding, unlike material Process clauses, in the Published corpus there is more variation with regard to the interplay with Theme. In half the instances (3 out of 6), Phenomeon composes Theme. When Theme, the shell noun invariably refers anaphorically to information in preceding clauses, as in 5.17. As N-Rheme, the lexicalisation of the shell noun is always realised as Qualifier, as in 5.18. In the Submissions corpus there are no instances in which Phenomenon composes Theme, and this constitutes a significant difference between the two sub-corpora. (See Appendix B.8). In addition, unlike in the Published sub-corpus, in the Submissions Phenomenon composing N-Rheme is found to refer anaphorically (5.15) as well as esphorically. Further significant differences are found between the two sub-corpora with regard to the probability that Phenomenon enacts a particular interpersonal function. In the Published sub-corpus, Phenomenon with shell noun as Head is more likely to enact Subject than Complement, but in the Submissions sub-corpus the inverse situation occurs. This difference attains high statistical significance. (Appendix B.7.)

The Senser in these mental Process clauses in which a nominal group with shell noun as Head construes Phenomenon falls into one of four categories: the writer, the reader, 'the paper' or a historical figure. The historical figures – important economists in the history of economics – are all from two texts: P_Econ and S_Econ. As with verbal Process clauses, the prevalence of mental Processes in these texts is easily explained by their subject matter. The former deals with the influence of theory from Physics on the development of Economic theory and the latter with the discussion and debate amongst reputable academics in the field of economics as they sought to make Econometrics a legitimate discipline within the field of Economics.

A comparison of the distribution of shell nouns construing Phenomenon in mental Process clauses in the Published and Submissions sub-corpora is given in table 5.9. An extra shell noun in the count for the Submissions sub-corpus is due to a factual shell noun – **findings** – that is realised within a Qualifier to the Head – **implications** – construing Phenomenon. None of the differences in the relative share of a shell noun type attains statistical significance.

	Phenomenon in Published		Phenomenon in Submissions					
	%	N	%	N	T Stat	Signif.	ChiSqu	Signif.
SHELL-NOUN-TYPE	N=6		N=11					
factual	33.3	2	27.3	3	0.25		0.07	
linguistic	33.3	2	27.3	3	0.25		0.07	
mental	33.3	2	36.4	4	0.12		0.02	
modal	0.0	0	9.1	1	0.00		0.58	
eventive	0.0	0	0.0	0	0.00		0.00	
circumstantial	0.0	0	0.0	0	0.00		0.00	

Table 5.9 Distribution of shell noun types construing Phenomenon in Published and Submissions

While there is some overlap in the classes of shell nouns found to construe Phenomenon in the two sub-corpora, there is no overlap in the actual choice of shell noun. All but one instance in the Published sub-corpus come from one text: P_Econ; the exception – the circumstantial shell noun **process** – is from P_Rand. By contrast, shell nouns are found to construe Phenomenon in four of the five texts in the Submissions sub-corpus. The factual shell nouns include **fact**, **problem**, **example** and **findings**, linguistic shell nouns **observations**, **statement**, **proposal** and **suggestion**, mental shell nouns **hopes**, **disinclination**, **strategy**, **hypothesis** and **implications**, modal shell nouns **need** and **probabilities**, and circumstantial shell noun **process**. Their distribution across the texts is shown in Appendices C11; C.12.

5.3.4. Shell nouns as Participant in existential Process clauses

There are very few instances of shell nouns construing Existent in the corpus. In all the instances Existent conflates with Complement and N-Rheme. (See Appendices B9; B.10.) In each case, the lexicalisation (underlined) of the shell nouns construing Head in Existent is realised esphorically as Qualifier, as in 5.20.

5.20 There is [Existent] **evidence** in the literature on executive compensation in firms that executive pay levels vary by industry. (P_CEO)

5.21 In any event there is [Existent] a **consensus** about the effects of social change on trust, causing “depletion” in the overall level of social capital, described as “an inverted U shape curve between the density of social capital and the level of development”. (S_Net)

Example 5.21 comes from the Submissions sub-corpus, and it is one of the few instances in which there are infelicities of expression. Although somewhat clumsily worded, **consensus** is functioning as a shell noun and must be understood as one if the text is to make sense. This becomes apparent in a rewrite, as in 5.21’.

5.21’ In any event there is [Existent] **consensus** that the effect of social change on trust is to cause “depletion” in the overall level of social capital. The effect can be described as “an inverted U shape curve between the density of social capital and the level of development”.

The rewrite maintains Existent as N-Rheme but reconstrues the lexicalisation of the shell noun **consensus** as a finite clause. In fact, it involves two shell nouns: **consensus** and **effect**. While **consensus** construes the Existent, **effect** construes Value in a relational Process clause that is rankshifted to construe Qualifier to the shell noun **consensus**. In the projected clause the lexicalisation of **effect** – ‘to cause “depletion” in the overall level of social capital’ – construes Token. Shell nouns in relational Process clauses will be described next, but first, the distribution of shell nouns construing Existent in the Published and Submissions sub-corpora is shown in table 5.10. Although differences in the distributions of shell nouns across Existent do not attain statistical significance, overall there is a slightly higher proportion of shell nouns construing Existent in the Published sub-corpus than in the Submissions.

	Existent in Published		Existent in Submissions					
	%	N	%	N	T Stat	Signif.	ChiSqu	Signif.
SHELL-NOUN-TYPE	N=7		N=4					
factual	42.9%	3	0.0%	0	0.00		2.36	
linguistic	0.0%	0	0.0%	0	0.00		0.00	
mental	14.3%	1	50.0%	2	1.25		1.64	
modal	28.6%	2	25.0%	1	0.12		0.02	
eventive	0.0%	0	25.0%	1	0.00		1.93	
circumstantial-shell-noun	14.3%	1	0.0%	0	0.00		0.63	

Table 5.10 Distribution of shell noun types construing Existent in Published and Submissions

Unsurprisingly, the range of shell noun classes used to construe Existent is wider in the Published sub-corpus than in the Submissions. The Published includes factual shell nouns **features** and **evidence**, the mental shell noun **motivation**, modal shell nouns **need**, and the circumstantial shell noun **ways**. The Submissions includes the mental shell nouns **consensus** and **intention**, the modal shell noun **probabilities**,

and the eventive shell noun **attempts**. Appendices C.13 and C.14 show the instances of shell nouns construing Existent across the 10 texts making up the corpus.

5.3.5. Shell nouns as Participant in relational Process clauses

As noted earlier, by far the greatest number of shell nouns are found in relational Process clauses. In both sub-corpora, in identifying relational Process clauses, the shell noun construes Value more often than Token. When the shell noun does construe Token, the Process is, in general, realised by a marked relational Process rather than the neutral *be*. The verb 'be', however, accounts for more than half of the instances of relational Processes overall, but the results show that the Submissions sub-corpus uses relatively more marked relational Processes and fewer neutral 'be' than the Published sub-corpus. This finding attains high statistical significance (see Appendix A.6).

In most cases the lexicalisation of the shell noun is realised either anaphorically (5.22, 5.23) or esphorically (5.24). The lexicalisation may also be realised cataphorically (5.25) though this is much rarer. As in previous examples, the functional label 'Token' precedes the nominal group with shell noun as Head, the shell noun is in bold, and the lexicalisation, which is not the focus of the transitivity analysis here, is underlined.

- 5.22 In the 1920s people at NBER who looked at distributions of price changes noted that they had "tails" that were much too fat to be qualified as Gaussian; (Mills 1927, p. 336) was perceptive enough to notice that [Token] this **phenomenon** might imply "infinite probable error". (P_Econ)
- 5.23 Core et al (1999) analyse the relationship between CEO compensation and board composition, concluding that the board composition explains a significant amount of cross-sectional variation in CEO compensation, after controlling for standard economic determinants of pay. [Token] Their **result** reveals that CEOs earn greater compensation when governance structures are less efficient. (S_CEO)
- 5.24 [Token] The **fact** that itinerant agents could be assigned to assisting resident agents recalls the importance of principals establishing contacts with reliable peers, (S_Net)
- 5.25 [Token] Three main **aspects** confirm their prominence. Firstly, the highest average values of remittances received were delivered by agents having centrality in the network. Secondly, ... Third ... (S_Net)

The marked variants of relational Processes, as in 5.24 and 5.25, are realised by a variety of subtypes. In both sub-corpora the subtype 'sign' is most frequent with the following verbs, presented in decreasing order of frequency, used in the Submissions sub-corpus: suggest, imply, reflect, validate, confirm, prove, indicate, show, recall, establish, demonstrate, and reveal; and in the Published sub-corpus: suggest, demonstrate, highlight, and imply. These relational intensive identifying clauses are all decoding. Construing anaphorically referring shell nouns as Token in a relational Process clause realised with a marked verb, as examples 5.32-5.24 contributes to the unfolding argument of the writer. The ideational selections explicitly construe the reasoning process through which the writer builds his argument.

Comparison between the two sub-corpora shows that there is a difference in their use: the Submissions sub-corpus has a higher proportion of verbs with high caused modality (e.g. validate, confirm, prove) and a lower proportion of verbs with low caused modality (e.g. suggest, imply) than the Published sub-corpus. These differences attain, respectively, medium and high statistical significance. The results for the statistical analysis are available in Appendix A.6. As well as verbs of the 'sign' subtype, there are verbs of the 'significance' subtype: constitute, contend, mean, and assert in the Submissions sub-corpus; and represent in the Published. No other subtypes, i.e. symbol, naming, defining, or kind/part, are found.

Interestingly, when the shell noun construes Token, Value is typically realised by either a nominal group with a shell noun as Head or a nominalisation that does not function as a shell noun as Head, i.e. a material abstraction. For example, in 5.26 the nominal group with the shell noun **results** as Head [Token] refers anaphorically to the results presented in previous sections of the paper, while the shell noun **conclusions** [Value] refers cataphorically. While the clause functions to link the information in the preceding and following clauses, the shell nouns not only help create the link, but interpret the preceding and following text as **results** and **conclusions**, respectively. In this sense, the shell nouns are contributing to maintaining field. In 5.27 the Value is realised by a nominal group with the nominalisation 'utility' as Head. In 5.28 both Token and Value are nominal groups with a shell noun as Head, and the lexicalisations of both shell nouns are realised esphorically as Qualifier. Note that in each case the

relational Process verb (in italics) is a marked variant, not ‘be’. In the examples the Participant function (Token/Value) precedes the nominal group, the shell noun is in bold, and its lexicalisation, which is not the focus of the transitivity analysis here, is underlined.

- 5.26 [Token] The **results** of this research *suggest* [Value] some **conclusions** about Liverpool’s metropolitan business networks, and also about networks more generally. (P-Net)
- 5.27 [Token] The **need** to analyze or summarize multivariate data that are spatially aligned *suggests* [Value] the utility of factor-analytic models that can incorporate spatial covariation. (P_Pov)
- 5.28 Again [Value] the **fact** that membership of these institutions was highly instrumental *is highlighted* [Token] by the **fact** that between 1770 and 1777, Watts invested in fifteen out of thirty-three slave trade voyages with the following members of the Unanimous Club: Joseph Brooks junior, William Crosbie, Alexander Nottingham, William Pole, Charles Pole and Clayton Case. (P_Net)

So far the examples of nominal groups with a shell noun as Head construing Token have all been from intensive identifying relational clauses. While there are no instances of possessive identifying relational clauses in which the shell noun construes Token, shell nouns are found to construe Token in circumstantial relational clauses in both sub-corpora, the circumstantial element being realised in the Process, as examples 5.29 and 5.30 illustrate. As in the preceding examples, the relational Process is in italics, the Participant function (Token) precedes the nominal group, the shell noun is in bold, and its lexicalisation, which is not the focus of the transitivity analysis, is underlined.

- 5.29 [Token] The **fact** that one person enjoys the services of a public good does not *preclude* others from doing the same. (P_CEO)
- 2.30 The governance of non-profit organisations is a theme that has come to prominence recently Birchall and Simmons (2004), Cornforth (2004), Spears (2004). [Token] This **development** *follows* the research in governance of market enterprises observed elsewhere Berglof (1997), Brickley et al (1997). (S_CEO)

Token with shell noun as Head in the nominal group typically enacts Subject, as in the examples 5.22-5.27, and no significant differences are found between the two sub-corpora. Similarly, no statistically significant differences are found for the textual metafunction; Token with shell noun as Head typically composes Theme. (Appendices B.11; B.12.) There are two instances of shell nouns in predicated Themes, both from

the Submissions sub-corpus. In 5.31 the shell noun **efforts** refers anaphorically; its lexicalisation is underlined. In 5.32 the shell noun **feature**, which composes predicated Theme, also refers anaphorically although the lexicalisation is not shown in the example. The predicated Theme selects negative polarity, and sets up a contrastive focus with Value – also containing a shell noun **probability** – in the relational Process clause that follows it.

- 5.31 At least for the first decades of the journal’s existence, Frisch was the sole driving force behind its publication: he set the agenda, corresponded with the authors, asked for articles, was the referee in most cases, discussed the papers and made suggestions, and, finally, decided on publication, changed the notation for coherence and even corrected the galley proofs. He worked immensely hard and it was [Token] his **efforts** that determined the survival and development of the journal. (S_Econ)
- 5.32 But it is not [Token] this **feature** which explains the best performance of their business. [Token] It is rather [Value] the higher **probability of establishing strong ties when the range of options is larger**. (S_Net)
- 5.33 In the 1920s people at NBER who looked at distributions of price changes noted that they had “tails” that were much too fat to be qualified as Gaussian; (Mills 1927, p. 336) was perceptive enough to *notice* that [Token] this **phenomenon** might imply “infinite probable error”. (P_Econ)

The clause in which a shell noun construes Token may be free or bound. In 5.33, for example, the relational clause is projected by the mental Process *notice*.

Table 5.11 shows the distribution of shell noun types construing Token in the two sub-corpora. Factual shell nouns account for most of the instances, followed by mental shell nouns. By contrast, circumstantial shell nouns construing Token appear only in the Published sub-corpus, and this finding attains high statistical significance.

	Token in Published		Token in Submissions					
	%	N	%	N	T Stat	Signif.	ChiSqu	Signif.
SHELL-NOUN-TYPE	N=23		N=38					
factual	43.5%	10	44.7%	17	0.09		0.01	
linguistic	4.3%	1	21.1%	8	1.80	+	3.18	+
mental	26.1%	6	18.4%	7	0.70		0.50	
modal	8.7%	2	10.5%	4	0.23		0.05	
eventive	0.0%	0	5.3%	2	0.00		1.25	
circumstantial-shell-noun	17.4%	4	0.0%	0	0.00		7.07	+++

Table 5.11 Distribution of shell noun types construing Token in Published and Submissions

The factual shell nouns in the Published sub-corpus include: **fact, result(s), case, effect,** and **phenomenon**. In the Submissions sub-corpus they are: **fact, result(s), proof, aspect, findings, advantage, development, elements,** and **profile**. The mental shell nouns include **analysis(es), idea** and **view** in the Published sub-corpus and **strategy, lemma, corollary, hypothesis, solution, conceptualization,** and **perspective** in the Submissions. The linguistic shell nouns consist of **description** in the Published sub-corpus, and **proposition, questions, remarks, suggestion,** and **comment** in the Submissions. The modal shell nouns **need** and **probability** are common to both sub-corpora, but **ability** is found only in the Submissions. Finally, circumstantial shell nouns, which are only present in the Published sub-corpus, include **constraint** and **way**. See Appendices C.15 and C.16 for their distribution across the texts.

Of all participants, Value is the one which attracts the greatest use of shell nouns in both sub-corpora; it accounts for more than a quarter of the instances in which a shell noun construes a participant. Yet, while the overall frequency of shell nouns construing Value is similar in both sub-corpora, there are marked differences in the realisations of the identifying relational clause.

Intensive identifying relational clauses with a shell noun construing Value give rise to two of the patterns identified by Schmid: Th-N *be* cl and Th *be* N, although the latter pattern can also be realised by an attributive relational Process clause. It is therefore expected that a great number of the instances will be of one or the other of these patterns. In fact, of the 65 instances in which a shell noun construes Value in the Published sub-corpus, Th-N *be* cl makes up 42, i.e. just under two thirds. By contrast, the pattern accounts for only one third - 22 out of 66 - of intensive identifying relational Process clauses containing shell nouns as Value in the Submissions sub-corpus. In this lexicogrammatical pattern Value conflates with Theme and Subject (5.34 – 5.40), and Token is typically realised by a fact clause, which also constitutes the lexicalisation of the shell noun (underlined). Note that in 5.35 the lexicalisation could be reconstrued as a non-finite fact clause ‘to be a Nash equilibrium in every game’.

5.34 Since [Value] one of the clearest **lessons** from the study of deterministic procedures is [Token] that understanding such incentives is crucial to understand the behaviour of the market, (S_Rand)

- 5.35 The proof of the above result reveals [[that [Value] a sufficient **condition** for an ordinal Nash equilibrium in the game (P,?~,P) is in fact [Token] being a Nash equilibrium in every game (P,?,P)]].
- 5.36 [Value] One of the **reasons** for centralizing this market, and later for reorganizaing it, was [Token] that market outcomes did not seem to be 'stable' (P_Rand)
- 5.37 [Value] The primary **motivation** for our proposed methodology, therefore, is [Token] to formalize index development in the context of a model ... (P_Pov)
- 5.38 [Value] A major **consideration** in both policy and research is [Token] whether model-based indices lead to increases in validity ... (P_Pov)
- 5.39 Consequently, [Value] the **priority** was [Token] to set up the assemblies of econometricians. (S_Econ)
- 5.40 However [Value] the final **conclusion** is [Token] that the situation is more complex than our conceptualization suggests ... (S_CEO)

5.38 illustrates an identifying intensive relational Process clause of the exemplifying type. It is possible to reconstrue it as 5.38'.

- 5.38' [Token] Whether model-based indices lead to increases in validity is [Value] a major **consideration** in both policy and research (P_Pov)

Although these identifying intensive relational Process clauses appear as bound clauses in a hypotactic clause nexus (5.34) and in embedded fact clauses (5.35), most instances are clause simplexes (5.39), or a free clause in a hypotactic clause nexus (5.37, 5.38) or in a paratactic clause nexus (5.40). Despite the prevalence of shell nouns in this function, only a few shell nouns are common to both sub-corpora: **conclusion** (linguistic); **purpose** (mental); and **way** (circumstantial). In the Published sub-corpus the shell nouns include: **limitation, criterion, advantage, problem, reason, example, thing, aspects, point, similarity,** and **irony** (factual); **suggestion, question, explanation,** and **proposal** (linguistic); **concern, idea, rationale, objective, motivation, consideration,** and **resolution** (mental); and **condition, possibility, task** and **step** (modal); and in the Submissions: **lessons, alternative, matter, result, fact,** and **modification** (factual); **motto** (linguistic); **priority, aim, strategy,** and **concept** (mental); **action** (eventive); and **function** (circumstantial). It is to be noted that both sub-corpora have a high proportion of factual and mental shell nouns in this lexicogrammatical pattern. (See Appendices C.17 and C.18 for the distribution across texts.) For Schmid, this pattern represents a focussing function; the shell noun characterises the

information and enables the lexicalisation to be presented as New. Halliday and Matthiessen (2004: 474) see the shell noun as ‘interpreting’ the embedded clause construing Token, which suggests Schmid’s semantic function. Because Value is conflated with Subject, the validity of the statement rests upon the shell noun (Halliday & Matthiessen, 2004), and as Theme, it represents the point of departure of the message and is unlikely to be disputed. The lexicalisation, composing N-Rheme, becomes the focus of the message.

In addition to intensive identifying relational Process clauses, shell nouns are found to construe Value in possessive identifying relational Process clauses, in which the lexicalisation (underlined) construes Token (5.41). Possessive identifying relational Process clauses with a shell noun as Value, however, are not common.

5.41 [Value] **Limitations** of the Townsend index, spelled out in Section 1, *include* [Token] (a) that each census variable contributes with equal weight to the summary index; (b) census tracts are considered independent, despite similarities between neighboring tracts, possibly resulting in less efficient use of information; and (c) measures of uncertainty, particularly due to varying census tract population , are not part of the index. (P_Pov)

The second pattern identified by Schmid (2000), *Th be N*, is found in both sub-corpora, but much less frequently than that which was previously described (5.42, 5.43). While mainly factual shell nouns – particularly **case** – appear in this use, there is one instance in of an intensive identifying relational clause of the exemplifying subtype with a linguistic shell noun realising Value (5.44). In each case the lexicalisation of the shell noun is realised anaphorically, encompassing both the text reference ‘This’ and the preceding text to which the text reference refers.

5.42 This was [Value] the **case** with Schumpeter and Fisher. (S_Econ)

5.43 However, this is not necessarily [Value] the **case**. (P_Net)

5.44 This suggests a complicated **picture** of metropolitan business networks. (P_Net)

Another pattern of identifying relational Process clauses common to both sub-corpora is when Value, realised by a shell noun whose lexicalisation appears esphorically as Qualifier, conflates with N-Rheme. The relational Process is typically realised by a marked variant such as *reveal* or *reflect* (5.45) or a circumstantial verb

such as *stem from* or *result in* (5.46). This pattern is more frequent in the Submissions sub-corpus.

5.45 and *represents* [Value] the first serious **attempt** to measure them. (P_Net)

5.46 [Token] One of the difficulties that arises in attempting to apply the common game theoretical tools *stems from* [Value] the **need** to compare the probability distributions over matchings generated by a random rules when preferences are ordinal. (S_Rand)

In the Submissions sub-corpus, but not in the Published, this pattern is found in hypotactic enhancing clauses (5.47), and hypotactic elaborating clauses in which the verb is neutral, i.e. *be* rather than marked (5.48). The instances include clauses where the lexicalisation of the shell noun, which construes Qualifier, is finite (5.47) as well as non-finite (5.48). The lexicalisation is underlined in the examples.

5.47 resulting in CEO entrenchment and [Value] the **opportunity** for the CEO to demand compensation in excess of the equilibrium wage rate. (S_CEO)

5.48 which is [Value] another **way** of discussing what makes principal-agent relations work. (S_Net)

Hypertactic elaborating clauses are also found to contain anaphorically lexicalised shell nouns that construe Value (5.49).

5.49 , which is [Value] the **aim** of this paper. (S-Net)

As noted earlier, in both sub-corpora there are instances in which both Token and Value are realised by shell nouns (5.50, 5.51). In 5.50 the lexicalisation of **example** and **case** overlaps. In other words, the same lexicalisation is shelled twice: once by **example**, and once by **case**. The lexicalisation appears cataphorically in the following clause. The sentence with shell nouns serves to announce it. Similarly, in 5.51 'establishing the harmonisation and cohesion of social statistics in order to achieve comparability at European levels' is shelled as a *first advantage* and a **possibility**.

5.50 [Value] A classic **example** of how difficult it is to measure performance for managers in the non-profit sector is [Token] the **case** of a manager of a nursing home. (P_CEO)

5.51 [Value] The first **advantage** of this instrument is related [Token] to the **possibility** of establishing the harmonisation and cohesion of social statistics in order to achieve comparability at European level (S_Pov)

There are five instances in which the shell noun construes Value in an impersonal intensive identifying relational Process clause, where the Subject is

postposed. The shell nouns **case**, **purpose**, **intention**, and **focus** are found in this pattern (5.52).

5.52 It is also [Value] the **case** that although in the simple cross section nonprofits with higher levels of government grants pay their heads more, this is not true once organization fixed effects are accounted for. (P_CEO)

Finally, there are three instances in which shell nouns construing Value are found to refer cataphorically (5.53, 5.54). In this case, the shell noun always composes N-Rheme. Two of the instances select for interrogative mood, and they come from the same text.

5.53 , which consisted of [Value] the following **steps**. (P_Econ)

5.54 What are [Value] the **implications** of these results? (S_CEO)

While both the Published and Submissions sub-corpora present similar proportions of a shell noun construing Value (28% and 23% respectively), analysis of the interplay between the textual function, the interpersonal function and Value brings to light significant differences between the two sub-corpora. The Published sub-corpus is found to have a much higher proportion of instances in which Value enacts Subject and composes Theme, the lexicalisation of the shell noun [Token, Complement, N-Rheme] being realised as a fact clause (e.g. 5.34-5.41). This difference attains high statistical significance. A second significant difference is that the Submissions sub-corpus has a higher relative share of instances in which Value enacts Complement and composes N-Rheme, the lexicalisation of the shell noun being realised as a Qualifier to the shell noun (e.g. 5.45-5.49). This difference is found to be of medium significance. Finally, the Submissions sub-corpus has an instance of Value conflating with Adjunct in a circumstantial relational Process clause. The absence of this coupling in the Published sub-corpus is deemed to be of low statistical significance. (See Appendix B.13) With regard to the textual function, shell nouns construing Value are much more likely to compose Theme than N-Rheme in the Published sub-corpus, but this trend is reversed in the Submissions sub-corpus. The different trends are found to be statistically highly significant (Appendix B.14.).

There are also significant differences found in the distribution of shell noun types between the two sub-corpora (table 5.12). In the Published sub-corpus factual

shell nouns account for most uses followed by mental shell nouns, mental shell nouns accounting for only a small share. By contrast, in the Submissions sub-corpus, mental shell nouns make up most instances, followed by factual and modal shell nouns. The lower proportion of factual shell nouns and the higher proportion of modal shell nouns in the Submissions sub-corpus are found to be statistically significant.

Common to the two sub-corpora are: **reason, advantage, case, example, and fact** (factual); **conclusion** (linguistic); **aim** and **purpose** (mental); and **way** (circumstantial). In addition the Published sub-corpus includes **problem, point, irony, similarity, thing, focus, aspects, and dilemma** (factual); **explanation, suggestion and proposal** (linguistic); **intuition, analysis, picture, intention, idea, objective, motivation, rationale, consideration, concern, resolution, and interpretation** (mental); **step** and **possibility** (modal); **attempt, choice and task** (eventive); and **condition, approach, constraint, limitation, and criterion** (circumstantial). The Submissions sub-corpus includes **difficulty, importance, factor, matter, and implication** (factual); **motto** (linguistic); **assumption, lesson, issue, strategy, function, inference, hypothesis, and concept** (mental); **probability, condition, requirement, inability, potential, opportunity, ability, and impossibility** (modal); and **alternative, action, difficulty, improvement, modification, and priority** (eventive).

	Value in Published		Value in Submissions					
	%	N	%	N	T Stat	Signif.	ChiSqu	Signif.
SHELL-NOUN-TYPE	N=65		N=66					
factual	41.5%	27	24.2%	16	2.13	++	4.44	++
linguistic	7.7%	5	9.1%	6	0.29		0.08	
mental	29.2%	19	28.8%	19	0.06		0.00	
modal	4.6%	3	24.2%	16	3.30	+++	10.17	+++
eventive	6.2%	4	10.6%	7	0.91		0.84	
circumstantial-shell-noun	10.8%	7	3.0%	2	1.76	+	3.07	+

Table 5.12 Distribution of shell noun types construing Value in Published and Submissions

As well as in identifying relational Process clauses, shell nouns are found to construe Carrier and Attribute in attributive relational Process clauses in both sub-corpora. The Carrier typically composes Theme, and while most instances are in intensive attributive clauses (5.55, 5.58), there are also instances of circumstantial (5.57) and possessive attributive clauses (5.56). The lexicalisation of the shell noun

construing Carrier is found to be realised anaphorically (5.56), esphorically as Qualifier (5.55) and cataphorically (5.57, 5.58). In the examples the transitivity function precedes the element, the shell noun is in bold, and its lexicalisation – not the object of the transitivity analysis here – is underlined.

- 5.55 [Carrier] The **fact** that both the advocates and the detractors of statistical/probalilistic work could invoke the mantle of Science [rel. Pro.: intensive] is indicative of the conflicting and confused images of science in the period. (P_Econ)
- 5.56 [Carrier] the first **alternatives** mentioned above (included in category A) [rel.Pro: possessive] had drawbacks. (S_Net)
- 5.57 [Carrier] The following existence **theorem** [rel.Pro: circumstantial] is due to Klaus and Klijn. Theorem 3.2. Let (P H, PC) be a weakly responsive couple's market ... (P_Rand)
- 5.58 Financing was even harder to attain given the general ignorance of econometrics and the widespread dismissal of its potentiality: even later, when the society was beginning to impose itself, the treasurer, Roos, noted the difficulties in obtaining funding from public and other sources, since the referees of projects were very sceptical – mathematicians were rather critical and, if asked, Jacob Viner and Carl Snyder could be 'quite unfavourable', whereas Mitchell and Taussig's attitude was supposed unpredictable. [Carrier] The **paradox** of the **situation** [rel. Pro.: intensive] was obvious. Snyder, Mitchell and Taussig were members of the society and yet were suspected of not favouring the financing of its projects and activities. (S_Econ)

In example 5.58, Carrier, which composes Theme and enacts Subject, is realised with two shell nouns: **paradox** as Head and **situation** in the Qualifier. Although the lexicalisation of **situation** is realised anaphorically and **paradox** cataphorically, both are presented as Given and take specific deixis. The effect is to synthesise the writer's claim and highlight it as unlikely to be disputed. Such evaluative uses will be further discussed in the section on discourse semantics.

There are also a few instances in which the attributive relational clause in which the shell noun construes Carrier has an Attributor. In some cases the Attribute is explicitly realised (5.59) and in others it is fused with the verb, for example, reinforce = make stronger (5.60); highlight = make clear (5.61). In 5.61, Attributor is realised by an anaphorically referring shell noun. Attributive clauses with double agency occur only in the Submissions sub-corpus.

- 5.59 and [Attributor] he considered [Carrier] the **criteria** for selection [Attribute] a dangerous **limitation** for econometrics (S_Econ)
- 5.60 [Attributor] This just *reinforces* [Carrier] the **argument** that there were high probabilities of agents residing in Brazil, freed from scouting devices, incurred often in moral hazardous behaviour. (S_Net)
- 5.61 [Attributor] The present **result** *highlights* [Carrier] the urgent **need** for a code of governance practice to be introduced in this sector. (S_CEO)

Analysis of the interplay between the interpersonal function of shell nouns construing Carrier finds that in the Published sub-corpus Carrier always enacts Subject while in the Submissions sub-corpus it also enacts Complement. The latter instances correspond to the clauses that are realised with an Attributor and the difference between the two sub-corpora attains medium statistical significance. (Appendix B.15). In the textual metafunction, Carrier realised with a shell noun typically composes Theme in both sub-corpora; however, the Submissions sub-corpus shows a higher relative share of Carrier in N-Rheme (Appendix B.16) This difference attains medium significance and is, again, due to the attributive clauses with double agency. Carrier is found to compose Other, i.e. neither Theme nor N-Rheme, when the topical Theme is realised by a Circumstance (e.g. 5.62).

- 5.62 Finally, once computed, this index lacks a measure of uncertainty. For policy making decisions, [Carrier] this last **feature** may be problematic (P_Pov)

	Carrier in Published		Carrier in Submissions					
	%	N	%	N	T Stat	Signif.	ChiSqu	Signif.
SHELL-NOUN-TYPE	N=29		N=39					
factual	51.7%	15	28.2%	11	2.00	++	3.90	++
linguistic	17.2%	5	20.5%	8	0.33		0.12	
mental	13.8%	4	15.4%	6	0.18		0.03	
modal	6.9%	2	23.1%	9	1.81	+	3.21	+
eventive	0.0%	0	0.0%	0	0.00		0.00	
circumstantial-shell-noun	10.3%	3	12.8%	5	0.31		0.10	

Table 5.13 Distribution of shell noun types construing Carrier in Published and Submissions

Table 5.13 shows the distribution of shell noun types that construe Head in Carrier across the two sub-corpora. In the Submissions sub-corpus the counts are slightly higher than those in the Appendices because shell nouns appearing in Qualifiers have been counted as well as those construing Head in the nominal group that construes Carrier. The Submission sub-corpus shows a statistically significant

lower proportion of factual shell nouns and a higher proportion of modal shell nouns, and there is little overlap in those used in each sub-corpus. Common in both are: **fact, feature, result, approach**, and **criterion/criteria**. The shell nouns found only in Published are: **proof, pattern** (factual); **claim, appeal, explanation, metaphor, point** (linguistic); **assumption, theorem, idea** (mental); **trend** (modal); **position**, and **occurrence** (circumstantial). In Submissions they are: **problem, alternative, case, paradox** (factual); **remark, observation, question, answer, suggestion, argument** (linguistic); **lemma, consensus, concern, hypothesis, idea** (mental); **probability, chance, power, need, task, agenda** (modal); **procedure, situation**, and **strategy** (circumstantial) (Appendices C.15; C.16).

Although not strictly a Participant, Attribute is realised with a shell noun with a frequency similar to that of Token in the Published sub-corpus and Medium in the Submissions. Unlike attributive clauses in which the shell noun construes Carrier, the relative share of intensive, possessive and circumstantial relational attributive Processes is quite evenly distributed among the three modes. The intensive mode gives rise to one of the lexicogrammatical patterns identified by Schmid (2000): *This be N*, where the lexicalisation of the shell noun is realised anaphorically. Instances of this pattern are found in both sub-corpora (5.63). By contrast, impersonal projections in which the Subject is post-posed appear only in the Published sub-corpus (5.64).

5.63 The next list of Fellows was only established four years later, in 1937: Cowles, Hicks, Montara, René Roy and H. Staehle were all elected. In 1938 it was the turn of Lange, Lontief, J.C. Stamp and T.O. Yntema. By the end of the first decade of the society's existence, forty-two Fellows represented the Olympus of the 'econometric people'. This was, in any case, [Attribute] an immense success. (S_Econ)

5.64 but it is [Attribute] a **fact** that the entire discourse was conducted by continuous reference to images of "science". (P_Econ)

As previously mentioned, in addition to intensive relational clauses (5.63, 5.64), shell nouns as Attributes appear in possessive (5.65) and circumstantial relational Process clauses (5.66, 5.67). In the latter case, the circumstantial element is part of the Attribute.

5.65 501(3) nonprofits have [Attribute] the added **benefit** that contributions to the organization are deductible to the contributor. (P_CEO)

5.66 This may be [Attribute] due to the **fact** that membership of the drinking club was no longer generally acceptable, or indeed, necessary, once actors had moved up the socio-economic ladder. (P_Net)

5.67 [Attribute] Among the main empirical **findings** are that pay levels of top officers may trade lower pay for higher social benefits. (P_CEO)

The lexicalisation of a shell noun construing Attribute may be realised anaphorically (5.63 above), esphorically as Qualifier (5.65, 5.66, 5.68) or fact clause as Token (5.67), or cataphorically (5.69).

5.68 This would have translated into [Attribute] a **belief** that one could ultimately reconcile the deterministic neoclassical model with stochastic considerations in the economic context as well (P_Econ).

5.69 In such an altered environment, neo-classical economists had [Attribute] two **choices**. Either they could withdraw within the deep obscurity of the Walresian ramparts, or else they could try and strike some sort of accommodation with the new stochastic world view. (P_Econ)

The results for the interplay between Attribute and the interpersonal function show that the distribution of the interpersonal functions of Attributes differs between the two sub-corpora. The Published sub-corpus shows a relatively higher proportion of Attributes that enact Adjunct and a lower proportion of Attributes that enact Complement than the Submissions sub-corpus (Appendix B.17). This difference attains medium statistical significance and is due to the greater relative share of circumstantial relational Process clauses in the Published sub-corpus in which the circumstantial element is realised in the Attribute, as in 5.66 (e.g. due to the **fact** that), rather than fused with the verb. With regard to the textual function of Attributes, no significant differences are found between the two sub-corpora. Attributes typically compose N-Rheme (Appendix B.18).

	Attribute in Published		Attribute in Submissions					
	%	N	%	N	T Stat	Signif.	ChiSqu	Signif.
SHELL-NOUN-TYPE	N=22		N=25					
factual	50.0%	11	24.0%	6	1.88	+	3.43	+
linguistic	18.2%	4	24.0%	6	0.48		0.24	
mental	18.2%	4	16.0%	4	0.19		0.04	
modal	4.5%	1	16.0%	4	1.27		1.62	
eventive	4.5%	1	4.0%	1	0.09		0.01	
circumstantial-shell-noun	4.5%	1	16.0%	4	1.27		1.62	

Table5.14 Distribution of shell noun types construing Attribute in Published and Submissions

There are minor differences between the two sub-corpora in the distribution of shell noun types found to construe Attribute (table.5.14). The Published sub-corpus reveals a higher proportion of factual shell nouns while the Submissions sub-corpus has a higher proportion of modal shell nouns. These differences attain low statistical significance. Furthermore, there is practically no overlap in the shell nouns used in the two sub-corpora; only **problem** and **benefit** – both factual shell nouns – are common to both. By far the most common shell noun in the Published sub-corpus is **fact**, which surprisingly is absent from the Submissions. All the other shell nouns represent single instances. The shell nouns found to construe Head in Attribute in the Published sub-corpus are: **fact**, **findings**, **evidence** (factual); **recommendation**, **tidings**, **dictum**, **metaphor** (linguistic); **view**, **theory**, **belief** (mental); and **choices** and **method** (circumstantial). In the Submissions sub-corpus they are: **result**, **property**, **advantage**, **success**, **limitation** (factual); **proposition**, **remark**, **explanation**, **encouragement**, **commitment**, **insinuation** (linguistic); **agreement**, **implication**, **solution**, **subject**, **concern** (mental); **condition**, **opportunity**, **potential**, **certainty** (modal); and **way**, **means** and **approach** (circumstantial) (See Appendices C.21, C.22 for the distribution across texts). It is to be noted that while most of the instances in Published come from one text, P_Econ, in Submissions they are more evenly spread among S_Econ, S_Net and S_Rand.

5.3.6. Shell nouns as Medium

The final experiential function of shell nouns at clause rank is that of Medium. Most instances of Medium are found in passive clauses and the lexicalisation of the shell noun is realised anaphorically, as in 5.70. However, there are instances of the lexicalisation being realised esphorically, as Qualifier (5.71) and cataphorically (5.72). There are also instances of self engendering Processes (5.73). In the examples the experiential function precedes the ranking element, the shell noun is in bold, and its lexicalisation, which is not the focus of the transitivity analysis, is underlined.

5.70 According to the view of the foremost theorists, the development of the doctrines of utility and value had laid the foundations of scientific economics in exact concepts and it would be possible to erect upon the new foundation a firm structure of interrelated parts which, in definiteness and cogency, would be suggestive of the severe beauty of the mathematic-physical sciences. But [Medium] this **expectation** [Process] has not been realized. (P_Econ)

- 5.71 [Medium] the **proposal of publishing a journal** [Process] had been constantly evoked. (S_Econ)
- 5.72 [Medium] A **miracle** [Process] was needed in order to publish *Econometrica* and it came in the form of a complete **surprise**: Alfred Cowles III, the son of a millionaire, the president of an investment counselling firm, Cowles & Co., and a competent statistician interested in stock market predictions, offered to pay twelve thousand dollars a year for the journal. (S_Econ)
- 5.73 Informal institutional networks were still important for co-investment opportunities during the 1770s, but [Medium] this **trend** [Process] declined in the 1780s. (P_Net)
- 5.74 In spite of these early efforts, [Medium] the decisive **steps** in the econometric movement [Process] were not taken until it became an American enterprise.

Medium is always found to enact Subject, and with one exception, it composes Theme (Appendices B.19; B.20). The exception is in the Submission sub-corpus, where there is a Circumstance as marked Theme (5.74). The lexicalisation of the shell noun in this instance is realised cataphorically. The distribution of shell nouns in Medium across the Published and Submissions is shown in table 5.15.

	Medium in Published		Medium in Submissions					
	%	N	%	N	T Stat	Signif.	ChiSqu	Signif.
SHELL-NOUN-TYPE	N=14		N=22					
factual	42.9%	6	31.8%	7	0.66		0.45	
linguistic	7.1%	1	22.7%	5	1.21		1.50	
mental	7.1%	1	18.2%	4	0.92		0.87	
modal	21.4%	3	13.6%	3	0.60		0.37	
eventive	7.1%	1	4.5%	1	0.32		0.11	
circumstantial-shell-noun	14.3%	2	9.1%	2	0.47		0.23	

Table 5.15 Distribution of shell noun types construing Medium in Published and Submissions

The results show that none of the differences between the two sub-corpora in the shell nouns that construe Medium are statistically significant. Only one shell noun is common to both sub-corpora: **result** (factual). The Published sub-corpus includes the shell nouns **proof**, **reason**, **feature**, and **difficulty** (factual); **question** (linguistic); **expectation** (mental); **trend** (modal); and **position** and **situation** (circumstantial). The Submissions sub-corpus includes: **corroboration**, **difficulty**, and **miracle** (factual); **proposition**, **proposal**, **remarks**, **comment**, and **argument** (linguistic); **hypothesis**, **strategy** and **policy** (mental); **steps** and **need** (modal); **initiative** (eventive); and **approach** and **criterion** (circumstantial) (Appendices C.23, C.24 show their distribution across texts).

The preceding section has described the Participant functions that shell nouns construe in the two sub-corpora, but with the exception of exemplifying Qualifiers that realise the lexicalisations of the shell nouns, it has not explored the ways in which shell nouns are modified by means of elements preceding the Head or Qualifiers realised by embedded elaborating clauses or prepositional phrases. A description of these modifiers to shell nouns as Participants follows. Modifiers to shell nouns construing Head in a prepositional phrase that construes a Circumstance will follow the lexicogrammatical analysis of shell nouns in Circumstances (section 5.5).

5.4. Modification to Participants realised with shell nouns

Both sub-corpora present instances in which the shell noun construing Head of Participant or Attribute is modified by elements preceding it and/or elements following it. The examples include modification by means of Post-Deictic (5.75); Numerative (5.76; 5.77); Epithet (5.78); and/or Classifier (5.79; 5.80), and they appear across all classes of shell noun.

- 5.75 the [Post-Deictic] following [Classifier] existence [Mental shell noun] **theorem** (P_Rand)
- 5.76 [Numerative] One of the [Epithet] clearest [Mental shell noun] **lessons** (S_Rand)
- 5.77 its [Numerative] first [Modal shell noun] **task** (S_Econ)
- 5.78 an [Epithet] important [Linguistic shell noun] **question** (P_Pov)
- 5.79 this [Classifier] endogeneity [Factual shell noun] **problem** (S_CEO)
- 5.80 the [Classifier] college admission [Factual shell noun] **problem** (S_Rand)

The following sections will explore these instances of modification. A general comparison of the elements tagged at group level is given in Appendix D.1, and tables summarising the instances of modification are available in Appendices D.2 – D.5.

5.4.1. Modification realised by Post-Deictic

Post-Deictics appear most frequently with factual and mental shell nouns in both sub-corpora, and despite the fact that there are no statistically significant differences between the two sub-corpora (table 5.16), a qualitative analysis shows that in the Submissions sub-corpus, Post-Deictic typically serves to locate the shell noun within the text (5.81; 5.82), while in the Published sub-corpus, it more frequently

assesses the familiarity of the shell noun (5.83) or similarity to a defined group. In the examples the experiential function precedes the element, the shell noun is in bold, its lexicalisation, which is not the focus of the analysis here, is underlined, and the post-Deictic is in italics.

- 5.81 Yet the [Post-Deictic] *next* **initiative** would not fail. In the autumn of 1927 Frisch met Schumpeter for the first time, at Harvard. ... (S_Econ)
- 5.82 However this [Post-Deictic] *additional* **advantage** of counting on people in the colony represents the typical situation that raises agency problems derived from asymmetric information (S_Net)
- 5.83 At level II, which in our case characterizes between area variation, we make the [Post-Deictic] *customary* **assumption** that $\pm i \gg N.0; 1=mi/$. (P_Pov)

The five most frequent Post-Deictics in the Submissions sub-corpus are alternative (5), following (4), above (4), initial (2), and next (2). In the Published sub-corpus the five most frequent Post-Deictics are following (2), next (2) potential (2), same (2), and other (2). All other instances are single occurrences. The Post-Deictics found in nominal groups in which a shell noun construes a Participant are listed in Appendix D.2.

	Participant containing Post-Deictic in Published		Participant containing Post-Deictic in Submissions					
	%	N	%	N	T Stat	Signif.	ChiSqu	Signif.
SHELL-NOUN-TYPE	N=24		N=33					
factual	50.0%	12	39.4%	13	0.79		0.63	
linguistic	8.3%	2	3.0%	1	0.88		0.78	
mental	20.8%	5	33.3%	11	1.03		1.08	
modal	12.5%	3	6.1%	2	0.84		0.72	
eventive	0.0%	0	12.1%	4	0.00		3.13	+
circumstantial-shell-noun	8.3%	2	6.1%	2	0.33		0.11	

Table 5.16 Distribution of shell noun types among Participants realised with Post-Deictic in Published and Submissions

5.4.2. Modification by Numerative

Numeratives occur most frequently with cataphorically lexicalised shell nouns, (5.84), but they are also found with esphorically and anaphorically referring shell nouns. The lexicalisation of an esphorically referring shell noun may be realised in the Qualifier (5.85) or as a fact clause construing Token (5.86). When the shell noun refers

anaphorically, there is typically a Qualifier or Classifier in addition to the Numerative (5.87), which further identifies the lexicalisation of the shell noun. As in previous examples, the experiential function precedes the element, the shell noun is in bold, the lexicalisation is underlined, and Numerative is in italics.

- 5.84 [Medium] [Numerative] *Several* **steps** were taken to ensure validity and reliability of the data. First, the point of departure was a questionnaire already tested in other economic fields ... (S_CEO)
- 5.85 There have been [Existent] [Numerative] *many* **attempts** to go beyond the bounds of the simple consideration of income resources. (S_Pov)
- 5.86 [Value] [Deictic] The [Numerative] *third* **reason** nonprofits are an interesting context in which to study compensation is that there is a theoretical literature that deals, in part, with how managers in nonprofits may be paid, but there is almost no empirical work in the area. (P_CEO)
- 5.87 Each of the 115 principals used [Goal] [Deictic] the [Numerative] *four* **solutions** [Qualifier] previously pointed out in section 2. (S_Net)

In cataphorically lexicalised shell nouns, the Numerative helps to create the expectation that the lexicalisation of the shell noun will be enumerated either by means of conjunction or an ordinal Numerative.

SHELL-NOUN-TYPE	Participant with Numerative in Published		Participant with Numerative in Submissions		T Stat	Signif.	ChiSqu	Signif.
	%	N	%	N				
	N=19		N=22					
factual	42.1%	8	36.4%	8	0.37		0.14	
linguistic	5.3%	1	13.6%	3	0.89		0.81	
mental	15.8%	3	27.3%	6	0.87		0.78	
modal	0.0%	0	9.1%	2	0.00		1.82	
eventive	15.8%	3	9.1%	2	0.64		0.43	
circumstantial-shell-noun	21.1%	4	4.5%	1	1.62		2.59	

Table 5.17 Distribution of shell noun types among Participants realised with Numerative in Published and Submissions

Although Numeratives appear most frequently in Participants whose Head is a factual shell noun, they are not restricted to this class (table 5.17). The most common shell noun to be modified by a Numerative is **reason**, followed by **advantage** and **ways** in the Published sub-corpus, and in the Submissions, it is **hypothesis**, **advantage** and

remarks. Most shell nouns modified by a Numerative, however, are single instances. (See Appendix D.3)

5.4.3. Modification by Epithet

Nominal groups that are realised with a non-specific Deictic, Epithet and shell noun as Head appear in the lexicogrammatical pattern identified by Schmid (2000): *Th – be – N*, where the shell noun construes Attribute (5.88, 5.89).

5.88 But Frisch understood rapidly that this was [Deictic] an [Epithet] unfair [Linguistic shell noun] **insinuation** (S_Econ)

5.89 This suggests [Deictic] a [Epithet] complicated [Mental shell noun] **picture** [Qualifier] of networks

These Epithets clearly serve an evaluative purpose and they will be picked up and discussed in detail in the analysis at the discourse semantics stratum. An exception to the combination of non-specific Deixis and Epithet is found with the Epithets *main* (5.90) or synonyms *fundamental*, *central* and *predominant*. In this case, the nominal group typically construes Value in an intensive relational Process clause in another of the patterns identified by Schmid (2000) *Th-N be cl* (5.91).

5.90 We are now ready to state [Deictic] our [Epithet] main [Factual shell noun] **result** (P_Rand)

5.91 [Deictic] The [Epithet] main [Linguistic shell noun] **conclusion** is that adoption of governance principles related to board composition is urgently needed in the Portuguese NPO sector. (S_CEO)

SHELL-NOUN-TYPE	Participant containing Epithet in Published		Participant containing Epithet in Submissions		T Stat	Signif.	ChiSqu	Signif.
	%	N	%	N				
	N=32		N=38					
factual	28.1%	9	13.2%	5	1.56		2.43	
linguistic	15.6%	5	18.4%	7	0.30		0.10	
mental	34.4%	11	21.1%	8	1.24		1.56	
modal	0.0%	0	31.6%	12	0.00		12.20	+++
eventive	9.4%	3	10.5%	4	0.16		0.03	
circumstantial-shell-noun	12.5%	4	5.3%	2	1.07		1.16	

Table 5.18 Distribution of shell noun types among Participants realised with Epithet in Published and Submissions

A comparison of the frequency of Epithets with shell noun types in the two sub-corpora shows that the Submissions sub-corpus has a statistically significant higher proportion of modal shell nouns with Epithets and the Published sub-corpus a higher proportion of circumstantial shell nouns. The prevalence of modal shell nouns with Epithets is due mainly to repeated uses of **probability** in S_Net, where the writer was assessing the average probability of investors having a common acquaintance. Other than this difference, Epithets occur most frequently with mental and factual shell nouns. The Epithets found in nominal groups in which a shell noun construes a Participant are listed in Appendix D.4.

5.4.4. Modification by Classifier

Classifiers such as *existence*, *endogeneity*, or *college admission* add ideational meaning to the shell noun, but they do not realise its lexicalisation. This feature is exemplified in 5.92.

5.92 Traditionally eq. (1) was estimated by OLS, but Core et al. (1999) and Gosh and Sirmans (2005) found that in the pay-performance model, the board composition is endogenous to the system and as such, an OLS model is subject to mis-specification. Therefore, estimated OLS parameters obtained by the traditional approach are not consistent, either because of measurement errors in the Board variables or because the explanatory variables, such as board composition or individual education, are correlated with the unobserved disturbances in the equation. Alternative approaches to deal with [Deictic] this [Classifier] endogeneity [Thing: factual shell noun] **problem** fall into three categories: (S_CEO)

The lexicalisation of the shell noun **problem** is realised anaphorically in the preceding clauses, which explain the problem. The Classifier *endogeneity*, itself a nominalisation of *endogenous*, does not compact the entire lexicalisation, but it does suggest the origin of the problem. Such Classifiers may contribute to establishing the field of discourse. For example S_Rand explores Nash equilibrium game theory in the context of ‘the college admission problem’, so the Classifier helps to construe the field of discourse as the object of research of the writer. The relation of Classifier to shell nouns and field will be developed in chapter 6 in the analysis of shell nouns at the stratum of discourse semantics. The Classifiers found to modify Participants are listed in Appendix D.5.

	Participant with Classifier in Published		Participant with Classifier in Submissions					
	%	N	%	N	T Stat	Signif.	ChiSqu	Signif.
SHELL-NOUN-TYPE	N=19		N=17					
factual	21.1%	4	23.5%	4	0.17		0.03	
linguistic	0.0%	0	11.8%	2	0.00		2.37	
mental	26.3%	5	17.6%	3	0.61		0.39	
modal	5.3%	1	35.3%	6	2.39	++	5.17	++
eventive	0.0%	0	0.0%	0	0.00		0.00	
circumstantial-shell-noun	47.4%	9	11.8%	2	2.44	+++	5.36	++

Table 5.19 Distribution of shell noun types among Participants realised with Classifier in Published and Submissions

Although the results show statistically significant differences for the use of Classifiers with shell nouns construing Participants, the number of instances is low, so not too much should be read into them.

5.4.5. Modification by Qualifier

The function of adding ideational meaning can be realised by means of Qualifiers, which may construe the logico-semantic function of expansion or projection. In the latter case the Qualifier constitutes the lexicalisation of the shell noun and it exemplifies Schmid's (2000) pattern *N-cl*. Shell nouns whose lexicalisation is projected as Qualifier are found to construe a range of Participant functions, there being few differences between the Published and Submissions sub-corpora (Appendix D.6). In descending order of frequency of their four main functions, which account for more than half the instances, they construe Value, Goal, Carrier, and Token in the Submissions sub-corpus, and Goal, Value, Carrier, and Token in Published.

The projected Qualifiers in the corpus are found to be finite and non-finite. 5.93 illustrates a finite clause projected as Qualifier. In the following examples the shell noun is in bold and the lexicalisation, which also construes Qualifier, is underlined.

5.93 The [Thing: factual shell noun] **pattern** [Qualifier] that the top officer or director earns more than the top nonofficer, nondirector, or nontrustee as reported in Table 2 is still evident in Table 3. (P_CEO)

The analysis undertaken of projected Qualifiers in is extended to include instances in which the embedded clause is non-finite and imperfective, introduced by means of a preposition. This pattern is exemplified in 5.94.

5.94 However, he had the [Thing: factual shell noun] **advantage** [Qualifier] of collecting information on the spot. (S_Net)

In addition, for the purposes of the analysis, embedded clauses construing Qualifier to nouns of expansion such as **way**, which are deemed to be functioning as shell nouns, are counted as projected rather than expanding. 5.95 shows a non-finite, perfective clause projected as Qualifier to the shell noun **way**.

5.95 There is a host of other potential [Thing: circumstantial shell noun] **ways** [Qualifier] to measure performance in nonprofits. (P_CEO)

The results for the realisation of projected Qualifiers in Participants or Attribute reveal significant differences between the two sub-corpora (table 5.20). In the Published, the proportion of finite clauses is higher than non-finite clauses while in the Submissions sub-corpus this trend is reversed. The difference is found to be highly statistically significant.

Feature	Participant with Qualifier in Published		Participant with Qualifier in Submissions					
	Percent	N	Percent	N	T Stat	Signif.	ChiSqu	Signif.
PROJECTED-TYPE	N=59		N=77					
finite	69.5%	41	27.3%	21	5.36	+++	24.00	+++
non-finite	30.5%	18	72.7%	56	5.36	+++	24.00	+++
NON-FINITE-TYPE	N=18		N=56					
perfective	88.9%	16	50.0%	28	3.07	+++	8.55	+++
of-+-imperfective	11.1%	2	50.0%	28	3.07	+++	8.55	+++

Table 5.20 Distribution of realisation types of projected Qualifiers for shell nouns construing Participants

While these results might in part be attributed to the fact that compared to the Submissions sub-corpus, the Published sub-corpus contains a higher proportion of factual shell nouns, which typically take finite clauses as Qualifier, and a correspondingly lower proportion of modal shell nouns, whose Qualifier is typically realised by a non-finite clause, a further difference is found in the realisations of non-finite projected Qualifiers. The Submissions sub-corpus shows a lower proportion of

perfective clauses and a higher proportion of imperfective clauses introduced with the preposition *of* than the Published sub-corpus, a difference that again attains high statistical significance. The difference does seem to reflect a systemic choice by the writer, given that all of the writers are clearly aware of the resources of the lexicogrammatical system involved. By way of illustration, examples 5.96 and 5.97, which are non-finite and imperfective, could be realised respectively as perfective, as in 5.96', and finite, as in 5.97'.

5.96 the higher the discretionary **power** of the CEO in establishing the Board.
(S_CEO)

5.96' the higher the discretionary **power** of the CEO to establish the Board.

5.97 In this approach, it is also possible [[to estimate the **probabilities** of a household (with a specific well-being profile) belonging to different deprivations groups (latent classes)]]. (S_Pov)

5.97' In this approach, it is also possible [[to estimate the **probability** that a household (with a specific well-being profile) belongs to different deprivation groups (latent classes)]].

Projecting a clause as Qualifier to a shell noun enables the writer to construe complex ideas as a single clause; the projected clause being rank-shifted to function within a group rather than as a clause in its own right. The Qualifier specifies and lexicalises the shell noun, and its contribution to the discourse is ideational. This suggests that analysis of these esphoric lexicalisations may reflect shifts in field. I have not attempted a characterisation of the field construed in these Qualifiers here. The relation between a lexicalisation construed as Qualifier, the shell noun and field will be addressed in Chapter 6.

The systemic resource of rankshifting a clause to function as Qualifier to a shell noun also makes it possible to draw on the resources of the textual system and information status. The results show that in both sub-corpora shell nouns realised with a projected Qualifier – i.e. shell nouns whose lexicalisation construes Qualifier – compose N-Rheme more frequently than Theme (Appendix D.7). Almost two thirds of the instances in the Published sub-corpus and a slightly higher proportion in the Submissions sub-corpus are found to compose N-Rheme. This makes it possible to present both the lexicalisation and the shell noun that projects it as New. This feature

is illustrated in 5.99, where the shell noun is in bold and the lexicalisation is underlined.

5.99 This supports the [Thing: linguistic shell noun] **assertion** [Qualifier] that the use of drinking clubs for trade-related networking declined, but that the Lyceum was important culturally. (P_Net)

The tendency for lexicalisations construed as Qualifiers to modify Participants who compose N-Rheme is echoed in the interpersonal analysis, where these Participants tend to enact Complements rather than Subject (Appendix D.8).

Analysis of the corpus reveals that Qualifiers that expand are realised mostly by prepositional phrases, but also by elaborating embedded clauses. These expanding Qualifiers do not constitute the lexicalisation of the shell noun. Like elements that precede the Head, they are found to appear across all shell noun classes, but not with equal frequency. In descending order of number of instances, expanding Qualifiers are found with factual, mental, linguistic, circumstantial, modal, and eventive shell nouns. The order is the same for both sub-corpora. In the main these Qualifiers serve to help specify, but not lexicalise, the shell noun by specifying the object of study, thus they can contribute to establishing the field of the discourse. In example 5.100 the shell noun is lexicalised anaphorically, in 5.101 cataphorically, and in 5.102 esphorically as a fact clause that construes Token.

5.100 This suggests a complicated **picture** [Qualifier] of networks. (P_Net)

5.101 We have to use a **refinement** [Qualifier] of the Nash equilibrium concept. (S_Rand)

5.102 One of the most curious **aspects** [Qualifier] of the rise of neoclassical theory is that many of the leading lights of marginalism were also instrumental in the development of probability theory and statistics. (P_Econ)

In examples 5.100 – 5.102 the Qualifier helps to construe the field of the text as the object of study; however, not all instances of Qualifiers realised by prepositional phrases have this function. Some prepositional phrases realising Qualifiers have a semiotic object as Head, as shown in 5.103. The most frequent semiotic object to construe Head in a Qualifier to a shell noun is *paper*, and this is the same in both sub-corpora. In the Submissions sub-corpus, *section* is used in the same way. In all instances of the corpus the shell noun is mental and deixis is specific. The function of the Qualifier in these instances seems rather to locate the study within the field of

research or an aspect of the study within the paper (5.104). Note that in 5.103 the Subject is split, and its second element – ‘to present a detailed overview of the field of graph theory’ – constitutes the lexicalisation, underlined.

5.103 It is not the **intention** [Qualifier] of this paper to present a detailed overview of the field of graph theory. (P_Net)

5.104 Second, the **aim** [Qualifier] of the last section is to shed some light on what happens once we move towards allowing for history-dependent strategies. (S_Rand)

There are also instances in both the Published and Submissions sub-corpora in which a Qualifier to a shell noun contains a shell noun itself. This dual use of shell nouns can create cohesive ties in two directions. While the shell noun in the Qualifier typically refers anaphorically, the shell noun as Head can and often does refer esphorically or cataphorically as in 5.105 and 5.106 respectively.

5.105 The[Thing: factual shell noun] **problem** [Qualifier] with this [Thing: mental shell noun] **interpretation** is that it wilfully disregards the fact that the main point of contention between the disputants was the validity of neoclassical economic theory, and not all theory tout court. (P_Econ)

5.106 What are the [Thing: factual shell noun] **implications** [Qualifier] of these [Thing: factual shell noun] **results**? (S_CEO)

These examples illustrate how, like an accordion, both the prior and following text is condensed in the shell nouns, and the text moves on with the esphoric or cataphoric reference element grounded or anchored by the anaphorically referring shell noun in the Qualifier. Shell nouns also appear as Qualifiers to nominal groups with other nouns as Head, as in 5.107 and 5.108.

5.107 The struggle to establish econometrics has frequently been misrepresented as a battle between the misguided partisans of “measurement without theory” versus the level-headed exponents of a judicious and balanced empiricism. The canonical text [Qualifier] in this **interpretation** was the famous debate in the 1940s between Tjalling Koopmans and Rutledge Vining. (P_Econ)

5.108 The agents [Qualifier] in the college admissions **problem** are two finite and disjoint sets. (S_Rand)

Once again the Head of the Qualifier refers anaphorically, and the shell noun grounds the information presented as Theme and Subject. In these two examples the Qualifier could have been realised as a Circumstance, composing a marked Theme, but the writer has chosen to construe it at group rank rather than clause rank.

Similar to Qualifiers realised by prepositional phrases, Qualifiers realised by an embedded elaborating clause may also contribute to field. In 5.109 the embedded clause elaborates on the object of study: CEO pay in non-profit organisations. Schmid (2000:102-104) considers the embedded clause to be a lexicalisation of the shell noun **reason**. He argues that **reason** has two gaps to be filled: one which lexicalises the *effect* and one which lexicalises the *cause*. In this case, the embedded clause lexicalises the effect, and the gap for cause is lexicalised cataphorically.

5.109 Another **reason** [Qualifier][[that an organization may form as a non-profit]] comes out of the idea of public goods. (P_CEO)

5.110 The **task** [Qualifier] [[this small group was setting itself]] was immense in three different fields. Firstly they endeavoured to create a new discipline inside economics: quite originally the Society was created precisely to define its own subject. ... (S_Econ)

By contrast, in 5.11, the Qualifier helps to identify the shell noun and enables it to be realised as Given information. A full list of expanding Qualifiers found in the two sub-corpora is included in Appendices D.9 – D.12.

5.5. Circumstances realised with shell nouns

In addition to construing Head in a Participant, shell nouns are found to construe Head in a nominal group in a prepositional phrase construing Circumstance. The results show that shell nouns are present in enhancing, extending and projection Circumstances in both sub-corpora, but elaborating Circumstances are only found in the Submissions sub-corpus.

Examples 5.111 – 5.112 illustrate how a shell noun is used in a Circumstance composing marked Theme. In line with the other analyses, the functional label precedes the element under analysis, the shell noun is in bold and the content it encapsulates – its lexicalisation, which is not the focus of analysis here – is underlined.

5.111 Evaluation of PBF was based on the logarithm of the conditional predictive ordinates (CPO) computed from [equation 9], where $zip, p=1, \dots, I$ denote the observed vectors for unit $i=1, \dots, np$ and [equation 10] represent the simulated values of $?$ from its posterior distribution, given all observations but zip .

[Circumstance] According to these **criteria**, the best models are those with the lowest values for DIC or BIC and the highest value for InCPO. (S_Pov)

5.112 on November 1, 1926 Frisch wrote to four colleagues, Ladislau von Borkiewicz, Charles Jordan, Arthur Bowley and Eugene Slutsky – no one from the US. Slutsky, whom Frisch had already met in Oslo, was the most enthusiastic about the new association (Bjerkholt, 1998:31-32), although later on he never adhered to it. The same day, Frisch informed Divisia of the initiative of this letter. [Circumstance] In spite of these early **efforts**, the decisive steps in the creation of the econometric movement were not taken until it became a European-American enterprise. (S_Econ)

Most of the examples in which a Circumstance composes a marked Theme involve anaphoric reference, i.e. the shell noun refers to information preceding the Circumstance. 5.111 is a straight-forward example in which the shell noun **criteria** refers to the preceding clause complex, and in 5.112 the shell noun **efforts** refers to a portion of text larger than a clause complex. In some cases, the lexicalisation of the shell noun appears esphorically, projected by the shell noun as Qualifier. This use is shown in example 5.113.

5.113 [Circumstance] Despite, or perhaps due to, the **fact** that it was a durable subnet, it relied on information passed by others to a greater extent than the other cross-institutional subnets. (P_Net)

The use of anaphorically referring shell nouns in a Circumstance as marked Theme makes two contributions to the development of the text. Not only does encapsulating the preceding information in a noun enable the discourse to move forward (Halliday & Matthiessen, 2004; Francis, 1994), but the choice of Circumstance in which the shell noun appears makes the reasoning explicit. Examples 5.111 and 5.112 illustrate how the text develops by means of projection and contingency respectively, made clear by the prepositions **according to** and **in spite of**. The exception to this is the shell noun **case**. When case refers anaphorically, the distinction between location and contingency disappears. This is shown in 5.114.

5.114 However, it may be that the couple together with the hospital and the old match of the best student's partner does not form a blocking condition. [Circumstance] In this **case**, we prove that the couple together with the hospital and the best student's previous match forms a blocking coalition. (P-Rand)

In 5.114, the shell noun **case** refers anaphorically to the preceding clause. Reconstruing the Circumstance as a hypotactically bound clause shows that there

are two possibilities (5.114'; 5.114''). The Circumstance with shell noun **case** neutralises this distinction.

5.114' However, **if** the couple together with the hospital and the old match of the best student's partner does not form a blocking condition, we prove that the couple together with the hospital and the best student's previous match forms a blocking coalition.

5.114'' However, **when** the couple together with the hospital and the old match of the best student's partner does not form a blocking condition, we prove that the couple together with the hospital and the best student's previous match forms a blocking coalition.

In the seven instances in which the shell noun in the Circumstance composing marked Theme refers or is lexicalised cataphorically, there is always modification in the prepositional phrase. This applies to both the Published and Submissions sub-corpora. Modification may be by means of a Numerative, a Post-Deictic indicating location or a postmodifying Numerative (Table 5.21).

Published	Submissions
Finally, in the third stage In the following example For a variety of reasons However, for at least two reasons	In Proposition 7 In Proposition 8 As a first approach

Table 5.21 Modification in cataphorically referring or lexicalised shell nouns in Circumstances composing marked Theme

While the shell noun in Circumstances composing a marked Theme is more likely to refer anaphorically, the shell noun in Circumstances composing N-Rheme is more likely to be lexicalised cataphorically or esphorically. And, similar to those lexicalised cataphorically in the marked Theme, in the N-Rheme a cataphorically lexicalised shell noun is also modified. The modifiers are more varied; in addition to Numeratives, Post-Deictics indicating location and postmodifying Numeratives, there are Epithets and expanding Qualifiers. As previously noted, while the Epithets provide an evaluative function that will be picked up at the stratum of the discourse semantics, the Qualifiers that are expanding help to specify the shell noun but do not constitute the lexicalisation. Table 5.22 shows the modifiers from the two sub-corpora.

Published			
(P_Econ)	from a number of points of	view	
(P_Net)	for many	reasons	
(P_Pov)	in two	ways	
(P_Econ)	to the	issues	discussed in this paper
(P_CEO)	out of the	idea	of public goods
(P_CEO)	from the screening	hypothesis	of Hansmann U
(P_Econ)	in his famous	manifesto	
(P_Rand)	by means of a simple	example	
(P_Net)	in the	analysis	below
Submissions			
(S_Pov)	due to several	strategies	
(S_Net)	for two	reasons	
(S_Rand)	in two	ways	
(S_CEO)	in two	stages	
(S_Econ)	by a battery of	reasons	
(S_Econ)	in three different	fields	
S_Econ	to any of these three	goals	
S_Pov	according to the following	criteria	concerning allocation of the household's profiles to latent classes
S_CEO	for the present	purpose	
S_Net	from this social	theory	
(S_Rand)	in the following natural	way	
S_Rand	in the college admissions	problem	
(S_Rand)	in	Proposition	5
(S_Econ)	in view of a	remark	made by Divisia
S_Econ	in the form of a complete	surprise	

Table 5.22 Modification to shell nouns in Circumstances composing N-Rheme

These findings are in line with those of Francis (1994) and Dueñas (2003-2004), who argue that Numeratives such as *first* or *two* and Epithets such as *different* have a textual function in that they “contribute directly to the organizational role of labels [shell nouns]: they help to order messages with respect to each other and signal the relationships between them” (Francis, 1994:98). On the other hand Epithets such as *simple*, *natural* and *famous* are clearly interpersonal. The remaining modifiers – either Qualifiers or the Classifiers *screening* and *college admissions* – are ideational. While they help to specify the shell noun, as pointed out earlier, they are not its lexicalisation.

In almost a third of the instances in which a Circumstance with shell noun composes N-Rheme (10 out of 32) in the Submissions sub-corpus, the lexicalisation of the shell noun is realised esphorically as Qualifier. In the Published sub-corpus this pattern accounts for more than half the corresponding instances (17 out of 26). The effect is to draw attention to the lexicalisation, i.e. the projection. A couple of

examples suffice to illustrate this. Again the transitivity label precedes the element under analysis, the shell noun is in bold, and its lexicalisation, which is realised by the Qualifier, is underlined.

5.115 These doubts began [Circumstance] with the **observation** that it was humanly impossible to know all of the initial conditions for any moderately complicated Hamiltonian. (S_Econ)

5.116 since the static model could not legitimately be extended through time [Circumstance] due to the perennial **complaint** that the very structure of the fundamental determinants of the economy had changed in the interim. (P_Econ)

The remaining instances of shell nouns in Circumstances compose neither Theme nor N-Rheme. Of the nine instances, four involve anaphoric lexicalisations of the shell noun, two involve esphoric lexicalisations, and the remaining four are cataphoric. Interestingly with the exception of one instance, the latter group contains interpersonal Epithets while the exception is followed by a Circumstance of location (5.117 – 5.119).

5.117 as simple **proposals** for valuation of two criteria for the choice of next Fellows

5.118 by an apparently an apparently naïve **suggestion** by Morgenstern to his Fellows,

5.119 for our **analyses** in section 4

Comparison of the types of Circumstance in the two sub-corpora shows that there is little difference between them. Most of the Circumstances – around 80% - are of the enhancing type, and a further 10% are of the projecting type. The Submissions sub-corpus is found to contain a slightly higher proportion of Circumstances of manner, significant at weak level, and the Published sub-corpus is found to contain a higher proportion of Circumstances of location. The full results for the distribution of Circumstances containing a shell noun across the two sub-corpora are given in Appendix E.1.

When the Circumstance is analysed for the kind of shell noun that appears in it, two differences between the two sub-corpora are found to be significant at a medium and low level respectively: the Submissions sub-corpus has a lower proportion of mental shell nouns in Circumstances and a higher proportion of circumstantial shell

nouns. Exploration of the instances of the latter result finds that the text *S_Rand* accounts for more than half of the occurrences, and in this text there are repeated instances of the shell noun **case**. Although the other differences do not attain statistical significance, the Published sub-corpus shows a higher proportion of all other shell nouns except linguistic.

	Circumstance in Published		Circumstance in Submissions					
	%	N	%	N	T Stat	Signif.	ChiSqu	Signif.
SHELL-NOUN-TYPE	N=38		N=67					
factual	26.3%	10	19.4%	13	0.82		0.68	
linguistic	13.2%	5	25.4%	17	1.48		2.18	
mental	36.8%	14	17.9%	12	2.19	++	4.66	++
modal	7.9%	3	7.5%	5	0.08		0.01	
eventive	2.6%	1	1.5%	1	0.41		0.17	
circumstantial-shell-noun	13.2%	5	28.4%	19	1.79	+	3.18	+

Table 5.23 Shell noun types in Circumstantial element at clause rank

In order to shed light on the use of shell nouns in Circumstantial elements, it is necessary to consider how they interact with other metafunctions. In this case, the textual status of the Circumstance – i.e. whether it is Theme or not – and its information status – i.e. whether it composes N-Rheme or not – can help to clarify the way in which the shell noun is being used. The results show that a Circumstance containing a shell noun is more likely to compose N-Rheme than marked topical Theme in both sub-corpora. In the Published sub-corpora, N-Rheme makes up over two thirds of the instances of Circumstances realised with shell nouns while in the Submissions sub-corpus N-Rheme makes up just under half. A small number of Circumstances (4 in the Published sub-corpus and 5 in the Submissions) compose neither Theme nor N-Rheme. Notwithstanding these results, compared to the Published sub-corpus, the Submissions sub-corpus shows a higher relative share of Circumstances composing marked topical Theme and a lower relative share of Circumstances composing N-Rheme. These results attain high and medium statistical significance (Appendix E.2).

5.6. Relators realised with shell nouns

The last ideational function realised by shell nouns at the stratum of the lexicogrammar is logical rather than experiential. In a few instances shell nouns appear in Relators. Five of the six instances are from the Submissions sub-corpus, and only one is from the Published. In all cases, the Relator composes textual Theme, and the lexicalisation of the shell noun, underlined, is esphoric.

- 5.120 This model was suggested by a referee [Relator] on the **grounds** that conditional spatial correlation corr2.±i ; j± jf±k : k 2 Ri; k 6D i; jg/D !2 is not dependent on neighbourhood size. (P_Pov)
- 5.121 Pay performance contracts are settled [Relator] with the **aims** to align the principal-agent relationship between management and stockholders. (S_CEO)
- 5.122 This was mailed to a random sample of Portuguese NPO administrators, stratified by organization type (Misericórdias, foundations, mutual associations, solidarity associations, other associations and cooperatives) [Relator] with the central **aim** of determining their socio-economic characteristics. (S_CEO)
- 5.123 His comment to Morgenstern's suggestion reflects his essential mathematical work, rejecting the proposal [Relator] on the **grounds** that it might narrow econometric work down to statistical work. (S_Econ)
- 5.124 Since 1994, Eurostat has conducted an annual survey in the EU member states [Relator] with the **purpose** of collecting information on the incomes of households and individuals. (S_Pov)
- 5.125 [Relator] With the **aim** of understanding the effect of the weights, we compared point estimates from weighted and non-weighted data. (S_Pov)

In five of the instances the bound clause introduced by the Relator is the second clause of the clause complex. The effect is to draw attention to the lexicalisation of the shell noun. In the other instance, the bound clause precedes the free clause in the clause complex and composes clausal Theme (5.125). Examples 5.121, 5.122, 5.124 and 5.125 could have been realised with a perfective non-finite clause while 5.120 and 5.123 could have been realised with an enhancing hypertactic marker. What the shell noun does is make the relation of the bound clause explicit by naming the lexicalisation as an **aim, purpose** or **grounds**.

5.7. Textual Themes with shell nouns

In addition to the logical Relators described above, conjunctive Adjuncts are also found to compose textual Theme. However, because they do not construe an ideational function at the rank of the clause, they have not been picked up in the

preceding analysis. The use of shell nouns in conjunctive Adjuncts that compose textual Theme is exemplified in 5.126. The lexicalisation is underlined, the textual function precedes the ranking element of the clause, and the shell noun is in bold.

5.126 The founders took opposite camps: Roos and Fisher argued for an open society, whereas Schumpeter, Frisch and Bowley preferred a closed centre of excellence with mathematically trained scientists. [textual Theme] As a consequence of his different **view** of the nature of the society, Schumpeter opposed a number of names proposed by Fisher (Bjerkholt, 1998: 38,40). (S_Econ)

The shell nouns in textual Themes realised by a conjunctive Adjunct always refer anaphorically. According to Halliday and Hasan (1976), the shell noun does not establish a cohesive tie. Its function here seems rather to add interpersonal or ideational meaning by characterising the lexicalisation. Table 5.24 lists the shell nouns found in conjunctive Adjuncts that compose textual Theme. They include the factual shell nouns **result** and **advantage**, and mental shell nouns **purpose** and **view**.

Published	Submissions
In view of this result (P_Rand)	With this purpose in mind (S_Rand) For that purpose (S_Econ) As a consequence of his different view of the nature of the society (S_Econ) Besides these advantages (S_Net) For this purpose (P_Pov)

Table 5.24 Textual Themes realised by conjunctive Adjuncts containing shell nouns

5.8. Summary and discussion

This chapter has presented the results for the analysis of shell nouns at the stratum of the lexicogrammar. It provides a global frequency analysis of shell nouns in the Published and Submissions sub-corpora and examines the experiential, interpersonal and textual functions of shell nouns at clause rank, as well as describing modification to shell nouns at group rank.

The results show that the Submissions sub-corpus has a relatively higher proportion of shell nouns than the Published sub-corpus; and there are differences in the type of shell noun found. While the relative shares of linguistic, mental, eventive and circumstantial shell nouns are similar in the two sub-corpora, the Published sub-corpus has a higher share of factual shell nouns and a lower share of modal shell nouns than the Submission sub-corpus. The latter two differences were found to be

statistically highly significant. Despite these differences, there is considerable overlap in the actual shell nouns found to be used most frequently in most types, the exceptions being linguistic and mental shell nouns. Many of these shell nouns derive from the field of research, such as **result(s)**, and they include nouns related to methodology, such as **approach** or **hypothesis**. The results for the most frequent factual shell nouns are in line with those of Aktas and Cortes (2008) and Swales (2005).

Turning to the qualitative analysis, the results show that there are few differences in the distribution of experiential functions construed by shell nouns at clause rank between the Published and Submissions sub-corpora. In the Published sub-corpus, shell nouns are found to construe Value, Carrier, Goal, Token, Attribute, Medium, Verbiage, Actor, Existent, Phenomenon, Agent, Scope, and Sayer in descending order of frequency. The Submissions sub-corpus presents a slightly different order: Value, Goal, Carrier, Token, Attribute, Medium, Phenomenon, Verbiage, Actor, Attributor, Agent, Existent, Assigner, Scope, Beneficiary, and Sayer. These findings corroborate those of Gray & Cortes (2010), who found that shell nouns typically appear in relational clauses, and, despite the higher frequency of Value, they do not conflict with Francis (1994), who argues that shell nouns typically construe Carrier. The difference is probably due to the fact that Francis only considers anaphorically referring shell nouns while this study additionally includes esphorically and cataphorically lexicalised shell nouns. Furthermore, the frequency of Value – accounting for around a quarter of the instances in each sub-corpora – confirms Halliday and Matthiessen’s (2004) remark that fact nouns typically construe Value.

In relational Process clauses, which make up over half of the clauses in which a shell noun construes Participant, shell nouns are more likely to construe Value than Token, and Carrier than Attribute. Although this trend is common to both the Published and Submissions sub-corpora, the Published sub-corpus shows a significantly higher proportion of clauses in which Value conflates with Subject and Theme, and Token, which conflates with Complement and N-Rheme, is realised as a fact clause that constitutes the shell noun’s lexicalisation. This lexicogrammatical pattern corresponds to Schmid’s (2000) *N be cl* pattern, whose function is to present the lexicalisation as the focus of the clause (Schmid, 2000). In an SFL perspective, it is the systemic choice

of composing the lexicalisation, which construes Token, as N-Rheme that confers the element with prominence.

The presence of this pattern in both sub-corpora differs from the findings of Aktas and Cortes (2008), who found no instances of it in a corpus comprising published articles from the fields of Art and Design, Biology, Computer Science, Economics, Environmental Engineering, and Physics and Astronomy and papers written in the same fields by international EAL graduate students. Furthermore, although the higher proportion of this lexicogrammatical pattern in the Published sub-corpus than the Submissions may lend weight to Willis' (1993) argument that the pattern is so frequent that it should be taught, the fact that it is also found in the Submissions sub-corpus suggests that the writers in this study are aware of how to use it, but are choosing not to do so so frequently. In the Submissions sub-corpus, shell noun as Value more often conflates with Complement and N-Rheme, in a relational Process clause whose Process is realised by a verb such as *suggest*, *imply* or *confirm*. In this case, the lexicalisation of the shell noun is typically realised as Qualifier to the shell noun. Not only does the concomitant selection of these resources focus the reader's attention on the lexicalisation projected by the shell noun, but it also contributes to the development of the writer's argument and reasoning. An anaphorically referring shell noun construing Token in a relational clause whose Process is realised by a marked variant verb such as *confirm* or *reveal* similarly contributes to the structure of the writer's argument, the ideational selections directly making this possible.

When the shell noun construes Attribute and conflates with N-Rheme in an intensive relational Process clause, and the Carrier is realised by demonstrative reference, the shell noun interprets the lexicalisation in Schmid's (2000) *Th be N* pattern. Only a few of instances of this use are found: 2 in Published and 5 in Submissions. This result contrasts with that of Aktas and Cortes (2008), who found no instances of it in their academic corpus.

With regard to the interpersonal and textual functions, shell nouns are found to enact Subject and compose unmarked topical Theme more frequently in the Published sub-corpus than in the Submissions. These results respectively attain medium and high

statistical significance. The Submissions sub-corpus, by contrast, shows a significantly higher proportion of Circumstances realised with a shell noun as marked topical Theme. Halliday and Matthiessen (2004) argue that marked Themes constitute a resource for reorienting the direction of the discourse. In this study, the shell nouns in these marked Themes typically refer anaphorically, thus creating a cohesive tie with the preceding text. Notwithstanding, the marked Theme enables the text to develop along a different path.

In both sub-corpora, Participants, Attributes and Circumstances are found to be realised with shell nouns whose lexicalisation is found in the preceding text (anaphoric), in the immediately following text of the clause (esphoric) and in a following clause (cataphoric). Similar to other studies, anaphorically referring shell nouns typically coincide with Theme and Given information (Flowerdew, 2003a; Francis, 1994; Gray & Cortes, 2010; Godsen, 2009; and Moreno, 2004), and cataphorically lexicalised shell nouns typically with N-Rheme and New information (Flowerdew, 2003a). Anaphorically referring shell nouns illustrate a textual, linking function (Schmid, 2000) and enable the discourse to proceed on the basis of shared information compacted in the shell noun and referenced through the specific Deictic (Halliday & Matthiessen, 2004; Francis, 1994). The results show that in addition to when they compose Theme, anaphorically referring shell nouns also contribute a linking function when they construe Head in a Qualifier to another noun. The reference grounds the nominal group, and, if Thematic, provides a shared point of departure for the clause. This use exemplifies what Jordan (1985) has called 'perspective reentry' and Martin (2000) 'bridging'. If Rhematic, the anaphorically referring shell noun in the Qualifier specifies the Head, establishing its relevance in the development of the argument by means of the cohesive tie. While Ravelli (2004) has noted this function of Qualifiers in hyper-Themes, this study finds that it is not restricted to hyper-Themes.

Anaphorically referring shell nouns also appear in textual Themes realised by a conjunctive Adjunct. These textual Themes both move the discourse on and make explicit the reasoning behind the developing argument. They will be dealt with in more depth in Chapter 6.

Around a quarter of the shell nouns in each sub-corpus are instances in which the lexicalisation of the shell noun is esphorically realised as Qualifier to the shell noun. This corresponds to Schmid's (2000) pattern N-cl, which typifies his cognitive function of temporary concept formation, and constitutes one of the patterns that defines whether a noun can function as a shell noun or not. In a Systemic Functional Linguistics perspective, the clause realising the lexicalisation of the shell noun has been rank-shifted to function at group rank, and it is projected by the noun, which is a semiotic abstraction. Being able to project a clause enables the shell noun and its lexicalisation to function as a Participant or indirect Participant in another clause, and this considerably expands the resources of the system to construe complex ideas in a single clause. Furthermore, as an element of the clause – Participant or Circumstance – the shell noun and its projection can draw on the resources of the textual system and information status. The results of this study show that shell nouns in this pattern enter into a variety of Processes where they construe a Participant or Circumstance, and all the writers in the study make strategic use of the resources of the textual system for them. Notwithstanding, there is a clear trend in both sub-corpora for shell nouns whose lexicalisation is realised as Qualifier to compose N-Rheme rather than Theme.

With regard to the realisation of the Qualifier, the Submissions sub-corpus has a statistically significant higher proportion of non-finite clauses. While in part this may be explained by the higher proportion of modal shell nouns, which typically take a non-finite clause as projection, it also seems to be a choice made by the Portuguese-speaking writers. There are several instances in which the shell noun could project a finite clause, but the writer has opted for a non-finite. The reason for this difference may lie in the properties of Portuguese grammar. Unlike English, Portuguese is a highly inflected language, permitting both the omission of subject, and enabling the infinitive to be marked with number and person – 'o infinitivo flexionado' (Raposo et al., 2012), the so-called 'personal infinitive'. As detailed in Chapter 3, section 3.2.3, a non-finite clause, realised with an infinitive, that functions as a complement to a noun must be introduced by a preposition, the default preposition being 'de' (of). In Portuguese, it is possible to use the personal infinitive to avoid ambiguity. Many of the shell nouns in the Submissions sub-corpus that are found to take non-finite imperfective Qualifiers –

e.g. advantage, importance, probability, chance – can and do take an infinitive complement introduced by the preposition ‘of’ in Portuguese. Although in English non-finite imperfective forms introduced by ‘of’ are possible, it is likely that the Portuguese writers are transposing the Portuguese structure into their texts.

In the corpus, cataphorically lexicalised shell nouns constitute a resource in both sub-corpora for moving the discourse forward by means of prediction rather than encapsulation. This finding is in line with findings by Tadros (1994) and Winter (1977), and helps support the argument that encapsulation and prediction are the main ways in which discourse moves forward (Álvarez-de-Mon y Rego, 2001; Sinclair, 1993). Schmid (2000) calls this use of shell nouns signposting. Similar to the results of Álvarez-de-Mon y Rego (2001), cataphorically lexicalised shell nouns are always accompanied by modification of some kind: a Post Deictic such as *following*, a Numerative such as *several* or *two*, an Epithet such as *main*, a Classifier, or a Qualifier, which may, itself, be realised with an anaphorically referring shell noun. While the fact that a shell noun is a semiotic abstraction may help create a cohesive tie to a cataphorically realised lexicalisation, it may be that in the absence of selective or specific deixis, to use a shell noun alone is insufficient to create the expectation that its lexicalisation will follow. The modifier can serve different functions: it can highlight the location of the lexicalisation, for example by selecting a Post-Deictic such as *following*; it can evaluate the importance of the lexicalisation or some other aspect of it, for example by means of an Epithet such as *main*, or it can specify the shell noun such that its relevance to the text is established, for example by means of a Classifier. In the case of Numeratives, a cohesive tie can be created by selecting a conjunctive Adjunct such as *Firstly* or *Secondly* as textual Theme for the lexicalisation, or construing the Subject/Theme with a Numerative such as *first* or *second*. Both sub-corpora include instances of these selections.

While shell nouns as elements in a Participant, Attribute or Circumstance typically enter into free and bound clauses at clause rank, they are also found in rank-shifted clauses that construe a Participant (fact clauses) or Qualifier (embedded elaborating clauses). The Published sub-corpus shows a higher proportion of free clauses and lower proportion of bound clauses containing shell nouns than the

Submissions sub-corpus. These results are significant at the medium and high levels respectively. The varied ranks in which shell nouns enter illustrate their flexibility as a resource for making meaning.

The analysis of shell nouns at the rank of the lexicogrammar shows that they warrant the term given to them by Francis (1994): 'discourse organisers'. At a local level they constitute a resource for moving the text on by means of pointing backwards or forwards. Anaphorically referring shell nouns provide a shared point of departure for the reader while other shell nouns combine with selections from the group level to point to upcoming text. These uses represent Schmid's(2000) textual function. As a semiotic abstraction, the shell noun can project a clause as Qualifier, thus enabling complex information to be realised in a single clause. This use represents Schmid's cognitive function. Furthermore, the fact that shell nouns appear in nominal groups enables them to construe a range of Participant functions as well as enter into Circumstances. As a result, they are open to full interplay with the resources of the information structure. Like the published writers, the EAL academics in this study make full use of these functions, and despite the grammatical infelicities in some of the instances, the meaning is clear. This result contrasts with findings of studies that focus on learner writers (e.g.Straker, 2010).

This chapter has focused on the contribution of shell nouns to the discourse at the rank of the clause. The following chapter will present the results for the analysis at the stratum of discourse semantics.

Chapter 6. Results and discussion of shell nouns at the discourse semantics stratum

The analysis of shell nouns at the discourse semantics stratum seeks to explore how shell nouns contribute to meaning at a level higher than that of the clause. It aims to answer the research question: In what ways do shell nouns contribute to experiential, interpersonal and textual meanings at the discourse semantics stratum? The analysis is approached using the principles of the discourse semantic systems developed by Martin (1992), Martin and White (2005) and Hood (2010). This chapter presents the results of the analysis. The presentation is made somewhat complicated due to the fact that it is difficult to approach each metafunction without considering instances that involve coupling with resources from the other metafunctions. The presentation is therefore organised around each metafunction, but will include descriptions of the interplay with the other metafunctions. The results for the textual function will be presented first, followed by those of the experiential function and then the interpersonal. A summary of the results will follow, in which there is incorporated a discussion of how they relate to the results of other studies.

6.1. Shell nouns and textual meaning at the discourse semantics stratum

Although the analysis of the function of shell nouns in the lexicogrammar in chapter 5 made reference to the fact that shell nouns can refer anaphorically, and cataphorically, esphorically, i.e. the lexicalisation is realised in the immediate environment of the shell noun within the group or clause, the analysis of their textual function across clauses was not systematic. This section will explore the ways in which shell nouns contribute to textual meaning beyond that of the clause. It is divided in three parts: the interface with reference, hyper-Themes and hyper-News, and internal conjunction.

6.1.1. Reference

At the stratum of discourse semantics one of the functions of the textual system is to identify and track Participants in the discourse. In the case of shell nouns in this study, because of the strict definition given at the outset, i.e. that they are lexicalised by a clause or another nominal group with shell noun as Head, the nominal

group with a shell noun as Head is likely to be entering the discourse for the first time. The identification of a shell noun in this study distinguishes between neutralised, presenting and presuming reference (Martin, 1992).

A preliminary analysis shows that there are significant differences between the Published and Submissions sub-corpora as to how a nominal group with shell noun as Head is introduced into the discourse; the Submissions has a higher relative share of neutralised reference and a correspondingly lower relative share of effected reference (Appendix F.1). This difference stems from the greater use of headings, which take neutralised reference, in the Submissions sub-corpus. The headings include “Objectives” or “Limitations of existing measures and motivations for model-based indices” from P_Pov, for example. The use of shell nouns as headings will be dealt with in the section on macro- and hyper-Themes. Of the remaining instances in which reference is effected, three quarters of them presume the reader can identify them by recourse to the surrounding co-text, and one quarter use presenting reference. Although these results suggest that it is through presuming reference that a cohesive tie is established between the shell noun and its lexicalisation, they bring little insight into the full range of textual functions of shell nouns at the discourse semantics stratum. Exploration of the textual relation of the lexicalisation, the shell noun and its co-text is more informative. As noted in chapter 5, which explores the functions of shell nouns at the stratum of the lexicogrammar, shell nouns can be lexicalised anaphorically, esphorically or cataphorically. This section will explore how the type of lexicalisation contributes to the unfolding discourse and how the two sub-corpora differ.

Analysis of the distribution of the direction of the referent lexicalisation reveals that there are significant differences between the two sub-corpora. While cataphoric lexicalisations account for around a quarter of instances in both sub-corpora, in comparison with the Published sub-corpus, the Submissions sub-corpus has a lower proportion of esphoric lexicalisations and a higher proportion of anaphoric lexicalisations. These differences attain high statistical significance (table 6.1). As detailed in chapter 5, esphoric lexicalisations can be realised within the nominal group with shell noun as Head, i.e. construing Qualifier, or within the clause, as a fact clause.

Breaking the esphoric instances down by the type of lexicalisation explains this difference. In around a quarter of all the instances of shell nouns in each sub-corpus, the esphoric lexicalisation construes Qualifier. However, the Submissions sub-corpus has a lower proportion of esphoric lexicalisations realised as a fact clause, a result that attains high statistical significance (Appendix F.2).

	Published		Submissions					
DIRECTION	N=284		N=406					
esphoric	45.4%	129	33.0%	134	3.33	+++	10.92	+++
anaphoric	31.7%	90	42.4%	172	2.86	+++	8.08	+++
catphoric	22.9%	65	24.6%	100	0.53		0.28	

Table 6.1 Distribution of direction of LEXICALISATION of shell nouns in Published and Submissions

The results in table 1 suggest that the Portuguese writers are making greater use of shell nouns as a textual resource to compact information previously construed as a clause, and in doing so, move the discourse along. This use is exemplified in 6.1. In the example the lexicalisation of the shell noun is underlined, the shell noun is in bold, and the anaphorically referring nominal group is boxed.

6.1 Fourth, the questionnaire opted for a national random sample, with a response rate of 64.8%, which was considered an acceptable sample of respondents (Dillman, 1978). **This procedure** ensures the generalizability of the data, ... (S_CEO)

In 6.1 ‘This **procedure**’ refers back to the preceding clause complex, but the information is condensed in the nominal group, thus providing a known point of departure (Theme) for the next clause. Presuming reference ‘This’ tells the reader that the information is known, and the combination of presuming reference and the shell noun **procedure** tells the reader that the information about the procedure is available in the co-text.

The anaphorically referring shell nominal group does not have to compose Theme to contribute cohesively to the development of the text. It can also appear within Rheme, as in 6.2, where the anaphorically referring nominal group (boxed) construes Verbiage. The lexicalisation of the shell noun **results** is underlined.

6.2 Yermack (1986) demonstrates that larger boards are detrimental to performance, although Bhagat and Black (1999) fail to confirm **the results**. (S_CEO)

Again a cohesive tie is established through presuming reference, and the reader is alerted to the need to look in the co-text to find the lexicalisation.

In most of the instances of anaphoric reference, the link between the shell noun and its lexicalisation operates across sentence boundaries, and the lexicalisation is found in the adjoining clause, that is, in the final clause of the preceding sentence. This is not to say that the lexicalisation is limited to the preceding clause; it may stretch across more than one clause or even more than one sentence. Example 6.3 illustrates a lexicalisation appearing in the adjoining clause of the preceding sentence. In the example, as in the following ones, the lexicalisation is underlined, the shell noun is in bold, and the anaphorically referring nominal group with shell noun as Head is boxed.

6.3 Finally, once computed, the index lacks a measure of uncertainty. For policy making decisions in particular, **this last feature** may be problematic ... (P_Pov)

The anaphoric link may also be to the preceding clause within a clause complex, as example 6.4 shows.

6.4 Nonprofits that spend higher fractions of their total expenses on actual program services tend to pay their top officers, directors and trustees and top nonofficer, nondirector and nontrustee staff more in the cross section, but further analysis suggests that **this result** is not particularly robust. (P_CEO)

In the corpus there are a number of instances in which there is one clause or more that intervenes between the lexicalisation and the anaphorically referring shell nominal group.

6.5 In the 1750s four actors belonged to all three institutions and these actors used their advantageous position for career progression or investment in the slave trade. For example, the one slaving voyage that Robert Armitage organised in 1784 was financed by four investors, all of whom were members of the African Committee, Town Council or both. Matthew Stronge organised two voyages, and of the eight investors, four belonged to the African Committee, Ugly Face Club/Mock Corporation and Town Council. Thomas Rumbold belonged to the African Committee-Mock Corporation subnet and organised three voyages involving seventeen investors, fourteen of whom belonged to the same institutions as Rumbold. Robert Clay organised seven voyages involving eight investors, five of whom belonged to the same institutions as he did. **This trend** continued into the 1760s, if slightly to a lesser extent. (P_Net)

In 6.5 ‘This **trend**’ refers to the trend to belong to the same institutions and use the “advantageous position for career progression or investment in the slave trade”, which is realised several clauses earlier. The intervening clauses elaborate on the lexicalisation by giving examples. In each case of these anaphorically referring shell nouns there is a cohesive tie between the shell noun and its lexicalisation which is construed by presuming reference, i.e. it is presumed that the reader will recover the referent from the co-text. Presuming reference, construed as Deictic, can be realised as directed reference, *this*, undirected reference, *the*, or a possessive such as *Morgensten’s*.

Comparison between the two sub-corpora shows that while lexicalisations that immediately precede the clause that contains the shell noun account for most of the instances of anaphorically referring shell nouns, the Submissions sub-corpus has a higher proportion of lexicalisations that are separated from the shell noun by one clause or more and a correspondingly lower relative share of lexicalisations in the preceding clause. The statistics are reported in Appendix F.3. There are also twelve instances – four in Published and eight in Submissions – where the lexicalisation is partially realised in the same clause (6.6).

6.6 Travelling agents could collect information on the spot and assist the resident ones. Such a function was undertaken by merchant commissioners with good reputation (hubs in the network). This is **the reason** for mixed agency, integrated in Category B, becoming the pattern that assured best economic performance. (S_Net)

In 6.6 there is a cohesive tie across sentences that is composed by means of text reference, ‘This’, and the shell noun construes Value in an identifying relational Process clause. In this instance, the use is agnate to a cohesive conjunction, as exemplified in 6.6’.

6.6’ Travelling agents could collect information on the spot and assist the resident ones. Such a function was undertaken by merchant commissioners with good reputation (hubs in the network). **For this reason**, mixed agency, integrated in Category B, became the pattern that assured best economic performance.

Only two shell nouns are found to be used in this way: **reason** and **case**. By contrast a wider variety of shell nouns is found in a superficially similar

lexicogrammatical pattern, where, unlike 6.6, the shell noun does not take presuming reference (6.7).

6.7 ... according to Divisia (whom, as previously indicated, was not familiar with high mathematics), Marschak would not know a partial derivative²⁰, but Frisch rapidly understood that this was just an unfair insinuation. (S_Econ)

In 6.7 a cohesive tie is again achieved by means of text reference, but the shell noun takes presenting reference, and the shell noun construes, not Value, but Attribute in an attributive relational Process clause. Although for ease of identification the shell noun was tagged anaphoric, it cannot be said that it *refers* anaphorically to the underlined text given that there is no cohesive tie in the nominal group with shell noun as Head, and the relation is construed by means of a relational Process clause. Rather, the shell noun can be said to be *lexicalised* by the clause. Similarity to reference is apparent by the paraphrase: What insinuation? The insinuation that Marschak would not know a partial derivative. This textual pattern – an anaphorically lexicalised shell noun realised with presenting reference, construing Attribute, in an attributive relational Process clause – is always interpersonally orientated. In 6.7 Divisia’s insinuation that Marschak would not know a partial derivative is assessed as ‘unfair’. The interpersonal contribution to the discourse of such nominal groups with shell noun as Head depends on the shell noun and/or its modifiers, and this aspect will be addressed in the section 6.3 on the contribution of shell nouns to interpersonal meaning. Table 6.2 shows the instances of this textual pattern that are found in the corpus (table 6.2).

Shell noun type	Published	Submissions
factual		this is not a very illuminating result (S_Rand)
linguistic	This is, of course, a metaphor (P_Econ)	This was hardly an encouragement (S_Econ) [This was] even less a commitment (S_Econ) this was an unfair insinuation (S_Econ)
mental	This was a ‘scientific’ theory (P_Econ)	This is a well enough discussed subject (S_Net)
eventive		this was an immense success (S_Econ)

Table 6.2 Anaphorically lexicalised shell nouns as Attribute

The realisations of the lexicalisations of anaphorically referring shell nouns are in the main clausal, although there are some instances in which they refer to a nominal group with shell noun as Head that has previously been lexicalised as a clause. This is shown in 6.8.

- 6.8 In pay performance contracts, CEO compensation is an increasing function of performance measured by accounting numbers (Ames, 2003). Organisational non-profit operational insufficient performance may restricts the CEO's opportunity to increase his earnings. Under **this hypothesis**, CEO compensation is a function of non-profit performance. Core et al. (1999) and Gosh and Sirmons (2005) have tested **this hypothesis**. (S_CEO)

Such chains seldom go beyond two and tend to employ the same shell noun. Clausal lexicalisations are found to be realised by a single clause (6.9), a clause complex (6.10), or more than one clause that stretches across sentence boundaries (6.11). The following examples (6.9-6.11) illustrate these various realisations. Once again the cohesive tie is composed through presuming reference in each case, and the anaphorically referring nominal group is boxed, the lexicalisation underlined, and the shell noun is in bold.

- 6.9 The regressions in table 5, Panel B, show that the return to assets is much higher in for-profit firms than in nonprofits. **My results** are consistent with those of other studies on CEO pay in firms. (P_CEO)
- 6.10 Fourth, the questionnaire opted for a national random sample, with a response rate of 64.8%, which was considered an acceptable sample of respondents (Dillman, 1978). **This procedure** ensures the generalizability of the data, ... (S_CEO)
- 6.11 The co-variables determinant for the amounts of gold received were Category B and number of strong ties. The other co-variables had no relevant effect on the dependent variable when considered together (appendix 1). Category B defines the strategy defined by mixed agency, meaning resident agents were regularly assisted by people coming back and forth⁴⁴. Thus managing distance could not renounce setting up contacts with someone rooted in the colony.
The historical fundamentals of **these results** deserve a more detailed explanation. (S_Net)

In some cases directed reference (*this*) may not be thought specific enough for the reader to recover the referent from the preceding co-text. While it may be

expected that this is the case when the lexicalisation is more than one clause distant from the shell noun (6.12), it is not always so.

6.12 According to Roth and Vande Vate (1990), once the probability that a given pair of agents meets is bounded away from zero, each play of the game yields a matching stable with respect to the revealed orderings in the course of that play. Hence, given a profile of strategies that meets the above requirement, every outcome obtained with positive probability is stable for some revealed profile of preferences. (S_Rand)

In 6.12 the post-Deictic ‘above’ locates the lexicalisation as preceding the shell noun **requirement**. In a discourse semantics perspective, such post-Deictics identify the entity as belonging to a superset, and, rather than employ directed reference, ‘this’, the writer references the lexicalisation with undirected ‘the’. In fact, with the exception of a single instance from the Published sub-corpus, in all the instances in which there is reference to a superset, undirected reference is the norm. A reference to location may also be realised as a prepositional phrase that construes Qualifier to a shell noun, such as ‘in the preceding section’. In addition to location, a superset may be implied by an ordinal such as ‘first’ or ‘second’, or a post-Deictic such as ‘former’, ‘latter’ or ‘final’. In anaphorically referring nominal groups with shell noun as Head the following instances of reference to a superset are found, some of which occur more than once (table 6.3).

Reference by location		Reference by order	
Published	Submissions	Published	Submissions
	the above examples (S_Rand) the above definition (S_Rand) the above result (S_Rand) the results of the preceding section (S_Rand) the example above (S_Rand) the above requirement (S_Rand) the above results (S_Rand) the four solutions previously pointed out in section 2 (S_Net)	this last feature (P_Pov)	the latter case (S_Rand) the initial assumption (S_Rand) the first hypothesis (S_Net) the second hypothesis (S_Net) the first alternatives (S_Net)

Table 6.3 Textual reference to a superset with anaphorically lexicalised shell nouns in Published and Submissions

These modifiers to shell nouns are textual in that they contribute to the identification of the lexicalisation by means of the organisation of text. They are restricted to three texts, and interestingly, each writer seems to have a preferred

strategy for organising text: S_Rand seems to prefer reference by location whereas S_Net seems to prefer reference by order.

As previously mentioned, anaphorically referring nominal groups with shell noun as Head often enact interpersonal meaning as well as textual. This is the case in over three quarters of the instances of anaphoric reference in both sub-corpora (Appendix F.4). A mental or linguistic shell noun may contribute to the dialogic positioning of the writer, as 6.13 shows.

6.13 From a conceptual point of view, government grants may mean one of several different things for the pay of managers of nonprofits. First, they could be a sign of organizational quality or managerial ability. That is, in the cross section, we might expect that those organizations with grants or those with larger grants should have managers who are paid more. However, within organizations (where industry type, organization quality, and managerial ability are fixed) more government grants may be a sign of some sort of outside monitoring of the manager or an increase in funding of the manager. These **ideas** [heterogloss: expand: entertain] suggest the following simple modification of the earlier equation. (P_CEO)

‘These **ideas**’ refers to the propositions of the preceding text and presents them as entertained by the writer. Thus ‘ideas’ renders the text heteroglossic, opening up the dialogic space between the writer and other voices.

Over a third of the anaphorically referring shell nouns in the Published sub-corpus, and just over half of those in Submissions contribute to dialogic positioning. By contrast, this trend is reversed for the system of ATTITUDE, where, in the Published, almost half of the instances of anaphorically referring shell nouns couple with resources from ATTITUDE, and only a third of the instances in Submissions do so. The propensity of each sub-corpus to couple anaphorically referring shell nouns with resources from ATTITUDE or ENGAGEMENT is found to be statistically significant to the medium level (Appendix F.4). These results indicate that the Portuguese writers rely more heavily on anaphorically referring shell nominal groups to position themselves dialogically than to appraise an anaphorically referring shell noun through the system of ATTITUDE.

When an anaphorically referring shell noun is the object of appraisal, the evaluation can extend to encompass the lexicalisation. Attitudinal value can, for

example, be infused in the shell noun. In the following example the negatively inscribed shell noun ‘**difficulty**’ builds on a negative evaluation of the behaviour of the members of the Econometrics Society. In the examples the lexicalisation of the shell noun is underlined, the shell noun is in bold, the nominal group is boxed and the attitudinal resources are in italics.

6.14 In fact, not only did each of them pursue their own research agenda with little connection with each other’s, but they also had different visions as far as the future of the Society was concerned. This **difficulty** emerged immediately in 1931 as the Society took its first task: (S_Econ)

Alternatively the evaluative element may be realised elsewhere as Attribute:

6.15 The above **result** is as *uninformative* as large the set of individually rational matchings may be. (S_Rand)

Or invoked:

6.16 Several variables affect the CEO, namely, organization performance variables, board composition variables and individual variables. The present **result** *sheds light on* governance problems in Portuguese NPOs and *highlights* the urgent need for a code of governance practice to be introduced in this sector. (S_CEO)

In 6.16 the positive attitudinal value in ‘sheds light on’ invokes a positive value for ‘The present **result**’. In addition there are elements of GRADUATION which propagate the positive prosody in the co-text: graduated force (highlights = makes *very* clear; *urgent*).

Comparison between the Published and Submissions sub-corpora of the distribution of attitudinal resources in anaphorically referring shell nouns reveals that there are more instances of negative attitudinal values than positive in the Published sub-corpus whereas in the Submissions this trend is reversed. The full statistics are given in Appendix F.4. At this point I shall not attempt a discussion of these results. Rather, their significance will be discussed in the summary of results at the end of this chapter.

The distribution of shell nouns in anaphorically referring nominal groups is also found to differ between the two sub-corpora. In Published circumstantial shell nouns such as **approach**, **criteria** or **constraint** account for the greatest share, followed by factual shell nouns, then mental shell nouns, and linguistic shell nouns. In the

Submissions sub-corpus factual shell nouns are the most common, followed by mental shell nouns, linguistic shell nouns and then circumstantial shell nouns. Statistically, the relative high share of circumstantial shell nouns in the Published, and relative low share in the Submissions is deemed significant (table 6.4).

	Published		Submissions					
Feature	Percent	N	Percent	N	T Stat	Signif.	ChiSqu	Signif.
SHELL-NOUN	N=90		N=171					
factual	25.6%	23	29.8%	51	0.73		0.53	
linguistic	15.6%	14	22.8%	39	1.38		1.92	
mental	23.3%	21	28.1%	48	0.82		0.68	
modal	7.8%	7	2.9%	5	1.78	+	3.17	+
eventive	1.1%	1	3.5%	6	1.14		1.30	
circumstantial	26.7%	24	12.9%	22	2.81	+++	7.74	+++

Table 6.4 Anaphorically referring shell noun by shell noun type in Published and Submissions

Around a third of the shell nouns in the corpus are lexicalised cataphorically (see table 6.1), but this result includes shell nouns that appear in headings, which are minor clauses, as well as those in major clauses (Halliday & Matthiessen, 2004: 153). The following results pertain to cataphorically lexicalised shell nouns that appear in major clauses with effected reference. A cataphorically lexicalised shell noun may be introduced by presuming or presenting reference. Example 6.17 shows an example of presuming reference. The shell noun is in bold, the nominal group is boxed and the lexicalisation is underlined.

6.17 This is not to say that such institutions were not without **their problems**. They often became victims of their own success and protected vested interests. (P_Net)

In 6.17 a cohesive tie that links back to ‘such institutions’ is established by means of the possessive Deictic ‘their’ and the pronoun ‘They’. However, the **problems** of the institutions are clausal in nature, i.e. that ‘They often became victims of their own success and protected vested interests’. Presenting reference is illustrated in 6.18.

6.18 Hence the neo-classical research program at the turn of the century faced **an unpleasant choice**: either persist in a determinism bereft of the kind of results one associated with physics or else throw in the entire program of copying physics. (P_Econ)

Despite the absence of presuming reference, the lack of specificity in the shell noun **choice** indicates to the reader that he must look for the relevant meaning in the co-text. Because nothing appropriate has been given in the preceding text, it must appear in the following (Winter, 1977). The results of the analysis show that cataphorically lexicalised shell nouns are typically introduced with presenting reference, and there is little difference between the two sub-corpora in the distribution across textual functions of these cataphorically lexicalised shell nouns. (Appendix F.5). With regard to the distance between the shell noun and its lexicalisation, like the anaphorically referring shell nouns, most of the lexicalisations – three quarters or more – are realised in the adjoining clause, the link operating across sentence boundaries. This is exemplified in 6.19. The cataphorically lexicalised shell noun is in bold and the lexicalisation is underlined.

6.19 In what the above results is concerned, **a couple of remarks** is in order. The first observation concerns fairness and random matching mechanisms. In opposition to deterministic mechanisms, which are bound to favour one side of the market over the other, we have claimed that random mechanisms promote procedural fairness. Nevertheless, endstate justice is a different issue. + 8 clause complexes. (S_Rand)

In the Published sub-corpus around three fifths of the instances of cataphorically lexicalised shell nouns are of this type, about a fifth are realised more than one clause distant to the shell noun, and around one fifth are in an adjoining clause within a clause complex, i.e. there is no intersentential link. The results are slightly different for the Submissions: under half of the instances of cataphoric reference are to the adjoining clause of the next sentence, around a third to an adjoining clause within the clause complex, and a further quarter have a clause or more intervening between the shell noun and its cataphoric lexicalisation. The results for the statistically tested comparison between the two sub-corpora are reported in detail in the Appendix F.5.

Most of these cataphorically lexicalised shell nouns compose hyperThemes, although some of them operate more locally, with the lexicalisation spanning only a clause or clause complex. These general textual functions apply to cases where the lexicalisation immediately follows in the next clause as well as when there is a clause

or more intervening. In a few cases there is a large body of text between the shell noun and its lexicalisation. This is exemplified in 6.20.

6.20 **Our results** are presented and discussed in Section 6. Finally in Section 7 we draw **some conclusions** and consider **the implications of our findings** for further research. (S_Pov)

Example 6.20 comes from the end of the introduction to the research article and the shell nouns point to information in later sections of the text. Such uses might be considered more marginal examples of shell nouns, but they do commit the writer to instantiating the results, conclusions, implications and findings in Sections 6 and 7.

With regard to the type of realisation of these cataphorically lexicalised shell nouns, there are no significant differences between the Published and Submissions. Most of the lexicalisations – more than three quarters – are realised as extended text, i.e. they are made up of more than one sentence. Less than one quarter of the lexicalisations are realised as a clause complex, and a few are realised as a single clause. (See Appendix F.5.) An example of each kind of lexicalisation is given in 6.21 – 6.23. The cataphorically referring nominal group with shell noun as Head is boxed, the shell noun is in bold, and the lexicalisation is underlined.

6.21 **Further corroboration**, however, comes from Brickley et al. (1994). External directors enhance shareholder gains in tender offers. (S_CEO)

6.22 **The paradox of the situation** was obvious: Snyder, Mitchell and Taussig were members of the society yet were suspected of not favouring the financing of its projects and activities. (S_Econ)

6.23 Hansmann outlines **several reasons**. The first is that the buyer and the recipient of goods and services are not always the same person. For example people are more likely to donate funds to a relief organization organized as a nonprofit because they know that the managers are less likely to abscond with the residual money. Another reason that an organization may form as a nonprofit comes out of the idea of public goods (Hansmann, 1980). The fact that one person enjoys the services of a public good does not preclude others from doing the same. For example donors are more likely to contribute to nonprofit radio stations because they believe that the funds will be used for programming rather than for the private benefit of the managers. Museums and operas are similar in that they practice “voluntary price discrimination” (Hansmann 1980) by offering low-priced seats, but those who can afford to contribute do so with

the knowledge that the residual money cannot be distributed to those in control. Hansmann (1980) also describes “implicit loans” in higher education, where students are more likely to donate money back to their nonprofit colleges when they feel that those in charge cannot keep the money for themselves. (P_CEO)

In each case the shell noun announces and interprets the upcoming information, and the instances show coupling with interpersonal resources of Appraisal. This aspect will be discussed shortly. In 6.21 the lexicalisation is realised as a single clause, in 6.22 as a clause complex, and in 6.23 as extended text. The first two instances illustrate shell nouns operating at a more local level than in hyper-Themes for the lexicalisation is not developed beyond a clause or clause complex. In 6.23 the clause containing the cataphorically lexicalised shell noun composes a hyper-Theme, and the paragraph develops using more shell nouns as well as the resources of internal conjunction. Hyper-Themes and internal conjunction will be discussed in sections 6.1.2 and 6.1.3 respectively.

When shell nouns operate locally, two different shell nouns may be cataphorically lexicalised by the same lexicalisation. In 6.24, the lexicalisation, underlined, is shelled first as ‘a **miracle**’, and then as ‘a complete **surprise**’.

6.24 A **miracle** was needed to publish *Econometrica* and it came in the form of a complete **surprise**: Alfred Cowles III, the son of a millionaire, the president of an investment counselling firm, Cowles and Co., and a competent statistician interested in stock market predictions, offered to pay twelve thousand dollars a year for the journal. (S_Econ)

In 6.24 the nominal group with the shell noun **miracle** as Head is reinterpreted as ‘a complete **surprise**’ in the second clause of a paratactic clause nexus, and its lexicalisation appears as the final clause of the clause complex. Such reinterpretation followed by the lexicalisation is found in both sub-corpora. Again, there is interpersonal meaning being enacted in the nominal groups with shell nouns as Head: **miracle** is infused with positive value and **surprise** is graded; its focus is sharpened by means of the Epithet *complete*.

The relation of these cataphorically lexicalised shell nouns with the systems of IDENTIFICATION and APPRAISAL is worthy of note. Unlike anaphorically referring shell nouns, which overwhelmingly take presuming reference, almost half of the

cataphorically lexicalised shell nouns take presenting reference, and of the 55 instances, 51 involve interpersonal resources from APPRAISAL. In addition to selecting shell nouns that function to render the text heteroglossic, the interpersonal meaning may be realised by means of resources from the systems of GRADUATION or ATTITUDE. Graduation may be expressed as FORCE, as in 6.25, where the elements from Graduation are in italics, the shell noun is in bold, the cataphorically referring nominal group is boxed, and the lexicalisation is underlined.

6.25 Despite their intuitive appeal, empirical indices such as Townsend's are based on *a number of* important structural **assumptions**. Because the index is an unweighted sum of the Z scores, each variable contributes equally to measured deprivation; furthermore, even though areas may have different population sizes, no adjustment is made for differential precision of the component variables across areas. Moreover, it is assumed that for a specific area, information on deprivation depends exclusively on variables from that area, and not (for example) on variables from neighboring areas. Finally, once computed, the index lacks a measure of uncertainty. For policy making decisions in particular, this last feature may be problematic, for example, if decisions about resource allocation are based on cut off variables or percentiles of the index. (P_Pov)

There is a lot of evaluative work going on in 6.25 as the writer positions the reader to evaluate empirical indices like Townsend's negatively. The reader is alerted to an attitudinal reading by means of the resources of ENGAGEMENT, 'despite', which contracts the discourse, and the **assumptions**, which are to be lexicalised cataphorically are graded as quantity: 'a number of'; appraised as valuation: 'important'; and classified ideationally as 'structural'. A negative assessment of the structural assumptions and therefore the indices is propagated in the lexicalisation through resources from ENGAGEMENT that contract the discourse: 'no adjustment', 'not on variables from neighboring areas', as well as negatively inscribed selections from ATTITUDE: 'lacks', 'problematic'. The grading of the shell noun as FORCE thus helps the evaluative meanings to be maintained.

The most commonly used cataphorically lexicalised shell noun to couple with resources from APPRAISAL: FORCE is **reasons**. This shell noun is found in both sub-corpora. Table 6.5 lists the cataphorically lexicalised shell nouns that couple with elements from APPRAISAL: FORCE in Published and Submissions. The element of FORCE is in italics.

Shell noun type	Published	Submissions
factual	<i>several</i> improvements (P_Pov) <i>some</i> specific recommendations (P_Pov) <i>several</i> advantages (P_Pov) <i>a variety of</i> reasons (P_CEO) <i>at least two</i> reasons (P_CEO) <i>at least two</i> key features of nonprofits that distinguish them from for-profit organizations (P_CEO) <i>several</i> reasons (P_CEO) <i>many</i> reasons (P_Net) <i>some</i> brief biographical evidence (P_Econ)	<i>some</i> results that extend to deterministic mechanisms (S_Rand) <i>two</i> reasons (S_Net) <i>Three</i> main aspects (S_Net) <i>Further</i> corroboration (S_CEO) <i>Four</i> elements (S_CEO) <i>a battery of</i> reasons (S_Econ) <i>some of</i> the main findings (S_Pov)
linguistic	<i>some</i> conclusions (P_Net) <i>some</i> bad news (P_Econ)	<i>Some</i> concluding remarks (S_Rand) <i>Two</i> remarks (S_Rand) <i>a couple of</i> remarks (S_Rand) <i>a more</i> detailed explanation (S_Net) <i>some</i> conclusions (S_Pov)
mental	<i>a number of</i> important structural assumptions (P_Pov) <i>at least two</i> motivations (P_Pov) <i>a number of points of</i> view (P_Econ)	<i>several</i> issues (S_CEO) <i>several</i> steps (S_CEO) <i>several</i> strategies (S_Pov)
circumstantial		<i>two</i> ways (S_Rand) <i>three</i> different fields (S_Econ) <i>two</i> criteria (S_Econ)

Table 6.5 Cataphorically lexicalised shell nouns that couple with resources from APPRAISAL: FORCE in Published and Submissions.

Cataphorically lexicalised shell nouns with presenting reference also couple with resources from GRADUATION: FOCUS. Example 6.26 illustrates this option. Again the elements from GRADUATION are in italics, the shell noun is in bold, the cataphorically referring nominal group is boxed, and the cataphoric lexicalisation is underlined.

6.26 A miracle was needed to publish *Econometrica* and it came in the form of a complete surprise: *Alfred Cowles III, the son of a millionaire, the president of an investment counselling firm, Cowles and Co., and a competent statistician interested in stock market predictions, offered to pay twelve thousand dollars a year for the journal.* (S_Econ)

The Epithet ‘complete’ sharpens the focus of the shell noun **surprise** by characterising it as fully actualised. Instances of both sharpening and softening the focus of a cataphorically lexicalised shell noun are found in the corpus. They are listed in table 6.6. The element of GRADUATION: FOCUS is in italics.

Shell noun type	Published	Submissions
factual	Another <i>possible</i> reason (P_CEO)	a <i>complete</i> surprise (S_Econ)
linguistic	some <i>specific</i> recommendations (P_Pov)	a <i>partially</i> converse statement (S_Rand) a <i>particular</i> remark (S_Net) an <i>apparently</i> naive suggestion (S_Econ)
circumstantial		a <i>different</i> stability condition in the game induced by a random stable mechanism (S_Rand)

Table 6.6 Cataphorically lexicalised shell nouns coupling with resources from GRADUATION: FOCUS

Alternatively, cataphorically lexicalised shell nouns are found to draw on resources from the system of ATTITUDE. This is construed through Epithets, which inscribe the shell noun with positive or negative value. The prosodic value then spreads to the lexicalisation that follows the shell noun. In 6.27 the shell noun **choice** is appraised negatively as ‘unpleasant’.

6.27 Hence the neo-classical research program at the turn of the century faced an unpleasant choice: either persist in a determinism bereft of the kind of results one associated with physics or else throw in the entire program of copying physics. (P_Econ)

The shell noun here seems to provide scaffolding for the evaluation. While it is possible to simply delete the Epithet, (6.27’), it is also possible to avoid using a shell noun altogether (6.27’). However, both these options lack the writer’s assessment of the situation.

6,27’ Hence the neo-classical research program at the turn of the century faced a **choice: either persist in a determinism bereft of the kind of results one associated with physics or else throw in the entire program of copying physics.**

6.27’’ Hence the neo-classical research program at the turn of the century could either persist in a determinism bereft of the kind of results one associated with physics or else throw in the entire program of copying physics.

The cataphorically lexicalised shell nouns that take presenting reference and are inscribed with ATTITUDE are listed in table 6.7. The attitudinal resources are in italics.

Shell noun type	Published	Submissions
factual	at least two <i>key</i> features of nonprofits that distinguish them from for-profit organizations (P_CEO) a <i>serious</i> dilemma (P_Econ) a <i>simple</i> example (P_Rand)	Three <i>main</i> aspects (S_Net) some of the <i>main</i> findings (S_Pov)
linguistic	some <i>bad</i> news (P_Econ)	a more <i>detailed</i> explanation (S_Net) <i>simple</i> proposals (S_Econ) an apparently <i>naive</i> suggestion (S_Econ)
eventive	an <i>unpleasant</i> choice (P_Econ)	

Table 6.7 Cataphorically lexicalised shell nouns coupling with ATTITUDE

There are also a few instances in which the cataphorically lexicalised shell noun is infused with positive or negative value. In 6.28 the writer evaluates the methodology he uses in his study positively by means of the shell noun **improvements**. The lexicalisation then elaborates on what the good things about the methodology are.

6.28 Index construction based on a model allows several **improvements** over Townsend’s and similarly constructed indices: (1) the index can be represented as a weighted sum of (standardized) census variables, with data-driven weights; (2) by using posterior summaries, the indices can be reported with corresponding measures of uncertainty; and (3) incorporating information from neighboring areas improves precision of the posterior parameter distributions. (P_Pov)

In most of these cases, the shell noun is inscribed with positive value. Only one instance of negative value is found in the Submissions sub-corpus, yet, by casting a negative assessment over the context of the study, which is construed by the lexicalisation, the writer positively appraises the importance of his own study (6.29).

6.29 The motivation for the present research stems from several **issues**. First the sole code of governance in Portugal was established by the Stock Exchange. This procedure is contrary to observed codes of governance worldwide, which are drawn up by governments, directors’ associations, managers’ associations, professional bodies and investors and are usually overseen by autonomous watchdogs (Aguilera and Cuervo-Cazurra, 2004). Hence we can assert that Portugal lacks necessary codes of practice, there being no code whatsoever for non-profit organizations, which may affect negatively their efficiency. (S_CEO)

The following cataphorically lexicalised shell nouns that are infused with positive or negative value are found in the corpus (table 6.8). In the main, they are factual.

Shell noun type	Published	Submissions
factual	several improvements (P_Pov) several advantages (P_Pov)	a refinement of the Nash equilibrium concept (S_Rand) A miracle (S_Econ)
mental		several issues (S_CEO)

Table 6.8 Cataphorically lexicalised shell nouns infused with prosodic value in Published and Submissions

Similar to cataphorically lexicalised shell nouns that take presenting reference, those that take presuming reference are also found to couple with resources from APPRAISAL; however the selections are not necessarily the same. While resources from GRADUATION – in particular quantification – are commonly found with shell nouns with

presenting reference, they are much less frequent with shell nouns with presuming reference, where ordinal Numeratives tend to compose a textual function rather than interpersonal. In cataphorically lexicalised shell nouns with presuming reference, GRADUATION, which is boxed, is realised in terms of focus:

The first possible **reason** (P_CEO)
The next initiative (S_Econ)
the further bad **tidings** from the physicists' camp (P_Econ)

or force in intensified Epithets:

The fundamental **concerns** of corporate governance (P_CEO)
Our main **result** (S_Rand).

The realisations for ATTITUDE also differ. In addition to the options found in cataphorically lexicalised shell nouns with presenting reference, the evaluative element is found to be construed as Attribute. This is illustrated in 6.30, where the shell noun is in bold, the element from ATTITUDE boxed, and the cataphoric lexicalisation underlined.

6.30 The **paradox** of the situation was obvious: Snyder, Mitchell and Taussig were members of the society and yet were suspected of not favouring the financing of its projects and activities. (S_Econ)

About a fifth of cataphorically lexicalised shell nouns with presuming reference are found to involve reference to a superset. This is exemplified in 6.31, where reference to 'the next initiative' implies that there were other initiatives as well. In the example the textual reference to a superset is boxed, the shell noun in bold, and the lexicalisation underlined.

6.31 Yet the next **initiative** would not fail. In the autumn of 1927, Frisch met Schumpeter for the first time, at Harvard: their friendship and complicity in matters of the Society became a driving force behind the emerging movement. The American connection was to be the core of econometrics. In February of the next year, Frisch continued with his tour in support of econometrics and visited Irving Fisher at Yale, and then Charles Roos once again at Princeton: both would soon form part of the Society's first managerial board. Later that month, on February 29, Frisch met Schumpeter and Harberler at the Colonial Club in Harvard. (S_Econ)

Reference to a superset in cataphorically lexicalised shell nouns is restricted to four texts in the Published sub-corpus: P_Net, P_Econ, P_CEO, and P_Rand; and three in the Submissions: S_Econ, S_Pov, S_Rand. The shell nouns are varied. They include factual: **example, results**; linguistic: **observation**; mental: **analysis, theorem, lemma**; modal: **step**; eventive: **initiative**; and circumstantial shell nouns: **criteria**. The instances are listed in table 6.9. There seems to be a clear preference for positional reference with the cataphorically lexicalised shell nouns, in particular the post-Deictic ‘following’ appears repeatedly.

Reference by location		Reference by order	
Published	Submissions	Published	Submissions
the analysis below (P_Net)	the following natural way (S_Rand)	The first possible reason (P_CEO)	Our next results (S_Rand)
the following steps (P_Econ)	the following example (S_Rand)		The first observation (S_Rand)
The following existence theorem (P_Rand)	The results that follow (S_Rand)		the next initiative (S_Econ)
the following example (P_Rand)	The following lemma (S_Rand)		
	The following result (S_Rand)		
	the following criteria (S_Pov)		

Table 6.9 Cataphorically lexicalised shell nouns with reference to a superset

The distribution of shell noun types that function cataphorically shows that factual shell nouns make up the largest share, followed by linguistic, mental, circumstantial, eventive, and modal shell nouns, in decreasing order of frequency. Few of the shell nouns attain counts above 5; most are, in fact, single instances. The cataphorically lexicalised shell nouns that are most frequent in both sub-corpora are **result, reason, example, and remark**. In the case of the shell noun **proposition**, the high count is due to repeated instances in one text S_Rand. Table 6.10 shows the distribution of shell noun types in cataphorically referring nominal groups. No statistically significant differences are found between them. Appendix F.6 lists the shell nouns that compose a cataphoric reference function in major clauses in the corpus.

Feature	Cataphorically referring shell nouns in Published		Cataphorically referring shell nouns in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
SHELL-NOUN	N=45		N=68					
factual	42.2%	19	36.8%	25	0.58		0.34	
linguistic	22.2%	10	23.5%	16	0.16		0.03	
mental	20.0%	9	16.2%	11	0.52		0.27	
modal	2.2%	1	8.8%	6	1.43		2.03	
eventive	6.7%	3	2.9%	2	0.94		0.89	
circumstantial	6.7%	3	11.8%	8	0.89		0.80	

Table 6.10 Distribution of shell noun types in cataphorically lexicalised shell nouns in major clauses in Published and Submissions

Comparison of the distribution of shell nouns between anaphoric and cataphoric reference in major clauses across the corpus as a whole finds that they are not equally distributed. Factual shell nouns are more likely to compose cataphoric reference, a difference attains medium statistical significance. Further differences, which attain low statistical significance, are that circumstantial and mental shell nouns are more likely to compose anaphoric reference than cataphoric. These results are reported in table 6.11.

Feature	Anaphoric		Cataphoric		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
SHELL-NOUN	N=261		N=113					
factual	28.4%	74	38.9%	44	2.03	++	4.09	++
linguistic	20.3%	53	23.0%	26	0.59		0.35	
mental	26.4%	69	17.7%	20	1.83	+	3.32	+
modal	4.6%	12	6.2%	7	0.64		0.42	
eventive	2.7%	7	4.4%	5	0.88		0.77	
circumstantial	17.6%	46	9.7%	11	1.95	+	3.80	+

Table 6.11 Distribution of shell nouns in anaphoric and cataphoric functions in major clauses across whole corpus

This section has described the textual functions of shell nouns in major clauses at the stratum of discourse semantics with regard to how they link to either preceding or ensuing lexicalisations. In addition to anaphoric and cataphoric linking, shell nouns link esphorically, but because this is within the lexicogrammar of the clause, it is not addressed within the textual metafunction at the discourse semantics stratum, but

was described in sections 5.3.5 and 5.4.5 of chapter 5. Esphorically lexicalised shell nouns will nevertheless be taken into consideration in the sections on ideational and interpersonal meaning.

6.1.2. Macro-Themes, hyper-Themes and hyper-News

Analysis of the results shows that shell nouns constitute a resource for composing macro- and hyper-Themes in both the Published and Submissions sub-corpora, and that there are no significant differences in the way that they do this. In a research article, a macro-Theme may be realised by a heading. While it might be argued that shell nouns in headings are not true uses of shell nouns, in some cases a heading can be seen to be lexicalised by an ensuing clause or more. Consider 6.32, in which the shell nouns are in bold and the lexicalisations are underlined. Headings are boxed.

6.32 **Proposition 2**

Let φ be any individually random matching for (F,W,P) and let " φ " be a random stable matching rule. Then there exists an ordinal Nash equilibrium Q that supports φ in the game (P, φ, P) .

Proof

Let Q_w be such that $A(Q_w) = \{\varphi(w)\}$, for every $w \in W$, and let Q_f be such that $A(Q_f) = \{\varphi(f)\}$ for every $f \in F$. Clearly, $S(Q) = \{\varphi\}$ and φ is reached with probability one.... (S_Rand)

It is possible to paraphrase each of the headings with its lexicalisation in clausal form, as shown in 6.32'.

6.32' **Proposition 2** is to let φ be any individually random matching for (F,W,P) and let " φ " be a random stable matching rule. Then there exists an ordinal Nash equilibrium Q that supports φ in the game (P, φ, P) .

The **proof** is that if we let Q_w be such that $A(Q_w) = \{\varphi(w)\}$, for every $w \in W$, and we let Q_f be such that $A(Q_f) = \{\varphi(f)\}$ for every $f \in F$, then, clearly, $S(Q) = \{\varphi\}$ and φ is reached with probability one....

A heading with a shell noun may be realised as the shell noun without modification, like **Proof** in 6.32, or it may be modified with a Qualifier, as in 6.33.

6.33 **Reasons why pay levels for managers of non profits may be different**

The first possible reason for differences in pay between employees of non profit organizations and for profit firms was presented by Preston (1989) (P_CEO)

The heading in 6.33 clearly functions as a macro-Theme for the information that follows. The writer goes on to give three reasons why the pay levels might differ. Shell nouns as headings may also be modified with a Classifier, for example “Concluding **comments**” (P_CEO) or “Equilibrium **analysis**” (S_Rand). Shell nouns as headings do not occur in all texts of the corpus; rather they are concentrated in six: P_Rand, P_CEO, P_Pov, S_Rand, S_CEO, and S_Net; and they include mainly factual, linguistic and mental shell nouns. The shell nouns in macro-Themes all point forwards, in other words, the lexicalisation is realised cataphorically. They are listed in table 6.12.

Shell noun type	Published	Submissions
factual	features reasons Corollary Case Example	Proof Example Corollary Results
linguistic	comments Claim Remark	Definition Proposition
mental	motivations Analysis Objectives Theorem	analysis Lemma Hypothesis Hypotheses
circumstantial	Limitations	

6. 12 Shell nouns in headings composing macro-Themes in Published and Submissions

Hyper-Themes are realised as clauses. An example of a hyper-Theme is given in 6.34, where the hyper-Theme is boxed, reference to a superset is in italics, the shell noun is in bold, and its lexicalisation is underlined.

6.34 However, the present paper extends their contribution in *two ways*. First, we take equilibrium analysis further, going beyond the analysis of truncations. A concept of equilibrium based on first-order stochastic dominance is used, given that preferences are ordinal in nature and probability distributions over matchings are to be compared. The notion of ordinal Nash equilibrium guarantees that each agent plays his best response to the others' strategies for every utility representation of the preferences. ² Second the analysis is conducted in the context of the college admissions problem. ... (S_Rand)

In 6.34 the shell noun is lexicalised cataphorically. The Numerative, *two*, helps to predict the organisation of the lexicalisation, which also uses selections from internal conjunction. At the same time, it enacts an interpersonal function to grade the force of the shell noun in terms of quantity. While cataphorically lexicalised shell nouns account for the greatest share of hyper-Themes in both sub-corpora by far, there are

also instances in which the shell noun is realised esphorically. Example 6.36 illustrates a hyper-Theme in which the esphorically referring shell noun's lexicalisation is realised as a fact clause that construes Token (6.35). Although the lexicalisation for the shell noun **similarity** is realised esphorically, the entire clause acts as a hyper-Theme for the paragraph, the ensuing text elaborating on research on stable matchings in roommate problems and couples markets.

6.35 An important **similarity** between roommate problems and couples markets is that stable matchings need not exist. In a recent paper, Diamantoudi et al. (2004) showed that for roommate problems with stable matchings one can indeed construct an algorithm that will, starting from any matching, converge to a stable matching. Since an existing stable matching is explicitly used in the algorithm constructed by Diamantoudi et al. (2004), the algorithm is non-constructive and therefore differs considerably from Roth and Vande Vate's (1990) algorithm and the DPC algorithm for two sided matching markets. (P_Rand)

Shell nouns with esphoric lexicalisations construing Qualifier also feature in some hyper-Themes (6.36). The ensuing text elaborates on how the paper will analyse metropolitan networks and how it will measure them to assess change over time. Again the hyper-Theme is boxed, the shell noun is in bold, and its lexicalisation is underlined.

6.36 This paper pushes forward this more nuanced and sophisticated analysis of networks and represents the first serious **attempt to measure them to assess change over time**. It visualises Liverpool's metropolitan business networks comprising political, trade, social and cultural institutions during the period 1750-1810 and then measures the relationships between these individuals and groups. + 5 clause complexes (P_Net)

Most of the instances of hyper-Themes containing esphorically referring shell nouns occur in the Published sub-corpus. Only one instance is found in the Submissions.

The results for the distribution of shell nouns among hyper-Themes do not reveal statistically significant differences between the two sub-corpora (Appendix F.7). Factual shell nouns are the most common in both. The nominal groups with shell noun as Head functioning in hyper-Themes are reported in tables 6.13 and 6.14.

Shell noun type	Published	Submissions
factual	A classic example (P_CEO) the fact (P_Net) The problem with local and confidential networking (P_Net) An important similarity (P_Rand)	
mental		The aim of this paper (S_Pov)
eventive	the first serious attempt (P_Net) The struggle (P_Econ)	

Table 6.13 Esphorically referring shell nouns in hyper-Themes in Published and Submissions

Shell noun type	Published	Submissions
factual	a variety of reasons (P_CEO) at least two key features of nonprofits that distinguish them from for-profit organizations (P_CEO) the case of a manager of a nursing home (P_CEO) the results (P_CEO) many reasons (P_Net)	some results that extend to deterministic mechanisms (S_Rand) Four elements (S_CEO) the implications of these results (S_CEO) three different fields (S_Econ)
linguistic	some bad news (P_Econ)	a couple of remarks (S_Rand) a more detailed explanation (S_Net)
mental	a number of important structural assumptions (P_Pov) at least two motivations (P_Pov)	several issues (S_CEO)
modal		
eventive		the task this small group was setting itself (S_Econ)
circumstantial	two ways (P_Pov)	two ways (S_Rand) the following natural way (S_Rand)

Table 6.14 Cataphorically lexicalised shell nouns in hyper-Themes in Published and Submissions

There are surprisingly few hyper-News containing shell nouns in the corpus: only five in each sub-corpus. Only factual, linguistic and mental shell nouns are employed in this function. Shell nouns in hyper-News refer either anaphorically or esphorically. An example of a mental shell noun in a hyper-New is given in 6.37.

6.37 The extensive examination of the survey's validity, reliability and generalizability leads to the **inference that there is nothing in the evaluation to suggest that it is either invalid or unreliable.** (S_CEO)

Example 6.37 comes from the transition from the methodology section of the research article to the presentation of results. The shell noun refers esphorically, its lexicalisation construes Qualifier. In this instance the shell noun contributes interpersonal meaning, drawing on the system of ENGAGEMENT. The writer could have chosen not to use a shell noun at all, as in 6.37'. Use of the shell noun results in an extra layer of interpersonal meaning in the clause.

6.37' The extensive examination of the survey's validity, reliability and generalizability indicates/suggests that there is nothing in the evaluation to suggest that it is either invalid or unreliable.

The linguistic and mental shell nouns in hyper-News contribute to the dialogic positioning of the writer's voice. They include the mental shell nouns **picture**, **purpose** and **thesis** in the Published sub corpus, and mental shell nouns **aim**, **concerns** and **inference**, and the linguistic shell noun **conclusion** in the Submissions. A second example, also from the Submissions sub-corpus, comes at the point of transition from the descriptive report of others' research to the descriptive report of the writer's own study in the introduction section of the article.

6.38 Hence, agency problems can be critical in periods of transition from traditional systems of trust to market-based systems of law because they reduce incentives to invest in reputation mechanisms and raise transactions costs. If the system under analysis does not collapse, something must have been promoted to cope with this paradox. (S_Net)

The shell noun **paradox** is factual, and refers anaphorically. In this instance, **paradox** has an evaluative reading. The writer appraises the situation described in lexicalisation as a contradictory. In academia it is expected that research tackles situations that need explanation. In choosing to encapsulate the previous text in the shell noun **paradox**, the writer is evaluating her own study as worthwhile. The factual shell nouns in hyper-News in the corpus contribute to an evaluative reading either through attitudinal inscription such as 'these basic **findings**' (P_CEO) and 6.38 above or through using the shell noun to dialogically position the reader. While the anaphorically referring shell nouns draw on the system of ATTITUDE, esphorically referring factual shell nouns can be used to contract the dialogic position of the reader. This latter realisation of a factual shell noun in a hyper-New is exemplified in 6.39. In this case, the matrix realised by 'the point is' contracts the dialogic space so that the reader is railroaded into agreeing with the writer's point of view.

6.39 Rather, the **point** is that none of the first two generations of innovators of neoclassical price theory such as Jevons, Walras, Marshall, Edgeworth, Bowley and so forth felt compelled to link that particular theory to explicit empirical evidence or to pollute their value theory with stochastic concepts. (P_Econ)

This section has described how shell nouns are used in macro-Themes, hyper-Themes and hyper-News. It has been found that cataphorically lexicalised shell nouns constitute an important resource for realising macro- and hyper-Themes, though esphorically referring shell nouns can do so as well. Hyper-News, by contrast, draw on the textual resources of anaphoric and esphoric reference and they couple with evaluative resources from APPRAISAL. While mental and linguistic shell nouns – anaphoric and esphoric – and esphorically referring factual shell nouns enact the dialogic position of the writer, anaphorically referring factual shell nouns are appraised with selections from ATTITUDE.

6.1.3. Internal conjunction

In both sub-corpora shell nouns are found to function as a resource for internal conjunction. In this function the shell nouns often co-occur with hyper-Themes or the textual resources of superset. The combination of a hyper-Theme containing a shell noun with resources from internal conjunction is given in 6.40. The hyper-Theme is underlined, internal conjunction is boxed, and shell nouns are in bold.

- 6.40 For a variety of **reasons** nonprofits provide an interesting context in which to study the pay of managers in this period. **First**, non profits are an important part of the economy. There are more than 1 million nonprofits in the United States. They employ more than 10 percent of all workers and account for about six percent of gross national product (GDP) (Bowen et al. 1994). **Second**, although there has been increased scrutiny from the IRS over the pay of managers of firms, recent legislation has focused on the pay of top managers of charities in the United States (Taxpayer Bill of Rights 2, 1996). This new law not only requires organizations to carefully document how much they pay top managers, but it also requires the boards of the charities to be able to document how the salaries of their chiefs are determined. If the salaries are found to be higher than expected and higher than those found in similar charities, the heads could be fined and required to return the amount by which they were overpaid. Moreover, the members of the boards of directors of these charities also could be fined if pay levels of the top managers are found to be in question. The problem with this legislation is that there are no strong guidelines for applying these laws. **The third reason** non profits are an interesting context in which to study compensation is that there is a theoretical literature that deals, in part, with how managers in nonprofits may be paid, but there is almost no empirical work in the area. (P_CEO)

In 6.40 the opening clause of the paragraph is a hyper-Theme, and the nominal group with shell noun **reasons** as Head composes the starting point for the paragraph. The

ensuing text constitutes the lexicalisation of the shell noun **reasons**, in which resources from internal conjunction – First, Second – function to extend the argument. Possibly because the second reason is extensive, the third reason, which also draws on the resources of internal conjunction, does so by means of a shell noun rather than a conjunction realised as Numerative. By reiterating ‘reason’ the writer is able to remind the reader of the topic of the paragraph. Moreover, the Numerative ‘third’ implies reference to a superset, which reinforces the idea that the propositions introduced by ‘First’ and ‘Second’ are reasons as well.

Instances of hyper-Themes containing cataphorically lexicalised shell nouns that conflate with resources of internal conjunction are found in both sub-corpora. Similarly, there are instances of shell nouns operating a part of a superset and as a resource for internal conjunction in both Published and Submission. These shell nouns are always lexicalised esphorically as a fact clause or cataphorically, and they function to scaffold the development of the text. Most of these shell nouns are factual: **reason**, **advantage**; but there are two instances of linguistic shell nouns – **explanation** and **conclusion** – being used in this way. Table 6.15 lists the instances of shell nouns whose lexicalisation is realised esphorically or cataphorically and which compose internal conjunction in the corpus.

Shell noun type	Published	Submissions
factual	For a variety of reasons The third reason for at least two reasons several reasons Another reason The first possible reason Another possible reason for many reasons	for two reasons The first advantage The second advantage
linguistic	Another explanation	The final conclusion

Table 6.15 Shell nouns as a resource for internal conjunction in Published and Submissions

In addition to scaffolding the development of the text by means of signalling to the reader what information is coming up, anaphorically referring shell nouns can operate as internal conjunction. In 6.41 the shell noun **advantages** encapsulates several clauses of the immediately prior text and appraises them positively. Thus, its contribution is interpersonally orientated. The lexicalisation of the shell noun is underlined, the shell noun is in bold, and internal conjunction is boxed.

6.41 The historical fundaments of these results deserve a more detailed explanation. It tells that in spite of the risk implied in counting on agents placed at months of distance, the most rewarding returns needed people to stay there for long periods of time. They assured that goods flooding the market when fleets arrived from Portugal, were sold throughout the year, exploiting better opportunities, namely when prices rose long after the fleet had sailed back to Portugal. Moreover, the provisioning of interior regions, passing through a chain of middlemen merchants to supply an enlarging market, depended on credit facilities which could only be afforded by individuals settled for enough time to wait for the return of transactions in the mining areas⁴⁵ **Besides these advantages**, residing agents often overlapped trading activities with administrative tasks ... (S_Net)

Anaphorically referring shell nouns also combine with more general options of both internal and external conjunction. Although external conjunction falls within the ideational metafunction, it seems more appropriate to discuss the few instances found in the corpus here. In these cases, the conjunction could have been realised with text reference instead of an anaphorically referring shell noun. What the shell noun does is to add ideational or interpersonal meaning. An example of internal conjunction is given in 6.42, in which the lexicalisation of the shell noun is underlined, the shell noun is in bold, and internal conjunction is boxed.

6.42 Klaus and Klijn (2005a; 2005b) also showed that under a restricted unemployment aversion condition, the domain of weakly responsive preferences is maximal for the existence of stable matchings. **In view of this result**... (P_Rand)

In 6.42 'this result' refers anaphorically to the clause projected by the verbal Process 'showed'. In doing so, the writer both condenses the information and construes it as within the field of research.

Unlike the esphorically and cataphorically lexicalised shell nouns found to operate in internal conjunction, which are mostly factual, the shell nouns that refer anaphorically within internal and external conjunction options tend to be mental. Moreover, there are more instances of this use in the Submissions sub-corpus than in the Published. The instances are listed in table 6.16, where the shell noun is in bold.

Shell noun type	Published	Submissions
factual	In view of this result	Besides these advantages
mental		with this purpose in mind For that purpose For this purpose As a consequence of his different view of the nature of the society

Table 6.16 Anaphorically referring shell nouns combined with internal and external conjunction in Published and Submissions

6.2. Shell nouns and experiential meaning at the discourse semantics stratum

The aim of the analysis of shell nouns and experiential meaning at the discourse semantics stratum is to elucidate the relation between shell nouns, their lexicalisations, and field. As detailed in Chapter 4, each shell noun and each lexicalisation was coded for the field it construes. When the corpus is treated as a whole, the results show that almost a third of the instances of shell nouns in the corpus are found to be field specific (table 6.17). However, a closer, qualitative analysis reveals that the instances taken to construe a specific field are made up of repeated instances of a few shell nouns whereas the non-field specific shell nouns cover a wider range and there is less repetition.

Feature	Field specific		Non field specific					
	Percent	N	Percent	N	T Stat	Signif.	ChiSqu	Signif.
SHELL-NOUN	N=213		N=474					
factual	34.7%	74	29.3%	139	1.42		2.02	
linguistic	20.2%	43	16.5%	78	1.19		1.41	
mental	32.4%	69	21.3%	101	3.13	+++	9.70	+++
modal	0.5%	1	16.0%	76	6.13	+++	35.77	+++
eventive	0.0%	0	5.7%	27	0.00		12.63	+++
circumstantial	12.2%	26	11.2%	53	0.39		0.15	

Table 6. 17 Distribution of shell noun types as contribution to field in whole corpus

6.2.1. Shell nouns as a resource for construing a specific field

There are no statistically significant differences between the Published and Submissions sub-corpora in the shell nouns classed as contributing to the field of discourse (table 6.18). More than two thirds of the instances pertain to the field of research. The field of research is construed through the mental shell nouns: **objective, aim, hypothesis, thesis, premise, assumption, theory, theorem, and analysis**; factual

shell nouns: **corollary**, **result**, **finding**, **evidence**, **phenomenon**, and **proof**; and linguistic shell nouns: **claim**, **argument**, and **conclusion** in the Published sub-corpus. In addition to these shell nouns, the Submissions sub-corpus also construes the field of research through the mental shell nouns: **lemma** and **intention**; and linguistic: **contribution**.

Feature	Shell nouns in Published		Shell nouns in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
FIELD-SPECIFIC-TYPE	N=87		N=124					
research	69.0%	60	66.9%	83	0.31		0.10	
object-of-study	14.9%	13	23.4%	29	1.51		2.29	
context-of-object-of-study	0.0%	0	0.0%	0	0.00		0.00	
methodology	16.1%	14	9.7%	12	1.40		1.95	
mixed	0.0%	0	0.0%	0	0.00		0.00	

Table 6.18 Distribution of field construed through shell nouns in Published and Submissions

Similar to the field of research, there is little variation between the two sub-corpora in the shell nouns that construe field as methodology. They include mainly circumstantial shell nouns: **approach**, **method**, and **criteria**; as well as the modal shell noun: **specification**. However, these shell nouns – unlike those that construe the field as research which tend to be common across all the 10 texts in the corpus – are found in only three texts in the Published sub-corpus: P_Pov, P_CEO and P_Rand, but only two in the Submissions: S_Pov and S_CEO. The prevalence of these shell nouns which construe the field as methodology in the texts on poverty and CEO pay is perhaps understandable. P_Pov, for example, specifically states, “The main objective of our analysis is to use Bayesian factor analysis for computing and summarizing material deprivation at the census tract level” (P_Pov). Thus, the methodology is an important aspect of the paper.

The field of object of study and the field of the context of the object of study are construed through a wider variety of shell nouns that are restricted to two main pairs of texts, and there is little overlap in the shell nouns found. Table 6.19 lists the shell nouns found to construe field as the object of study by text.

Text	Shell noun type					
	factual	linguistic	Mental	modal	eventive	circumstantial
P_Rand			desire prospect	condition		
S_Rand			strategy			
P_Econ		observation statement assertion	interpretation belief consideration			
S_Econ		argument proposal remark suggestion comment	idea opinion interpretation			
P_CEO		law				

Table 6.19 Shell nouns that construe field as the object of study in Published and Submissions

The linguistic and mental shell nouns in P_Econ and S_Econ are nominalisations of linguistic and cognitive mental Processes. The implicit Sayers and Sensors are the economists who contributed to the discussions surrounding the evolution of Econometrics as a legitimate branch of economics. This is exemplified in 6.43 and 6.44.

6.43 These doubts began with the **observation** that it was humanly impossible to know all of the initial conditions for any moderately complicated Hamiltonian, and therefore it was simpler to treat an aggregate of mass points such as, say, an ideal gas as statistical averages (P_Econ).

6.44 Frisch presented just one candidate, Marschak, since he had previously abandoned this **proposal** in view of a **remark** made by Divisia: according to Divisia (whom, as previously indicated, was not familiar with high mathematics), Marschak would not know a partial derivative (S_Econ)

In the case of the Random matching articles, the Senser of the mental shell nouns is one side of the matching market in the matching problem, i.e. a hospital in P_Rand (6.45), or a worker in S_Rand (6.46).

6.45 Each hospital $h \in H$ has a strict, transitive, and complete preference relation $_h$ over the set of students S and the **prospect** of having its position unfilled (P_Rand).

6.46 Consider Q_w , an alternative **strategy** for w , such that wv and vQ_wv if and only if vQ_wv , for every $v, v \in F \setminus \{w\}$. (S_Rand)

6.2.2 Non-field-specific shell nouns

More than two thirds of the shell nouns in the corpus do not contribute to construing a specific field. Examples of such non-field specific shell nouns are **probability**, **case**, **fact**, or **need**. This section will explore how they interact with their

lexicalisations and field. The analysis will be presented by the field construed by the lexicalisation.

In contrast with the results for shell nouns, the results for the lexicalisations of shell nouns show that they do contribute to making ideational meaning by construing a particular field. By far the majority construe field as the object of study, followed by methodology, research, then the context of the object of study. There are also a number of instances in which the lexicalisation construes a mixed field: research and the object of study; or methodology and the object of study.

Lexicalisations that construe the object of study are found throughout the 10 texts of the corpus and they are dispersed throughout each section of the research article. Many of these lexicalisations are realised as an embedded clause projected by the shell noun. For example, in 6.47 from P_CEO, the object of study is managerial pay in nonprofit organisations, while in 6.48, from S_Econ, the object of study is the debate surrounding the establishment of econometrics as a legitimate field of study within economics.

6.47 The **pattern** that the top officer or director earns more than the top nonofficer, nondirector, or nontrustee as reported in Table 2 is still evident in Table 3. (P_CEO)

6.48 ... and established a constitutional **goal** to create a social science. (S_Econ)

The shell noun enables an activity sequence from the object of study to construe a ranking experiential function in another clause. In 6.47, for example, the **pattern** that the top officer or director earns more than the top nonofficer, nondirector, or nontrustee construes Carrier, and in 6.48, a constitutional **goal** to create a social science construes Goal. In the latter case, it is not possible to omit the shell noun. The shell noun, therefore, functions to reconstrue an activity sequence of the object of study as Thing. This is the temporary conceptual function described by Schmid (2000). As Thing, it is imbued with a certain amount of stability, and it can then be appraised, as in 6.47, or measured, for example. 6.49 illustrates the latter option.

6.49 ... which measures the average **probability** that two agents, having a common neighbour, being themselves connected ... (S_Net)

6.49 comes from S-Net, and the object of study is business networks and trust in the Portuguese-Brazilian gold market in the 18th century. Despite the grammatical infelicities in the projected clause, the shell noun **probability** creates a concept that makes it possible to measure how probable it was for two agents, who had a common neighbour, to be connected to each other. **Probability** is, of course, a grammatical metaphor. Had the writer construed the activity sequence congruently, the modal assessment would have seemed less exact, and more subjective (two agents, who had a common neighbour, were probably/possibly connected).

When these shell nouns projecting a Qualifier that construes field as the object of study are looked at from an interpersonal perspective, it is found that the shell noun more often than not contributes interpersonal meaning. Factual shell nouns, in particular **fact**, can function as a matrix to make the discourse contractive while mental and verbal shell nouns can render it expanding. The contractive and the expanding functions of the shell noun with field as the object of study are exemplified respectively in 6.50, in which the object of study is the change in business network structures among in Liverpool in the 18th century, and 6.51, in which the object of study is the historical development of econometrics as a legitimate branch of economics. In the examples the interpersonal function precedes the element enacting it, the shell noun is in bold and its lexicalisation is underlined.

6.50 Again, [ENGAGEMENT: HETEROGLOSSIC: contract: proclaim: reinforce: pronounce] the **fact** that membership of these institutions was highly instrumental is highlighted by [ENGAGEMENT: HETEROGLOSSIC: contract: proclaim: reinforce: pronounce] the **fact** that between 1770 and 1777, Watts invested in fifteen out of thirty-three slave-trade voyages with the following members of the Unanimous Club: Joseph Brooks junior, William Crosbie, Alexander Nottingham, William Pole, Charles Pole, and Clayton Case. (P_Net)

6.51 First and foremost, the success of quantum mechanics fostered [ENGAGEMENT: HETEROGLOSSIC: expanding: attribute] the **impression** that stochastic explanation was eminently scientific. (P_Econ)

A further example is given in 6.52, again from S_Net, whose object of study is business networks and trust in the Portuguese-Brazilian gold market in the 18th century. The shell noun expands the dialogic space and couples with a Qualifier that construes field as the object of study.

6.52 The analysis on networks gave theoretical support to [ENGAGEMENT: HETEROGLOSSIC: expand: entertain] the **hypothesis** that social capital, while a metaphor of systems of trust, does not fit the Portuguese-Brazilian market. (S_Net)

In addition to functioning as a systemic resource to position the writer with regard to other voices, shell nouns are found to grade the clause construing the object of study in terms of fulfilment. Instances are found of graded completion in both sub-corpora. In 6.53, from P_Econ, the object of study – establishing econometrics as a legitimate branch of economics – is assessed as not fully complete.

6.53 The [GRADUATION: focus: fulfilment: completion; downscaled] **struggle** to establish econometrics has frequently been misrepresented as a battle between the misguided partisans of “measurement without theory” versus the level headed exponents of a judicious and balanced empiricism. (P_Econ)

Projecting a clause that construes field as the object of study as Qualifier to a shell noun also enables the writer to assess the worth of the embedded figure attitudinally. This is clearest with attitudinally inscribed shell nouns like **benefit** (6.54), but it is also apparent when the attitudinal inscription is isolated (6.55). The two examples come from P_CEO, in which the object of study is CEO pay in non-profit organisations. The interpersonal function precedes the element that enacts it, the shell noun is in bold, and its lexicalisation is underlined.

6.54 501 (3) non-profits have the added [ATTITUDE: APPRECIATION: inscribed :+ve] **benefit** that contributions to the organization are deductible to the contributor. (P_CEO)

6.55 The **pattern** that the top officer or director earns more than the top nonofficer, nondirector, or nontrustee as reported in Table 2 is still [ATTITUDE: APPRECIATION: inscribed +ve] evident in Table 3. (P_CEO)

A figure construing field as the object of study may also be encapsulated in a shell noun that inscribes evaluation of the whole as a judgement of behaviour (6.56). The object of study in S_Pov is measuring poverty in Portugal.

6.56 One of them relates to [ATTITUDE: JUDGEMENT: social esteem: +ve] the household's **ability** to manage its general expenses, other than housing and clothing. (S_Pov)

One of the measures of poverty in the paper is whether a household can manage its general expenses, and the shell noun **ability** enacts a judgement on the household's behaviour. Instances of judgements of behaviour that construe field as the

object of study through a Qualifier are found in both sub-corpora. A further example of JUDGEMENT, this time of social sanction: veracity, is given in 6.57.

6.57 ... which measures [ATTITUDE: JUDGEMENT: social sanction: veracity +ve][GRADUATION: FOCUS: fulfilment: actualisation: downscaled] the average **probability** that two agents, having a common neighbour, being themselves connected ... (S_Net)

Although awkwardly expressed, the shell noun **probability** construes the fact that it is possible for two agents, who have a common neighbour, to be connected in the network. The shell noun concomitantly grades the truth of this situation as not fully actualised.

The instances described above illustrate a coupling of experiential and interpersonal meaning. The shell noun makes it possible to construe a figure from the field of the object of study as Qualifier, thus instantiating it as a Thing, a concept, while at the same time appraising it. The choice of shell noun specifically contributes interpersonal meaning by virtue of its function within one of the systems of ATTITUDE, GRADUATION, or ENGAGEMENT. A full description of the contribution of shell nouns to interpersonal meaning is given in section 6.4 of this chapter; however a brief comparison of the evaluative resources in shell nouns that project a figure construing the object of study the brings to light both similarities and differences within the corpus.

The results show the two sub-corpora use coupling of interpersonal resources with shell nouns whose lexicalisation, projected as Qualifier, construes field as the object of study, similarly, but there are differences in the relative shares of the interpersonal selections. Compared to the Published sub-corpus, Submissions makes greater use of selections from ATTITUDE (e.g. difficulty, advantage) and GRADUATION (e.g. propensity, prospect), and less use of selections from ENGAGEMENT (e.g. fact, assumption, observation) (Appendix F.8). The Published sub-corpus shows a higher propensity to construe the object of study using resources that contract the discourse. This is done through the shell noun **fact**, which constitutes all the instances in the Published, and is in fact the only shell noun with a count higher than 4. By contrast, although the shell noun **fact** is used to project a clause construing field as object of study 7 times in the Submissions sub-corpus, the trend is to open rather than close the

dialogic space. Repeated instances of the shell noun **case**, which is found 5 times, to realise a Circumstance (in the **case** that) contribute to this difference. A further difference between the two sub-corpora is that in expanding discourse, the Submissions sub-corpus mostly uses the shell noun to project field as the object of study to entertain the writer's view while the Published sub-corpus mostly uses the shell noun to attribute the proposition to another source. The higher proportion in Submissions of selections from GRADUATION and ATTITUDE, particularly JUDGEMENT, could be accounted for by the high counts of the shell noun **probability**, which is coded dually as JUDGEMENT and GRADUATION: downscaled fulfilment: actualisation; it appears 16 times.

The second most frequent realisation of lexicalisations that construe field as the object of study is extended text that stretches across sentence boundaries. In this case the shell noun refers anaphorically or is lexicalised cataphorically. The following instance shows how the shell noun **results** can be used to announce information spanning several clauses. Both the shell noun and its lexicalisation are field specific: the field of research, realised by the shell noun **results** (boxed), serves to project metaphorically the field of the object of research, which is realised by the cataphoric lexicalisation (underlined). The object of study is CEO pay in nonprofit organisations.

6.58 Columns 4, 5 and 6 examine the compensation of the top nonofficer, nondirector or nontrustee employee, and the results differ somewhat. In the simple cross section, the link between performance and pay is much stronger (0.591), and the result is still significant when 26 NTEE (industry) classification (26 possibilities) while still accounting for individual organization fixed effects. (I have not reported the results in the tables.) Only two of the individual industries show significantly positive returns on the fraction of expenses spent on program services. There is only a very weak link between pay and performance based on this measure. Nonprofits that spend higher fractions of their total expenses tend to pay their top officers, directors and trustees and top nonofficer, nondirector and nontrustee staff more in the cross section, but further analysis suggests that this result is not particularly robust.

Lexicalisations that construe field as the object of study are also realised as clauses or clause complexes. Again in 6.58, in addition to the nominal group 'the **results**' at the beginning of the segment, another shell nominal group 'this **result**' in the final clause complex refers anaphorically to a clause: Nonprofits that spend higher

fractions of their total expenses tend to pay their top officers, directors and trustees and top nonofficer, nondirector and nontrustee staff more in the cross section .

Field as the object of study also appears in lexicalisations realised as fact clauses. In an instance taken from the same text as 6.58 the shell noun **explanation** is non field-specific, but its lexicalisation is not (6.59).

6.59 **Another explanation** may be that it is easier for larger boards to monitor the operations of organizations, and therefore, organizations with larger boards may not require managers with high salaries. (P_CEO)

Another example of a non field-specific shell noun – **trend** – referring to a field-specific lexicalisation is given in 6.60. The object of study in P_Net is the changing structure of the social and business networks in Liverpool in the 19th century.

6.60 Usually, as group sizes increase, the density decreases if the numbers of actor degrees remain unchanged, i.e. the group gets too big for everyone to know each other. However, the network in this study does not follow **this expected trend** precisely. (P_Net)

There are some instances in which it is not possible to clearly classify the realisation of a lexicalisation as clausal or nominal. These instances are classed as ‘other’, and they are all from the Submissions sub-corpus. In each case there is awkwardness of expression although the meaning remains clear. This is illustrated in 6.61, where the shell noun is in bold and its lexicalisation is underlined.

6.61 In any event there is a **consensus** about the effects of social change on trust, causing “depletion” in the overall level of social capital, described as “an inverted U shaped curve between the density of social capital and the level of development”. (S_Net)

In addition to clauses, the object of study can be lexicalised in an equation (6.62). The few instances in which the lexicalisation is an equation all involve circumstantial shell nouns: **criterion**, **condition** or **case**, and they are found in S_Rand, P_Rand and P_Pov. Otherwise, the range of shell nouns whose lexicalisations construe field as the object of study spans all the types of shell nouns – both field specific and non field specific – and is not restricted in any way.

6.62 In fact, the stability of φ with respect to P implies that either $f \text{ Pf } w$ – in which case f declares w unacceptable – or, if $w \text{ Pf } f$, then $! \varphi(f) = q f$ and $w \text{ Pf } w$, for every $w \text{ } \varphi(f)$. In **the latter case**, since ... (S_Rand)

Comparison between the Published and Submissions sub-corpora of the lexicalisations that construe field as the object of study show that the Submissions sub-corpus has a lower proportion of fact clauses, i.e. embedded clauses that construe Token in an intensive identifying relational Process clause. This result attains high statistical significance, and is interesting. As will be shown in section 6.3.1, in intensive identifying relational Process clauses in which the shell noun construes Value, the structure can function as a matrix to open or close the dialogic space, in other words suppress alternative views to the proposition realised in the lexicalisation, or admit them. The lower proportion of these lexicalisations realised as fact clauses in the Submissions suggests that the writers are not engaging with the reader in the same way that the writers of the Published sub-corpus are. This could affect a reviewer's response.

Feature	Lexicalisation construing field as object of study in Published		Lexicalisation construing field as object of study in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
(LEXICALISATION)- TYPE	N=193		N=229					
(shell-nominal-group)	0.5%	1	1.3%	3	0.84		0.70	
(as-clause)	97.9%	189	95.6%	219	1.31		1.72	
(equation)	1.6%	3	0.4%	1	1.18		1.39	
((other))	0.0%	0	2.6%	6	0.00		5.13	++
(AS-CLAUSE-TYPE)	N=187		N=217					
(projected-clause)	33.3%	63	38.4%	84	1.05		1.11	
(fact-clause)	17.5%	33	5.9%	13	3.72	+++	13.47	+++
(clause)	12.2%	23	15.1%	33	0.85		0.72	
(>clause)	34.9%	66	37.4%	82	0.53		0.28	
(dem-ref+clause)	2.1%	4	3.2%	7	0.67		0.45	

Table 6.20 Realisation type of lexicalisations that construe field as object of study in Published and Submissions

Similar to lexicalisations that construe field as the object of study, lexicalisations that construe field as methodology are realised by a variety of clause types (table 6.21). However, the range of shell nouns is limited, and there are repeated instances of the same shell noun. Circumstantial shell nouns are the most frequent in both sub-corpora, followed by mental shell nouns, factual shell nouns, and modal shell nouns. Table 6.22 lists the shell nouns found in this environment in the corpus.

Feature	Lexicalisation construing field as methodology in Published		Lexicalisation construing field as methodology in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
(LEXICALISATION)-TYPE	N=25		N=19					
(shell-nominal-group)	4.0%	1	0.0%	0	0.00		0.78	
(as-clause)	96.0%	24	100.0%	19	0.87		0.78	
(equation)	0.0%	0	0.0%	0	0.00		0.00	
((other))	0.0%	0	0.0%	0	0.00		0.00	
(AS-CLAUSE-TYPE)	N=24		N=18					
(projected-clause)	25.0%	6	21.1%	4	0.30		0.09	
(fact-clause)	29.2%	7	21.1%	4	0.59		0.37	
(clause)	20.8%	5	5.3%	1	1.47		2.14	
(>clause)	25.0%	6	47.4%	9	1.53		2.34	
(dem-ref+clause)	0.0%	0	0.0%	0	0.00		0.00	

Table 6.21 Lexicalisations that construe field as methodology by realisation type in Published and Submissions

Shell noun type	Published		Submissions	
circumstantial	approach	3	approach	2
	limitation	2	procedure	2
	way	2	step	1
	specification	1	criteria	1
	criteria	1		
mental	motivation	3	strategy	4
	assumption	2	inference	1
	rationale	1	perspective	1
factual	advantage	3	feature	1
	feature	1	reason	1
	difficulty	1	case	1
			modification	1
modal	probability	1	probability	1
	need	1	possibility	1
linguistic	suggestion	1		
eventive			improvement	1

Table 6.22 Shell nouns whose lexicalisations construe field as methodology.

In the Published sub-corpus, most of the instances are found in one text, P_Pov, while they are spread more evenly across three texts of the Submissions sub-corpus: S_Pov, S_CEO and S_Net. Looking at these shell nouns from the perspective of the interpersonal metafunction, it is found that several carry infused attitude and they can be paired as positive and negative evaluations of methodology. This function is exemplified in 6.63 and 6.64, where in the former the methodological approach is appraised negatively as a **limitation**, and in the latter positively as an **advantage**. The

nominal group with shell noun as Head is boxed, the shell noun is in bold, and its lexicalisation is underlined.

6.63 A final **limitation** of empirical indices is that only data from the specific area are used. (P_Pov)

6.64 it [the Bayesian approach] confers the distinct **advantage** that uncertainty about factor scores, which in many cases are of direct interest, can be accurately summarized as a natural byproduct of the posterior parameter distribution. (P_Pov)

A mental or linguistic shell noun projects a researcher's view, thus attributing the methodology either to the writer of the paper (6.65) or another academic (6.66).

6.65 Our **suggestion** is to define "deprivation" in terms of percentiles of the latent index distribution. (P_Pov)

6.66 The lack of identifiability caused by the label-switching problem in finite mixture models has been handled by resorting to several strategies. The most usual **strategy** (e.g. Richard and Green, 1997; Bartolucci et al., 2003; Congdon, 2005; and Fruhwirth-Schnatter et al., 2002) consists of imposing artificial identifiability constraints on model parameters. S_Pov

A circumstantial shell noun such as **approach** can be considered to construe the field of methodology, so it is not surprising to find it lexicalised by clauses that construe the same field. Similarly, characteristics of the methodology can be specified by means of factual shell nouns in a cohesive relation. In 6.67, for example, **feature** refers anaphorically to the preceding clause.

6.67 Finally, once computed, the index lacks a measure of uncertainty. For policy making decisions in particular, this last **feature** may be problematic (P_Pov)

Besides the fields of the object of study and methodology, there are a few lexicalisations of shell nouns that construe field as the practice of research. They are exemplified in 6.68 – 6.70.

6.68 It is the **purpose** of this paper to investigate how one might begin to document this thesis. (P_Econ)

6.69 However, one **conclusion** from our analysis is that these results cannot be obtained. (P_Rand)

6.70 There is no **intention** here to add any contribution whatsoever to this interesting debate. (S_Net)

All of these instances imply the presence of a researcher. This is realised by means of the Qualifier to the shell noun in 6.68 and 6.69 (of this paper; from our

analysis) as well as mental shell nouns: **purpose, intention**; and linguistic shell noun **conclusion**, which can be traced back to the writer.

In some cases the lexicalisation of a shell noun construes field as a broader topic than the object of research, i.e. the context of the object of research. The example comes from the paper on trust and business networks in the Portuguese-Brazilian gold trade in the 18th century.

6.71 A widespread opportunistic behaviour, such as free riding, cheating, corruption and rent-seeking reflects the state's **inability** to enforce property rights and contracts or policies. (S_Net)

Instance 6.71 comes from the introduction section of the research article, and it appears in the descriptive report of other studies, i.e. it serves to describe the context in which the study is undertaken. Another instance from the introduction is given in 6.72, where the writer is referring to context in which non-profits operate, more specifically to recently introduced legislation regarding managerial pay in non-profits:

6.72 The **problem** with this legislation is that there are no strong guidelines for applying these laws. (P_CEO)

Lexicalisations that construe field as the context of the object of study also appear in the Conclusions section. The instances come from the Submissions paper on CEO pay in Portugal and the published paper on poverty. In both cases the lexicalisation broadens the scope of the research such that the results of the study are deemed relevant for practitioners in society.

6.73 The general **conclusion** of the present paper is that governance is obviously important in the Portuguese non-profit sector. (S_CEO)

6.74 The main **conclusion** is that adoption of governance principles related to board composition is urgently needed in the Portuguese NPO sector. (S_CEO)

6.75 An important **question** arises as to how researchers and policy makers should use this technology. (P_Pov)

Only one instance does not appear in the introduction or conclusion (6.76). It comes from the article that focuses on poverty in Portugal, and is found in a section of the article that describes the different kinds of non-profit organisations in Portugal. Such contextualisation is necessary for readers of the international community who would not be familiar with the Portuguese context.

6.76 In the present day, their main **function** is to provide social welfare to the elderly. (S_Pov)

Because shell nouns can be lexicalised by clauses, it is also possible for the lexicalisation to construe a mixed field, in which one field metaphorically projects another (Hood, 2010) although a finely tuned identification is not always easy. The most common mixed field in the lexicalisations of shell nouns is one in which the field of research projects the field of the object of study. The majority of these instances occur in either the introduction or conclusions section of each research article. This is the case for both the Published sub-corpus and the Submissions. Example 6.77 comes from the introduction of S_Net. The field of the object of study is the role of travelling in promoting trust in Portuguese-Brazilian gold transactions in the 18th century. The shell noun is in bold, the field of research is boxed, and the field of the object of study is underlined.

6.77 Network analysis arises as an operative **method** to discuss systems of trust. (S_Net)

While 'systems of trust' belong to the field of the object study, 'to discuss' them construes the field of research. Many of the instances of lexicalisations that construe field as research metaphorically projecting the object of study are realised by Processes that construe the field of research, and Participants that construe the field of the object of study, as in 6.77, where 'discuss' is a verbal Process. Other instances include verbal Processes '**ways** to report material deprivation at the area level' (P_Pov); mental Processes 'The **need** to analyze or summarize multivariate data that are spatially aligned' (P_Pov), 'the **opportunity** to gain a deeper insight into the deprivation phenomenon' (S_Pov) or material Processes 'intellectual **efforts** to submit the institutional level of an economic system to quantification' (S_Net), 'the **need** to measure poverty from a multidimensional perspective' (S_Pov). Once again, these instances appear in the introduction and conclusion sections of the articles.

The corpus includes some instances in which the writer refers to previous research. In 6.78, again the shell noun is in bold, the field of research is boxed, and the field as the object of study is underlined. The lexicalisation of the shell noun **reason** is realised by a fact clause that construes Token.

6.78 The third **reason** that nonprofits are an interesting context in which to study compensation is that there is a theoretical literature that deals, in part, with how managers in nonprofits may be paid, but there is almost no empirical work in the area. (P_CEO)

The shell nouns used to encapsulate lexicalisations where the field of research metaphorically projects the field of the object of study are listed in table 6.23. They are varied, and many are single instances. Repeated instances include shell nouns such as **result** or **aim**, which construe field as research.

Shell noun type	Published		Submission	
factual	result	2	result	3
	reason	2	implication	2
	feature	1	aspect	1
	fact	1		
	case	1		
	focus	1		
linguistic	remark	1	remark	1
	comment	1	observation	1
	description	1	discuss	1
mental	view	2	aim	4
	idea	1	purpose	1
	interpretation	1		
	picture	1		
	intention			
modal	need	1	need	1
			opportunity	1
eventive	tendency	1	effort	1
	struggle	1		
circumstantial	way	2	constraint	1
	approach	1	approach	1
	constraint	1	method	1
	method	1		

Table 6.23 Shell nouns whose lexicalisation construes the field of research projecting the field of the object of research in Published and Submissions.

The other main mixed field construed by the lexicalisations of shell nouns is that of methodology projecting the object of study. This is exemplified in 6.79. The field of methodology is boxed, the field of the object of study is underlined, and the shell noun is in bold.

6.79 Traditionally, eq.(1) was estimated by OLS, but Core et al. (1999) and Gosh and Sirmans (2005) found that in the pay performance model, the board composition is endogenous to the system and as such, an OLS model is subject to mis-specification. Therefore, estimated OLS parameters obtained by the traditional approach are not consistent, either because of measurement errors in the Board variables or because the explanatory variables, such as board composition or individual education, are correlated with the unobserved

disturbances in the equation. Alternative approaches to deal with this endogeneity **problem** fall into three categories: (S_CEO)

This instance exemplifies how field specific meanings can be encapsulated in a non-field specific shell noun. The shell noun reinterprets the lack of consistency in the OLS parameters as a **problem**. The Classifier *endogeneity* specifies the problem as belonging to the field of methodology, but the specification does not constitute the lexicalisation of the shell noun. The field construed by the lexicalisation is mixed: methodology projecting the object of study. Shell nouns that encapsulate lexicalisations that construe a mixed field include **problem**, **purpose approach** and **criterion/criteria** in Published, and **hypothesis**, **reasons**, **result** and **feature** in Submissions.

6.79 also exemplifies how shell nouns can be modified by field specific Classifiers. As previously stated, the Classifier does not constitute the lexicalisation of the shell noun. This becomes apparent in an example in which the field specific modification precedes the lexicalisation (6.80) and construes Qualifier. The cataphoric lexicalisation is underlined, the shell noun is in bold and the Qualifier boxed.

6.80 Another reason that an organization may form as a nonprofit comes out of the **idea** of public goods (Hansmann 1980). The fact that one person enjoys the services of a public good does not preclude others from doing the same. (P_CEO)

A public good is a good that, once consumed, does not preclude others from doing the same. By contrast, the idea of public goods is that just because one person enjoys the services of a public good does not mean that others cannot do the same. While 'of public goods' is field specific, it does not construe the shell noun's lexicalisation, which is not a concrete entity, but a semiotic entity. The shell noun here introduces a level of abstraction. A second example, 6.81, illustrates how the Qualifier 'of the paper' locates the anaphorically referring shell noun **aim**, whose lexicalisation is underlined, as within the field of research.

6.81 However, historical records, when providing data on a global network made up of principal-agents relations, bears information to compute cohesion (or lack of it) and to appraise the links of social capital to agency and economic performance, which is the **aim** of this paper. (S_Net)

Shell noun type	Published	Submissions
factual	Distinct <i>institutional</i> features (P_CEO) these <i>institutional</i> features (P_CEO) the case of a manager of a nursing home (P_CEO) the main <i>empirical</i> findings (P_CEO) The problem with <i>local and confidential networking</i> (P_Net) One of the most curious aspects of the <i>rise of neo-classical theory</i> (P_Econ) some brief <i>biographical</i> evidence (P_Econ) the proof of <i>convergence</i> (P_Rand)	the <i>college admissions</i> problem (S_Rand) the <i>converse</i> result (S_Rand) this <i>endogeneity</i> problem (S_CEO)
linguistic	<i>Concluding</i> comments (P_CEO)	<i>concluding</i> remarks (S_Rand)
mental	motivations for <i>model based indices</i> (P_Pov) a number of important <i>structural</i> assumptions (P_Pov) <i>Exploratory</i> analyses (P_Pov) the idea of <i>public goods</i> (P_CEO) the <i>screening</i> hypothesis (P_CEO) The <i>cluster</i> analysis (P_Net) a ' <i>scientific</i> ' theory (P_Econ) The following <i>existence</i> theorem (P_Rand)	<i>Equilibrium</i> analysis (S_Rand) this <i>social</i> theory (S_Net) Our <i>empirical</i> strategy (S_CEO) a <i>managerial</i> policy (S_CEO) a <i>constitutional</i> goal (S_Econ) the aim of <i>this paper</i> (S_Pov)
modal	the <i>posterior</i> probability (P_Pov) positive <i>posterior</i> probability (P_Pov)	a different <i>sufficient</i> condition (S_Rand) the <i>average</i> probability (S_Net) the <i>discretionary</i> power (S_CEO)
eventive		a refinement of <i>the Nash equilibrium concept</i> (S_Rand) <i>intellectual</i> efforts (S_Net) a possible modification in <i>the BCLM</i> (S_Pov)
circumstantial	Limitations of <i>existing measures</i> (P_Pov) A final limitation of <i>empirical indices</i> (P_Pov) The <i>model choice</i> criterion (P_Pov) The <i>model-based</i> method (P_Pov) Hansmann's <i>non-distribution</i> constraint (P_CEO) the <i>financial</i> criteria (P_CEO) the <i>compensation</i> criteria (P_CEO) These <i>selection</i> criteria (P_CEO) the three <i>weak responsiveness</i> conditions (P_Rand)	an <i>operative</i> method (S_Net) a <i>microeconomic</i> approach (S_CEO) the above-mentioned <i>model comparison</i> criteria (S_Pov)

Table 6.24 Field-specific modification to shell nouns in Published and Submissions

Table 6.24 lists the field specific modifiers used with shell nouns that are found in the corpus. The shell noun is in bold and the modifiers in italics. Many of the shell nouns in the list are also field-specific. For example the factual shell nouns: **proof**,

evidence, and **findings**, and mental shell nouns: **hypothesis**, **theory**, and **analysis** all construe field as research; while the circumstantial shell nouns **method** and **approach** pertain to the field of methodology.

This section has analysed the relation between shell nouns, their lexicalisations and field. It has been found that while shell nouns may or may not construe a particular field, their lexicalisations do. The most common field construed by the lexicalisations is that of the field of research, followed by methodology. It has also been found that the lexicalisation can construe a mixed field. This is not surprising given that most of the lexicalisations are realised by clauses. In addition modifiers can also help specify the shell noun as functioning in a specific field, but they do not provide its lexicalisation. The exception to this is when a Qualifier is projected by the shell noun. Finally, it has been suggested that the choice of shell noun can contribute interpersonal meaning that can spread to the lexicalisation. The contribution of shell nouns to interpersonal meaning is explored in the next section.

6.3. Shell nouns and interpersonal meaning at the discourse semantics stratum

Interpersonal meaning in discourse tends to be spread across different features such that the combination of these features lends an evaluative reading to the text (Martin & White, 2005). As a result, the contribution of shell nouns to interpersonal meaning does not act in isolation, but interacts with other elements. Due to the tagging procedures of the interpersonal function at the discourse semantics stratum, in particular that of the GRADUATION system, the analysis here will be mainly qualitative. Only where possible, will it be supplemented with a quantitative analysis.

6.3.1 Shell nouns and dialogic positioning

In Appraisal theory, dialogic positioning is construed through the ENGAGEMENT system. In this section, I shall first describe the results for how the writers use shell nouns to position themselves dialogically with regard to other voices, i.e. text attributed to other researchers, then describe those for how the writers use shell nouns to position their own voice.

The results of the analysis show that shell nouns – in particular mental and linguistic shell nouns – can function to attribute propositions to other voices, but that dialogical positioning is not restricted to these shell nouns. In some cases factual and modal shell nouns are found to do so as well. The following results pertain to discourse that is dialogically expanding.

Shell nouns as a resource for attributing propositions to other voices

Mental shell nouns, such as **assumption** (7.82), and linguistic shell nouns, such as **proposal** (6.83), clearly can be employed to attribute a proposition to another voice by means of projection. The means of attribution by a shell noun is boxed, and the attributed proposition is the lexicalisation, which is underlined.

6.82 Note that as for many classical matching markets with singles the **assumption** that agents preferences are strict is also crucial for matching markets with couples. (P_Rand)

6.83 Since the early days of the 1920s, when Frisch first tried to convince Divisia of his plans for the creation of a new international association, the **proposal** of publishing a journal had been constantly evoked. (S_Econ)

In 6.82 the **assumption** is understood to be a generalised belief held by researchers working in the field of classical matching markets with singles while in 6.83 the **proposal** can be traced back to the economists, like Divisia and Frisch, who were involved in establishing econometrics as a field of study within economics. Similarly, mental and linguistic shell nouns can acknowledge voices other than those of the writer or those sourced elsewhere, and which may have been previously identified, when they refer anaphorically or cataphorically to the attributed proposition(s). In this case, the attributed proposition is the lexicalisation of the shell noun (underlined). Attribution (boxed) can be realised by a shell noun (in bold) with directed anaphoric reference (6.84, 6.85):

6.84 It was he whom Henry Schultz declared was single-handedly responsible for the creation of the new field of the statistical study of demand (Schultz, 1938, p. 63). This **statement** has irked later generations of econometricians, and indeed, is false from a number of points of view: (P_Econ)

6.85 I have argued elsewhere that the so-called Marginalist revolution in the 1870s consisted of engineers directly appropriating the newly developed formalisms

of nineteenth century energy physics, changing the names of the variables, and renaming the result “mathematical economics” (Mirowski, 1984: forthcoming a). Precisely because **this claim** may appear controversial, I shall not attempt a full elaboration here, but merely take **this thesis** as a point of departure. (P_Econ)

anaphoric reference with a possessive Deictic (6.86):

6.86 In a related paper Roth and Vande Vate (1991), strategic considerations are made for the marriage market, focussing on the class of truncation strategies, i.e., strategies that are order-consistent with true preferences, but may regard fewer partners as acceptable. In a one period game in which every partner states a list of preferences and then a matching stable with respect to those preferences is selected at random, it is shown that all stable matchings can be reached as equilibria in truncations. However, certain unstable matchings can also arise in this way. A multi-period extension is then considered to rule out such undesirable outcomes. ... However, the present paper extends **their contribution** in two ways: (S_Rand)

esphoric reference with a fact clause (6.87):

6.87 Some researchers, including Main, O’Reilly and Wade (1993) have tested tournaments (Lazear and Rosen 1981) in large for-profit firms by including the number of vice presidents in a standard CEO pay regression. **The idea** is that, conditional on other factors, CEO pay should be higher in the presence of more vice presidents because more individuals are in competitions for the top spot in the firm. (P_CEO)

cataphoric reference (6.88):

6.88 Another reason that an organization may form as a non-profit comes out of **the idea of public goods** (Hansmann 1980). The fact that one person enjoys the services of a public good does not preclude others from doing the same. (P_CEO)

and in an existential Process clause (6.89):

6.89 In any event, **there is a consensus** about the effects of social change on trust, causing “depletion” in the overall level of social capital, described as “an inverted U shape curve between the density of capital and the level of development”. (S_Net)

As previously mentioned, not only mental and linguistic shell nouns are found to function to attribute propositions. In such cases there is always reference to the

external source in the co-text either by means of a footnote number (6.90), a possessive Deictic (6.91), an integral citation (Swales, 2005) (6.91), or a non-integral citation (Swales, 2005) (6.92). Again attribution is boxed, the shell noun is in bold, and the lexicalisation is underlined.

- 6.90 One of the **reasons** for centralizing this market, and later for reorganizing it ², was that market outcomes did not seem to be stable as indicated by unravelling (pre-NRMP) or a significant reduction in voluntary participation in the NRMP (pre-organized NRMP). (P_Rand)
- 6.91 Second is Hansmann’s “non distribution **constraint**”. Hansman (1980, 1986) notes that while nonprofit organizations are free to make profits, those profits cannot be distributed to those with formal control over the organization. (P_CEO)
- 6.92 There is some **evidence** that the makeup of the board of directors may have some influence on managerial pay or turnover in for profit firms (e.g. Weisbach, 1988; Main, O’Reilly and Wade, 1995; and Hallock, 1997, 1999). (P_CEO)

Table 6.25 shows the shell nouns that are found to attribute propositions in dialogically expanding discourse in the corpus.

Shell noun type	Published	Submissions
Linguistic	appeal, assertion, claim, description, dictum, generalization, manifesto, metaphor, observation, statement,	answer, argument, comment, contribution, discussion, encouragement, insinuation, motto, proposal, remark, suggestion,
Mental	assumption, belief, considerations, consensus, disinclination, expectation, hope, hypothesis, idea, impression, interpretation, premise, resolution, thesis, view	idea, intention, interpretation, consensus, opinion, strategy, theory, theorem, view
Modal	constraint	
Factual	evidence,	reason

Table 6.25 Shell nouns that expand discourse by attribution in Published and Submissions

Shell nouns are found to attribute propositions to other sources in dialogically contracting discourse both the Published and Submissions sub-corpora. The shell nouns include linguistic, factual and circumstantial ones, and the source is often named in the nearby text. In the following examples the endorsing shell noun is boxed, and the proposition endorsed, which is also the lexicalisation, is underlined. Endorsement of other researchers’ propositions can be realised by:

an anaphorically referring shell noun (6.93), (6.94):

6.93 Core et al (1995) analyse the relationship between CEO compensation and board composition, concluding that the board composition explains a significant amount of cross-sectional variation in CEO compensation, after controlling for standard economic determinants of pay. **Their result** reveals that CEOs earn greater compensation when governance structures are less efficient. (S_CEO)

6.94 Klaus and Klijn (2005a, 2005b) also showed that under a restricted unemployment condition, the domain of weakly responsive preferences is maximal for the existence of stable matchings. In view of **this result**, ... (P_Rand)

a cataphorically lexicalised shell noun (6.95):

6.95 **Further corroboration** comes from Brickley et al (1994). External directors enhance shareholder gains in tender offers. (S_CEO)

or an esphorically referring shell noun whose lexicalisation construes Qualifier (6.96):

6.96 This seemingly strange occurrence is, however, a reflection of **the fact** that most Liverpool slave traders engaged in the slave trade as only part of a wider business portfolio (Haggerty, 1980). (P_Net)

Shell noun type	Published	Submissions
factual	fact, result, reason	corroboration, result
linguistic		contribution
circumstantial	constraint	

Table 6.26 Shell nouns endorsing other writers' propositions in dialogically contracting discourse in Published and Submissions

The shell nouns found to endorse other academics' propositions are listed in table 6.26. Interestingly there are far fewer instances of attribution by means of shell nouns that contract the discourse than those that expand it, and there is no overlap between the shell nouns that construe the two dialogic functions. Moreover, there are no instances of mental shell nouns in dialogically contracting discourse. This suggests that the selection of the shell noun can be used to strategically position the writer in relation to information sourced to other voices – either to open up the dialogic space or to close it down.

Finally, in addition to endorsement, there is one instance in which a shell noun helps to signal concession to another voice. Evaluation of the preceding clause is enacted by the nominal group 'a well enough discussed **subject**', which also draws on the resources of GRADUATION (6.97). In the example the lexicalisation is underlined, elements of GRADUATION are in italics, and the shell noun is in bold.

6.97 Indeed personal graces and favours became part of business accounting. This is a well enough discussed **subject** in historical literature. The debate, however, has been overwhelmed by the fiscal consequences of such promiscuity. (S_Net)

The results for the ways in which shell nouns contribute to opening or closing the dialogic space between the writer's own propositions and other voices show that shell nouns can provide a matrix through which the writer shuts down the dialogic space. This can be achieved through esphorically lexicalised shell nouns in intensive relational Process clauses. In this case, the nominal group with shell noun as Head is always presented with presuming reference. Consider 6.98, where the matrix with shell noun is boxed, the shell noun is in bold, and its lexicalisation is underlined.

6.98 The **aim** is just to test the endurance of relations based on trust. (S_Net)

The matrix provided by the shell noun, the **aim**, could have been realised as Process rather than Thing as 6.98' shows.

6.98' We aim just to test the endurance of relations based on trust.

While it may be argued that the function of the shell noun is to render the text more objective sounding, I would argue that it also precludes other voices. This may well be due to the use of presuming reference, which identifies the shell noun as known by the reader, and composing the shell noun matrix as Theme. This coupling reduces the likelihood of being questioned. While different classes of shell nouns are found to function in this way, mental shell nouns make up most of the instances (table 6.27, below). Such instances constitute one of the defining patterns identified by Schmid (2000): *N be* cl. Examples of factual (6.99), linguistic (6.100), modal (6.101), and eventive (6.102) shell nouns functioning to present the writer's own claims while rendering the text contractive in this matrix follow. Again the matrix is boxed, the shell noun is in bold and the lexicalisation, realised as a fact clause, is underlined.

6.99 The inevitable **result** was that the divergence was consequently translated into the choice of Fellows, and the first years of the society were indeed dominated by the definition of criteria both for that selection and for the election process itself, which finally took place for the first time in February 1933. (S_Econ)

6.100 The main **conclusion** is that adoption of governance principles related to board composition is urgently needed in the Portuguese NPO sector. (S_CEO)

6.101 Thus, the next **step** in our analysis is to try to extend our results to more general preferences. (P_Rand)

6.102 The **priority** was to set up the assemblies of econometricians. (S_Econ)

Shell noun type	Published	Submissions
Factual	advantage, irony, point, problem	fact
Linguistic	question	conclusion
Mental	aim, concern, consideration, objective, purpose	aim, concept, function, purpose, solution, strategy
Modal	step	
Eventive		priority

Table 6.27 Esphorically referring shell nouns (lexicalisation realised as fact clause) that function to contract

In addition to the above examples, which were coded as CONTRACT: PROCLAIM: REINFORCE: pronounce, there are three similar instances in which the meaning of the shell noun contributes to justifying the assertion, hence, they are classified as CONTRACT: PROCLAIM: REINFORCE: justify. This is exemplified in 6.103.

6.103 The **motivation** for constraining ρ to have 1's along the diagonal is to maintain the same scaling for $\pm i$ in models with and without spatial correlation (P_Pov)

The instances of justification include only two shell nouns, **motivation** and **rationale**, which are both mental, and they are found in one text only from the Published sub-corpus. The other two instances are:

The primary **motivation** for our proposed methodology is (P_Pov)
The **rationale** for setting D1 is (P_Pov)

Similar to the matrix set in place by the shell nouns in relational Process clauses in the above examples are instances such as 6.104.

6.104 It is also the **case** that although in the simple cross-section non profits with higher levels of government grants pay their heads more, this is not true once organization fixed effects are accounted for. (P_CEO)

In this case the matrix is realised by a split Subject, 'It', composing Theme, and the nominal group with shell noun as Head. Again, only the Published sub-corpus includes instances of heteroglossic contracting matrices of this type. They are:

It is a little known **fact** (P_Econ)

It is a **fact** (P_Econ)

It is no **accident** (P_Econ)

It is not the **intention** of this paper (P_Net)

It is not the **focus** of the analysis (P_Net)

Another way in which some shell nouns can function to close down other voices is by using the shell noun to project a proposition (the lexicalisation) as Qualifier. The embedded clause is thus removed from the stratum of the clause and the shell noun once again acts as matrix. Such uses are typically presented with presuming reference.

6.105 However, the **fact** that Liverpool, and indeed many other British metropolitan areas were so successful in this period, suggests that business networks mostly fulfilled their purpose. (P_Net)

The shell noun **fact** is by far the most commonly selected shell noun to do this, and this result holds for both sub-corpora. However, other shell nouns also contribute to framing a proposition to render the text contractive. They include the factual shell nouns: **evidence**, and **proof**, in the Published sub-corpus; and the mental shell nouns **goal**, **purpose** and **aim** in the Submissions. An example of a mental shell noun used in this way – **goal** – is given in 6.106.

6.106 Secondly, they wanted to emulate physics and established a constitutional **goal** to create a social science ‘to promote studies that aim at a unification of the theoretical-quantitative and the empirical-quantitative approach to economic problems and that are penetrated by constructive and rigorous thinking similar to that which has come to dominate in the natural sciences’. (S_Econ)

In 6.106 the shell noun **goal** can be traced back to the economists who founded Econometrics, and these economists are entities that construe the field of the object of study rather than voices external to the text. The clause complex that is embedded and projected by the shell noun **goal** is removed from the rank of the clause and this reification of the proposition closes the dialogic space.

The shell noun **fact** is also employed in both sub-corpora in the context of justification in a Circumstance realised with ‘due to’.

6.107 This may be due to the **fact** that there are wide differences in the missions of most non-profits. (P_CEO)

The remaining instances of shell nouns that function to contract the dialogic space appear in headings, cataphoric reference or the context of Relators. Only the heading **Proof**, which is found in the Submissions sub-corpus only, is found to function as contractive. There are only two instances of cataphoric reference in which the shell noun is interpreted as a means of closing the dialogic space. In both cases the nominal group with shell noun as Head composes presenting information.

6.108 In Proposition 7 we establish **a partially converse statement**: the set of ordinal Nash equilibria in the game induced by a random stable mechanism includes all the strategy profiles that are simultaneously equilibria in the games induced by the rules that yield the firm-optimal and the worker-optimal stable matchings. (S_Rand)

6.109 In such an altered environment, neo-classical economists had **two choices**: either they could withdraw deep within the obscurity of the Walrasian ramparts or else they could try and strike some sort of accommodation with the new stochastic order. (P_Econ)

While the instance from S_Rand(6.108) is perhaps more marginal because the function of the shell noun **statement** can be interpreted as to carry the modifiers 'partially converse' and the Process 'establish' can be seen as carrying the force of the interpersonal function, in both cases it is possible to leave out the shell noun, which results in a different reading. The shell noun introduces another layer of meaning that frames the lexicalisation. This is most apparent in a rewrite of 6.109 in which the shell noun is omitted.

6.109' In such an altered environment, neo-classical economists either could withdraw deep within the obscurity of the Walrasian ramparts or else they could try and strike some sort of accommodation with the new stochastic order.

Without the frame of the nominal group 'two **choices**', the discourse is expanding. By framing the clause complex that realises the lexicalisation of the shell noun, alternative views to the two options given are excluded.

In the instances in which shell nouns appear in the context of Relators, the shell noun explicitly characterises the causal relation between the two clauses, thus once again, providing justification that precludes other voices. In 6.110 – 6.112 the shell noun acts as a frame that adds an additional layer of meaning.

- 6.110 This model was suggested by a referee on the grounds that conditional spatial correlation $\text{corr}(r_i, r_j) = \frac{1}{k} \sum_{k=1}^k R_i, R_j$ is not dependent on neighbourhood size. (P_Pov)
- 6.111 Pay performance contracts are settled with the aims to align the principal-agent relationship between management and stockholders. (S_CEO)
- 6.112 Since 1994, Eurostat has conducted an annual survey in the EU member states with the purpose of collecting information on the incomes of households and individuals. (S_Pov)

In 6.110, the shell noun matrix ‘on the grounds that’ could have been realised as ‘given that’ or ‘because’, while in 6.111 the Relator with shell noun could simply be left out, and in 6.112 the lexicalisation could have been realised as a hypotactic non-finite perfective clause.

The previous section has detailed how some shell nouns, particularly factual shell nouns, contribute to making the discourse contractive by means of composing headings or by creating particular matrices in which the lexicalisation is realised esphorically as a fact clause or Qualifier. The shell nouns that function to expand the dialogic space operate in similar matrices, but there is no overlap in the shell nouns that are used. Furthermore, there is a difference in the framing matrix: in expanding discourse the nominal group with shell noun as Head is composed with presenting reference, not presuming reference. An example is given in 6.113, where the shell noun matrix is boxed and the lexicalisation is underlined.

- 6.113 One reason why we would expect managers of non-profits with more assets to earn more is that they have responsibility over a much larger number of people and resources. (P_CEO)
- 6.114 A possible modification in the BLCM would be to incorporate the unknown number of latent classes as one additional parameter with an appropriate prior. (S_Pov)

Selection of presenting reference construes the nominal group as but one among others, which are not necessarily cited. In 6.114, for example, the modification is further graded as ‘possible’ while in 6.113 there may well be other reasons why managers of non-profits with more assets earn more. Exceptions to this feature of presenting reference are 6.115 and 6.116, in which the shell nouns are respectively linguistic and mental and reference is presuming.

- 6.115 **Our suggestion** is to define “deprivation” in terms of percentiles of the latent index distribution. (P_Pov)
- 6.116 **The intuition** is that some matching that would be unstable in a singles market is now stable. (P_Rand)

Although these shell nouns are composed with presuming reference, neither **suggestion** nor **intuition** presents the proposition as fully actualised; they are merely entertained by the writer. Table 6.28 lists the shell nouns used in matrices to present the writer’s voice as but one among others.

Shell noun type	Published	Submissions
factual	One advantage of a familiarity with Hermetian matrices One reason One of the most curious aspects of the rise of neoclassical theory	One of the difficulties
linguistic	Our suggestion Another explanation	
mental	The intuition	An alternative strategy
modal	one possibility	
eventive		His only alternative A possible modification
circumstantial	Another way to make the data available	one way to make a sharper prediction of equilibrium outcomes and guarantee stability

Table 6.28 Shell noun matrices in expanding discourse: entertain

Rather surprisingly, there are only five instances in which shell nouns whose lexicalisation is realised as Qualifier are found to entertain the writer’s voice. The instances include mostly mental or linguistic shell nouns, but there is one circumstantial shell noun. All are from the Submissions sub-corpus. In each instance the shell noun matrix is boxed, the shell noun is in bold, and the lexicalisation is underlined.

- 6.117 But this contradicts **the assumption** that under ? , each position of f is filled with a worker f finds better than w. (S_Rand)
- 6.118 **The question** whether or not the social strategies already identified were enough to mitigate agency problems deserves an in-depth analysis. (S_Net)
- 6.119 The analysis gave theoretical support to **the hypothesis** that social capital, while a metaphor of trust, does not fit the Portuguese-Brazilian market. (S_Net)
- 6.120 The extensive examination of the survey’s validity, reliability and generalizability leads to **the inference** that there is nothing in the evaluation to suggest that it is either invalid or unreliable. (S_CEO)

6.121 Diplomas edited in 1672 allowed ship captains to perform as agents, which was just a way of formalizing a situation that had been kept unchanged. (S_Net)

There is also one instance in which the matrix is realised by a split Subject, 'It', and a shell noun (6.122). In this case the modality in the verbal group enacts a subjective assessment of the situation, which rather than contracting the discourse, leaves it open should information to the contrary become available to the writer. It nevertheless indicates that the writer is committed to the view presented (Martin & White, 2005: 133).

6.122 It must be the case that w is matched to a firm under every matching in S(Q_w Q?s). (S_Rand)

Another matrix type for expanding discourse in which the writer entertains propositions is realised by the shell noun **case** in a Circumstance of condition:

6.123 In the case that firms make job offers, the algorithm arrives at the firm optimal stable matching ?F [P], with the property that all firms are in agreement that it is the best stable matching. (S_Rand)

The Circumstance containing the shell noun could have been realised as a hypertactic clause of condition (if firms make job offers) or time (when firms make job offers). The use of the shell noun neutralises this distinction. The shell noun **case** can be similarly employed in Circumstances to refer anaphorically rather than esphorically (6.124).

6.124 However, it may be the that the couple together with the hospital and the old match with the best student's partner does not form a blocking coalition. In this case we prove that the couple together with the hospital and the best student's previous match forms a blocking coalition. (P_Rand)

Other instances of anaphorically referring shell nouns that construe the writer's view as heteroglossically expanding are again mental or linguistic. They include **idea**, **analysis** and **thesis** (mental), and **question** and **answer** (linguistic) in the Published sub-corpus, and **conceptualization**, **hypothesis** (mental), **question** and **remark** (linguistic) in the Submissions. An example from each sub-corpus is given in 6.125 and 6.126. The lexicalisation is underlined and the nominal group with shell noun as Head is boxed.

- 6.125 The question whether or not the social strategies already identified were enough to mitigate agency problems deserves an in-depth analysis. Considering moral hazard is a more common problem in developing countries, being cheated for misinformation or for unauthorized applications of capital must have been a problem more widespread in early modern times than historiography has thought of. A risk assessment in principal-agent relations has not oriented an enquiry yet, although these **questions** dwell upon a set of items transversal to David Hancock and Avner Grief's works. (S_Net)
- 6.126 From a conceptual point of view, government grants may mean one of several different things for the pay of managers of non-profits. First, they could be a sign of organizational quality or managerial ability. That is, in the cross section, we might expect that those organizations with grants or those with larger grants should have managers who are paid more. However, within organizations (where industry type, organizational quality and managerial ability are fixed), more government grants may be a sign of some sort of outside managing of the manager or an increase in funding attracted by the manager. These **ideas** suggest the simple modification of the earlier equation: (P_CEO)

In addition to referring anaphorically, shell nouns can refer cataphorically to propositions entertained by the writer. In 6.127 two shell nouns – **remarks** and **observation** – alert the reader that the ensuing propositions are those of the writer, but they do not shut down other views. In 6.128 the shell noun **hypothesis** clearly identifies the proposition as hypothetical, therefore, entertained.

- 6.127 In what the above results are concerned, a couple of **remarks** is in order. The first **observation** concerns fairness and random matching mechanisms. In opposition to deterministic mechanisms, which are bound to favour one side of the market over the other, we have claimed that random mechanisms promote procedural fairness.
- 6.128 The **hypothesis** is addressed as follows. H1: Cohesion and the distribution of links in the network gives the probability of moral hazard and measures social capital in the system. (S_Net)

The analysis shows that this use of shell nouns is much more common in the Submissions sub-corpus than in the Published. The Submissions sub-corpus includes one instance of a mental shell noun, **hypothesis**, which is used repeatedly in S_Net and S_CEO, the linguistic shell nouns **proposition**, **remark**, **observation**, and **explanation**, and the factual shell noun, **implication**. The Published sub-corpus contains one instance: **recommendation**, which is linguistic.

Finally, the same shell nouns that have been described as functioning to alert the reader that information is entertained are found in Headings. Table 6.29 lists the shell nouns composing Headings that dialogically expand the discourse by attributing the ensuing text to the writer rather than another researcher.

Shell noun type	Published	Submissions
factual	Reasons why pay levels for managers in non profits may be different Example Corollary	Corollary
linguistic	Remark Concluding comments Claim	Definition Proposition,
mental	Theorem Analysis of Rhode Island Census data	Equilibrium analysis Hypotheses Hypothesis 1, 2

Table 6.29 Shell nouns in headings dialogically expanding discourse in Published and Submissions

The preceding analysis has approached the contribution of shell nouns to dialogic positioning in terms of whether information is sourced to another voice or to the writer's, and the analysis was purely descriptive. While linguistic and mental shell nouns predominate, factual shell nouns are also found to construe these functions by means of framing the proposition being put forward. A quantitative analysis of the same data that approaches the contribution of shell nouns to dialogic positioning in terms of whether they open or close the dialogic space shows that there are significant differences in the prevalence of shell noun types between the two dialogic positions when the corpus is treated as a whole (table 6.30).

Feature	Contract		Expand		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
SHELL-NOUN	N=124		N=232					
factual	65.3%	81	9.9%	23	13.41	+++	119.97	+++
linguistic	5.6%	7	46.1%	107	8.54	+++	60.82	+++
mental	21.0%	26	35.3%	82	2.84	+++	7.90	+++
modal	2.4%	3	0.4%	1	1.70	+	2.88	+
eventive	1.6%	2	0.9%	2	0.64		0.41	
circumstantial	4.0%	5	7.3%	17	1.23		1.51	

Table 6.30 Distribution of shell nouns functioning to contract or expand the discourse in whole corpus

The results show that factual shell nouns are statistically more likely to function as dialogic contraction, but linguistic and mental shell nouns are more likely to function as dialogic expansion. These results attain high statistical significance.

Together, these shell noun groups account for more than 90% of the instances. Of the remaining groups, circumstantial shell nouns are slightly more likely to expand the dialogic space while it is the reverse for modal shell nouns. The trend for circumstantial shell nouns to function as dialogic expansion may be due to repeated instances of ‘in this case’. No differences are found for eventive shell nouns, which make up only a very minor proportion of the instances.

Comparison between the Published and Submissions sub-corpora of the contribution of shell nouns to dialogic position in the discourse reveals that there are significant differences between them. An abridged table shows the relevant results (table 6.31); the full results are reported in Appendix G.1.

Feature	Published		Submissions					
	Percent	N	Percent	N	T Stat	Signif.	ChiSqu	Signif.
HETEROGLOSS-TYPE	N=165		N=187					
contract	40.0%	66	28.3%	53	2.32	++	5.32	++
expand	60.0%	99	71.7%	134	2.32	++	5.32	++
EXPAND-TYPE	N=99		N=134					
entertain	54.5%	54	69.4%	93	2.34	++	5.40	++
attribute	45.5%	45	30.6%	41	2.34	++	5.40	++

Table 6.31 Distribution of dialogic position by means of shell nouns in Published and Submissions

While in both sub-corpora there is a tendency for shell nouns to construe the text as heteroglossically expanding rather than contracting, the relative share of each is different in the Published and Submissions. The Published sub-corpus has a statistically higher proportion of dialogically contracting shell nouns and a lower proportion of dialogically expanding shell nouns than the Submissions sub-corpus. A second difference is that in dialogically expanding discourse, the Submissions sub-corpus is more likely to use shell nouns to entertain the writer’s own propositions rather than to attribute them to other voices while in the Published sub-corpus, these functions are more evenly distributed. The significance of these results will be discussed in the section that summarises the results at the end of the chapter.

A further breakdown of the corpus makes it possible to compare the distribution of shell noun types in contracting discourse in the two sub-corpora (table 6.32). The results show that there are no statistical differences in the distribution of

shell noun types between Published and Submissions. Factual shell nouns, which account for more than half of the instances, are the most frequent, followed by mental, linguistic, and circumstantial shell nouns in decreasing order of frequency. It is to be noted that the mental shell nouns appear only in the pattern *The N is cl*, i.e. in identifying relational Process clauses in which the shell noun construes Value and its lexicalisation Token.

Feature	Contract in Published		Contract in Submissions					
	Percent	N	Percent	N	T Stat	Signif.	ChiSqu	Signif.
SHELL-NOUN	N=66		N=56					
factual	67.6%	46	62.5%	35	0.60		0.36	
linguistic	5.9%	4	5.4%	3	0.13		0.02	
mental	16.2%	11	26.8%	15	1.44		2.09	
modal	2.9%	2	1.8%	1	0.41		0.17	
eventive	1.5%	1	1.8%	1	0.14		0.02	
circumstantial	5.9%	4	1.8%	1	1.15		1.33	

Table 6.32 Shell noun types enacting CONTRACT in Published and Submissions

Feature	Expand in Published		Expand in Submissions					
	Percent	N	Percent	N	T Stat	Signif.	ChiSqu	Signif.
SHELL-NOUN	N=99		N=133					
factual	15.2%	15	6.0%	8	2.32	++	5.30	++
linguistic	34.3%	34	54.9%	73	3.16	+++	9.64	+++
mental	42.4%	42	30.1%	40	1.95	+	3.79	+
modal	1.0%	1	0.0%	0	0.00		1.35	
eventive	0.0%	0	1.5%	2	0.00		1.50	
circumstantial	7.1%	7	7.5%	10	0.13		0.02	

Table 6.33 Shell noun types enacting EXPANDING discourse in Published and Submissions

By contrast, the Submissions sub-corpus shows a statistically significant higher proportion of linguistic shell nouns and lower proportion of factual and mental shell nouns to expand the discourse than the Published sub-corpus (table 6.33). These results attain respectively high, medium and low significance.

6.3.2. Shell nouns as a resource for grading propositions

In addition to opening up or closing down the dialogic space, some shell nouns can function to grade propositions. The instances found in the corpus include mental, modal and eventive shell nouns. Modal shell nouns such as **probability**, **possibility** or **opportunity** and even some mental shell nouns such as **prospect** and can grade

propositions with respect to how actualised they are while eventive shell nouns such as **attempt** can grade them with respect to completion. In the following examples the nominal group grading FULFILMENT: actualisation/completion is boxed, the shell noun is in bold, and its lexicalisation is underlined.

In 6.129 ‘being unemployed’ is not fully actualised, it is merely a **prospect**. Similarly, in 6.130 whether the CEO demands compensation in excess of the equilibrium wage or not is not known, but it could happen.

6.129 Each couple $c \in C$ has a strict, transitive and complete preference relation $_c$ over all possible combination of ordered pairs of (different) hospitals and the **prospect** of being unemployed. (P_Rand)

6.130 On the other hand, if the director election process is influenced by the CEO, independence of the board is compromised and CEO monitoring is rendered ineffective, resulting in CEO entrenchment and the **opportunity** for the CEO to demand compensation in excess of the equilibrium wage. (S_CEO)

In both these instances the lexicalisation of the shell noun is realised by a projected Qualifier, but there are also instances in which the lexicalisation is construed as expanding (6.131). For the purposes of the analysis ‘the **risk** of counting on resident agents without the assignment of scouting services’ is counted as projected because the lexicalisation, underlined, fills out the meaning of the shell noun, but in ‘the **risk** implied in category A’, ‘implied in category A’ helps specify the shell noun but does not constitute its lexicalisation.

6.131 Category A gathers the set of individuals that incurred the **risk** of counting on resident agents without the assignment of scouting services. Category B gathers together the individuals who never incurred the **risk** implied in category A. (S_Net)

In the corpus downscaled actualisation is realised by the modal shell nouns **probability, possibility, prospect, opportunity, potential, chance**, and the mental shell noun **prospect**.

While downscaled actualisation is realised mainly by modal shell nouns, downscaled completion is overwhelmingly realised by eventive shell nouns such as **attempt, struggle, task, effort** and **initiative**. There is, however, one instance of a modal shell noun, **challenge**, that grades the proposition as incomplete rather than not

fully actualised. The graded proposition tends to compose esphoric reference, and while most lexicalisations are realised as an embedded projected clause that construes Qualifier (6.132), there is also an instance in which it is realised as a fact clause that construes Token (6.133). Lexicalisations are underlined.

6.132 This paper pushes forward this more nuanced and sophisticated analysis of networks and represents the first serious **attempt** to measure them to assess change over time. (P_Net)

6.133 Mitchell helped found the NBER in 1920, a research organization whose **task** was to implement an empirically grounded economics and encourage the development of a modern 'scientific' economics. (P_Econ)

Table 6.34 shows the shell nouns that are found to function to grade the fulfilment of a proposition as either actualisation or completion. They are listed in descending order of frequency. The modal shell noun **probability** is by far the most common. This result is not surprising given that the corpus is made up of research articles, most of which are empirical studies. The other shell nouns consist of single instances or counts lower than 5. Not only does the Submissions sub-corpus show a higher count of shell nouns with this function, but it also shows a greater range of shell nouns.

Shell noun type	Published	Submissions
mental	prospect (3)	prospect (1)
modal	probability (3), possibility (1)	probability (19), opportunity (4), condition (2), potential (2), risk (2), possibility (1), chance (1), challenge (1)
eventive	attempt (2), struggle (1), task (1)	effort (5), attempt (2), initiative (1)

Table 6.34 Instances of shell nouns that grade FULFILMENT of a proposition in Published and Submissions

When the instances of shell nouns to grade the fulfilment of propositions in the two sub-corpora – Published and Submissions – are compared, the results show that the Submissions sub-corpus uses proportionately more modal shell nouns and comparatively fewer mental shell nouns to grade fulfilment than the Published. These results attain high statistical significance (table 6.35). This difference can be explained by the high counts of **probability** in the Submissions sub-corpus.

Feature	Published		Submissions					
	Percent	N	Percent	N	T Stat	Signif.	ChiSqu	Signif.
SHELL-NOUN	N=11		N=41					
factual	0.0%	0	0.0%	0	0.00		0.00	
linguistic	0.0%	0	0.0%	0	0.00		0.00	
mental	27.3%	3	2.4%	1	2.91	+++	7.53	+++
modal	36.4%	4	78.0%	32	2.81	+++	7.07	+++
eventive	36.4%	4	19.5%	8	1.17		1.39	
circumstantial	0.0%	0	0.0%	0	0.00		0.00	

Table 6.35 Distribution of shell nouns contributing to GRADUATION: FULFILMENT in Published and Submissions

There is one instance in the corpus in which the shell noun contributes directly to interpersonal meaning by grading the scope of the process in the lexicalisation in terms of quantity rather than focus. The shell noun **tendency** is found to do this. In the following example, **tendency** grades how frequently the assumption is made.

6.134 There is a **tendency** to assume that just because actors are linked, they must form a cohesive social network. (P_Net)

The clause complex could have been realised without the shell noun, reconstruing the embedded mental Process *assume* as Existent by drawing on the resources of grammatical metaphor. This option is exemplified in 6.134', but the meaning is slightly different. Interestingly there is still a shell noun – **assumption** – in the re-construal; its lexicalisation construes Qualifier. Thus the dialogic position is still heteroglossic, but there is no notion of how widespread or common the assumption is. The shell noun **tendency**, therefore, can be seen as a means of contributing an extra layer of meaning.

6.134' There is an **assumption** that just because actors are linked, they must form a cohesive social network.

6.3.3. Shell nouns as a resource for Attitude

The final way in which shell nouns contribute to making interpersonal meaning at the discourse semantics stratum is through the system of ATTITUDE. Some shell nouns are not attitudinally neutral. As a result, they can launch a positive or negative assessment that spreads to the co-text, in particular to the lexicalisation of the shell noun. The prosodic flow may be forwards or backwards. Example 6.135 illustrates how the shell noun **limitation** leads the reader to see that using only data from a specific

area is a negative feature of empirical indices. By contrast, the shell noun **advantage** in 6.136 establishes a positive prosody that spreads across the several clauses of the lexicalisation (underlined).

6.135 A final **limitation** of empirical indices is that only data from the specific area are used. P_Pov

6.136 Under certain distributional assumptions, this approach confers several **advantages**: first the model-based deprivation index retains its simple structure as a weighted average of census variables, but the weights are functions of model parameters and therefore are informed by data; second, different sample sizes across areas are incorporated naturally and reflected in posterior variability of the indices; and third, deprivation indices are summarized not as single numbers, but rather in terms of posterior distributions that reflect important uncertainties about deprivation status. (P_Pov)

These attitudinally inscribed shell nouns construe Head in nominal groups in a range of lexicogrammatical environments. In addition to the esphoric and cataphoric uses exemplified respectively in 6.35 and 6.136 above, they include: Value in relational Process clauses whose lexicalisation, realised as a fact clause, construes Token (6.137):

6.137 The **irony** was that prior to 1930 it was mainly (but not entirely) economists who distanced themselves from the neoclassical research program and were openly sceptical about neoclassical price theory who pioneered the explicit connection between a stochastically informed empiricism and some sort of economic theory. (P_Econ)

Attribute whose lexicalisation precedes the shell noun (6.138):

6.138 being cheated for misinformation or unauthorized applications of capital must have been a **problem** more widespread in early modern times than historiography has thought of. (S_Net)

nominal groups that refer anaphorically across clauses (6.139):

6.139 conditional specifications in particular imply marginal variance structures that are not compatible with the usual set of constraints in factor analysis, but the **difficulty** can be resolved by rescaling. (P_Pov)

and esphorically referring nominal groups where the lexicalisation construes Qualifier (6.140):

6.140 501(3) nonprofits have the added **benefit** that contributions to the organization are deductible to the contributor. (P_CEO)

These attitudinally charged shell nouns can be employed strategically to align the reader with the view put forward by the writer, tarnishing other researchers' studies and burnishing the writer's own (Hao & Humprey, 2012). For example, in 6.135 and 6.136 above, both of which are from P_Pov, the writer tarnishes studies that use empirical indices by means of the shell noun **limitation** and burnishes his own study by appraising its methodology positively with the shell noun **advantage**.

The shell nouns in the above examples are all appraising the semiotic entities of their lexicalisations, and thus have been coded as APPRECIATION. Factual, mental, eventive and circumstantial shell nouns are found to construe this function. They are shown in Table 6.36. Several of them, such as **advantage** and **problem**, exhibit multiple instances and are found across a range of texts.

Shell noun type	Published	Submissions
factual	advantage (5), benefit (1), difficulty (1), irony (1), problem (5)	advantage (5), benefit (1), difficulty (1), miracle (1), paradox (2), problem (3), importance (1)
linguistic		
mental		solution (6)
modal		
eventive	improvement (1), problem (1), success (1)	refinement (1), success (1)
circumstantial	limitation (2), shortcoming (2)	

Table 6.36 Attitudinally inscribed shell nouns enacting APPRECIATION in Published and Submissions

Inscribed positive or negative attitude does not need to be infused in the shell noun for it to establish prosodic value that spreads to the lexicalisation. It may also be realised as Attribute to a shell noun in a relational Process clause, where the shell noun refers anaphorically. In 6.141 from P_CEO, the writer is reporting his own results. By appraising them as 'consistent with tournament theory', he validates them, thus burnishing the outcomes of his study.

6.141 This **result** (ever increasing pay as one moves up the organizational hierarchy) is consistent with tournament theory (P_CEO)

Shell nouns with isolated inscribed appreciation can also function to disalign from particular views. In 6.126, for example, the writer refers to an assumption that emerges from the literature review, appraises it negatively, and thus provides a

starting point for his own study. The shell noun **assumption** refers esphorically to its Qualifier.

6.142 As mentioned before, the **assumption** that couple's preferences are weakly responsive is quite restrictive. (P_Rand)

And in 6.143 the inscription is realised as Epithet to the shell noun. Again the writer evaluates the findings from another academic's research negatively.

6.143 In the context of deterministic mechanisms, Roth (1985) shows that by suitably falsifying their preferences, agents can induce any individually rational matching with respect to the true preferences. Unfortunately, this is not a very illuminating result. (S_Rand)

Example 6.143 constitutes one of Schmid's (2000) patterns for shell nouns, *This be N*, and when a shell noun construes Attribute to a Carrier realised by text reference, it is always interpersonally oriented. Attitudinal assessment can be isolated, as in 6.143 or infused in the shell noun (6.144).

6.144 By the end of the first decade of the Society's existence, forty-two fellows represented the Olympus of the 'econometric people'. This was, in any case, an immense success. (S_Econ)

In 6.144 **success** is positively charged and it is further graded in force by the Epithet *immense*. Other instances of inscribed value in shell nouns construing Attributes in this lexicogrammatical pattern include 'a substantial **improvement**' (S_Pov) and 'an unfair **insinuation**' (S_Econ). The attitudinal assessment does not need to be inscribed or infused in the shell noun; it can be invoked. Examples 6.145 and 6.146 illustrate how options from other interpersonal discourse semantic systems invoke an attitudinal reading of the shell noun. In the examples the lexicalisation is underlined, the shell noun is in bold, the relevant resources of GRADUATION are in italics and of ENGAGEMENT boxed.

6.145 Haavelmo's proposal was to redefine the meaning of "population" and "sample" in economics: the population could be conceptualized as all possible economic decisions (or, what is equivalent, all virtual "economies"), whereas the sample would be all actual observed realizations (Haavelmo, 1994, pp 51-53). This is, of course, a **metaphor**: there is no *ghostly* distribution of 1988 GNP growth rates from which the *actual* experience is drawn. (P_Econ)

6.146 If Moore had a precursor, it was more nearly Cournot: both searched for phenomenological regularities cast in the form of a mathematical function; for Moore these were to be linked in a chain to develop an *explicitly* macroeconomic theory of business cycles. In Moore's opinion, this was a '*scientific*' theory. (P_Econ)

In 6.145 selections from ENGAGEMENT – 'of course' and 'no' – construe a dialogically contractive position that flows backwards such that **metaphor** is imbued with a negative value for complexity. In 6.146, while the Epithet *scientific* contributes ideational meaning, the scare quotes around it adjust its boundaries in terms of downscaled GRADUATION: FOCUS, with the result that the theory takes on a negative reading; it is not, in fact, a very scientific theory. There are two other instances of attitudinal value in shell nouns in this lexicogrammatical pattern (**encouragement** and **commitment**), but they fall under the system of JUDGEMENT rather than APPRECIATION and will be dealt with on page 273.

The corpus contains instances of both positive and negative assessments of results from research that serve to position the writer and reader with regard to the value of the research. Example 6.146 illustrates how the writer uses a positive assessment of a mental shell noun to validate the relevance of his study. The instance comes from the descriptive report stage of the introduction. The attitudinal resources are boxed, the shell noun is in bold, and its lexicalisation is underlined.

6.146 A major **consideration** in policy and research is whether model-based indices lead to increases in validity (P_Pov)

As in example 6.146, most of the instances of inscribed APPRECIATION that are realised as Epithets within the nominal group with shell noun as Head evaluate the semiotic entity with regard to its importance: e.g. *key*, *main*, *important*, or *major*. This is similar in both sub-corpora. Table 6.37 shows the instances of inscribed ATTITUDE realised as Epithet to a shell noun that are found in the corpus. The Epithets are in italics.

Shell noun type	Published	Submissions
factual	the <i>distinct</i> advantage (P_Pov) <i>Distinct</i> institutional features (P_CEO) at least two <i>key</i> features (P_Pov) the <i>main</i> empirical findings (P_Pov) these <i>basic</i> findings (P_Pov) a <i>serious</i> dilemma (P_Econ) One of the most <i>curious</i> aspects (P_Econ) our <i>main</i> result (P_Rand) The <i>main</i> problem (P_Rand) a <i>simple</i> example (P_Rand) An <i>important</i> similarity (P_Rand)	Three <i>main</i> aspects (S_Net) the <i>predominant</i> implication (S_CEO) some of the <i>main</i> findings (S_Pov)
linguistic	An <i>important</i> question (P_Pov) some <i>bad</i> news (P_Econ) the further <i>bad</i> tidings (P_Econ) his <i>famous</i> manifesto (P_Econ)	The <i>main</i> conclusion (S_CEO) an <i>unfair</i> insinuation (S_Econ) <i>simple</i> proposals (S_Econ) an apparently <i>naive</i> suggestion (S_Econ)
mental	the <i>primary</i> motivation (P_Pov) a number of <i>important</i> structural assumptions (P_Pov) The <i>main</i> objective (P_Pov) A <i>major</i> consideration (P_Pov) its <i>primary</i> purpose (P_CEO) The <i>ideal</i> resolution (P_Econ) the <i>prejudiced</i> point of view (P_Econ) The <i>biased</i> point of view (P_Econ) the <i>mistaken</i> premise (P_Econ) their <i>prime</i> objective (P_Econ)	one of the <i>clearest</i> lessons (S_Rand) our <i>main</i> goals_Rand a <i>complex</i> solution (S_Net) a <i>rational</i> strategy (S_Net) The <i>fundamental</i> concerns (S_CEO) their <i>main</i> function (S_CEO) the <i>central</i> aim (S_CEO) the <i>dominant</i> concept (S_Econ)
modal		an extremely <i>strong</i> condition (S_Rand) the <i>decisive</i> steps (S_Econ)
eventive	the first <i>serious</i> attempt (P_Net) an <i>unpleasant</i> choice (P_Econ)	
circumstantial	An <i>effective</i> way (P_Pov) This seemingly <i>strange</i> occurrence (P_Net) the radically <i>wrong</i> method (P_Econ)	in the following <i>natural</i> way (S_Rand) an <i>operative</i> method (S_Net)

Table 6.37 Inscribed ATTITUDE in shell nouns realised by Epithets in Published and Submissions

An evaluative reading of the shell noun may also be invoked by means of inscription or invocation in the lexicalisation. In 6.147 the inscribed assessment of net ending assets as ‘an appropriate measure’ invokes a positive appreciation of the whole nominal group. Moreover, the Relator ‘while’ alerts the reader to the evaluative reading such that the previous shell nominal group, ‘the **idea** that the relationship between firm size and managerial pay should be the same across all different nonprofit industries’ is retrospectively read as negative in value. In the example shell nouns are in bold, their lexicalisations are underlined and the element that triggers the evaluative reading is boxed.

6.147 Table 5 relaxes the **idea** that the relationship between firm size and managerial pay should be the same across all different nonprofit industries while retaining the idea that net ending assets are an [+ APPRECIATION: valuation] appropriate measure of size for each organization. (P_CEO)

In the main, because shell nouns are semiotic entities, inscribed and invoked values draw on the system of APPRECIATION. While most instances involve APPRECIATION as valuation, there are also instances of complexity and reaction. 6.148 illustrates APPRECIATION as valuation.

6.148 In the context of deterministic mechanisms, Roth (1985) shows that by suitably falsifying their preferences, agents can induce any individually rational matching with respect to the true preferences. Unfortunately, this is not a very illuminating result. (S_Rand)

By evaluating Roth's (1985) result as 'not very illuminating', the writer is casting a negative assessment of the value of the result for research. The lexicalisation of the shell noun, underlined, spells out what the result was. The following example (6.149) may not at first glance appear to be evaluative at all, but, in fact, it illustrates APPRECIATION as complexity, but also, in one sense, as valuation.

6.149 Hence, agency problems can be critical in periods of transition from traditional systems of trust to market-based systems of law because they reduce incentives to invest in reputation mechanisms and raise transactions costs. If the system under analysis does not collapse, something must have been promoted to cope with this paradox.

Example 6.149 is taken from the introductory section of the research article where the writer is about to move from a descriptive report of other research to a descriptive report of her own. The shell noun **paradox** constitutes the boundary between the generic stages. The aim of the writer's study is to explain why the situation characterised as 'this **paradox**' – which realises the lexicalisation of the shell noun and is underlined in the example – did not cause the economic system to collapse. APPRECIATION as complexity answers the question 'does it hang together?' while APPRECIATION as valuation answers the question 'was it worthwhile?' (Martin & White, 2005: 56). The shell noun **paradox** can be paraphrased as 'an illogical situation', implying complexity, and in academic research, complicated situations are valued as appropriate subjects for research. Thus, the shell noun **paradox** contributes to imbuing

the economic situation to be analysed with positive value, and enhances the value of the writer's own study.

The results show that there are much fewer instances in which the shell noun inscribes or invokes APPRECIATION as reaction, which answers the question 'did I like it?'. 6.150 is one of them. It comes from one the texts that focus on the emergence of econometrics as a legitimate branch of economics.

6.150 If this deficiency with respect to a legitimate dynamics were not distressing, the neo-classical research programme was further buffeted by **some bad news** from the physicists' camp. Just as the economists had come to pay homage to the physicists' belief in determinism, the physicists themselves were contriving to distance themselves from it, and in some cases, even expressing doubts concerning its validity (P_Econ)

The evaluation of the news as 'bad' clearly describes the writer's assessment of how the economists would have felt about the change in the physicists' beliefs. Instances of APPRECIATION as reaction are found in both sub-corpora.

When the distributions of shell nouns used in the two sub-corpora to enact APPRECIATION are compared, it is found that the Submissions sub-corpus has a higher proportion of modal shell nouns (table 6.38). This difference is found to have high statistical significance in the Chi square tests; however, the counts are small, so not too much should be read into it. Notwithstanding, the Submissions sub-corpus shows a much higher share of positive assessments than the Published sub-corpus: 77% vs 57% respectively. This finding attains high statistical significance, and the full results are reported in Appendix G.2. The Submission sub-corpus shows a higher proportion of positive attitude than the Published.

Feature	Published		Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
SHELL-NOUN	N=96		N=90					
factual	40.6%	39	36.7%	33	0.55		0.31	
linguistic	12.5%	12	7.8%	7	1.06		1.13	
mental	24.0%	23	34.4%	31	1.58		2.48	
modal	0.0%	0	6.7%	6	0.00		6.61	+++
eventive	5.2%	5	3.3%	3	0.63		0.40	
circumstantial	18.3%	17	9.3%	8	1.74	+	3.00	+

Table 6.38 Distribution of shell nouns to enact APPRECIATION, inscribed and infused, in Published and Submissions

As previously suggested, shell nouns inscribed with value not only evaluate entities; they can also function to invoke tokens of judgements of behaviour. In the corpus, appreciation of a circumstantial shell noun **occurrence** is found to invoke a judgement of SOCIAL ESTEEM: normality. In 6.151, which comes from the text on the evolution of business networks in 19th century Liverpool, the writer assesses the fact that William Roscoe was elected as MP as a ‘strange **occurrence**’ (-ve APPRECIATION: complexity). At the same time this assessment functions to indirectly evaluate the behaviour of all those involved in his election.

6.151 It is certainly difficult to explain William Roscoe’s election as MP in 1806. True, Roscoe campaigned on an anti-East India Company ticket (Wilson, 2008), but it must have been well known that he supported abolition. He was also nominated by his business partner in his bank, Thomas Leyland, and seconded by Thomas Earle, both prominent slave traders. This seemingly strange **occurrence** is, however, (P_Net)

A second instance of JUDGEMENT as SOCIAL ESTEEM: normality is inscribed rather than invoked. In 6.152, from the article on business networks in 18th century Portuguese-Brazilian trade, the modal shell noun **propensity** assesses how normal it would be to trust and cooperate.

6.152 One may ask what network analysis can tell about cohesion as metaphor of social capital or **propensity to trust and cooperate** in such contexts of market growth. (S_Net)

In contrast to the few instances of JUDGEMENT as SOCIAL ESTEEM: normality in the corpus, there are several instances of JUDGEMENT as SOCIAL ESTEEM: capacity in both the Published and Submissions. These judgements are overwhelmingly realised by modal shell nouns, but there is one instance of the eventive shell noun **difficulty**. The lexicalisation of the shell noun in these judgements is always realised esphorically as Qualifier to the shell noun (6.153 - 6.155). With one exception (6.153), in which it is the behaviour of the writers or readers that is being assessed, the behaviour that is being evaluated is that of a person or group construing the field of study. In 6.154 and 6.155 the object of study is business networks in 19th century Liverpool and 18th century Portugal-Brazil, respectively, while in 6.156 the object of study is CEO pay.

- 6.153 Moreover, these dimensions reveal the existence of intermediate deprivation groups, which allows us the **opportunity** to gain a deeper insight into the deprivation phenomenon. (S_Pov)
- 6.154 In particular, business networks are intended to produce various outcomes such as the rapid transfer and synthesis of knowledge or understanding, the conferment of status and legitimacy, lower information and transaction costs, and the **ability** to adapt to environmental changes (P_Net)
- 6.155 Hence, the most rewarding transactions were increasingly distant from Portugal, which was a factor for Portuguese merchant's **difficulties** in accessing to and assessing information. (S_Net)
- 6.156 The higher the number of stockholders, the more dispersed is the capital and the higher the **discretionary power** of the CEO in establishing the board. (S_CEO)

The shell noun makes it possible to explicate the behaviour being assessed in its lexicalisation, and the person involved can be either explicitly present or implied. For example, in 6.156 it is clearly the behaviour of the CEO that is being assessed, and in 6.155, it is that of the Portuguese merchants. In 6.154, however, the reader must infer that it is ultimately the capacity of the businessmen in the network. While many of the shell nouns that inscribe a judgement of capacity are nominalisations of modality (e.g. **ability, impossibility, inability**), there are also instances of lexical selections (e.g. **difficulty, opportunity, potential, power, capacity**). Negatively inscribed shell nouns such as **inability, impossibility** and **difficulty** clearly enact a negative assessment, but they may similarly be enacted by means of the resources of GRADUATION, such as in 'a *reduced capacity*'. Tables 6.39 and 6.40 list the instances of shell nouns functioning to enact JUDGEMENT: SOCIAL ESTEEM: capacity that are found in the corpus.

Shell noun type	Published
modal	the perceived ability of familial, ethnic and religious networks to reduce moral hazard (P_Net) the ability to adapt to environmental changes (P_Net) a group's ability to mobilise in the face of external pressure (P_Net) the inability to track every single individual particle (P_Econ)

Table 6.39 Shell nouns enacting JUDGEMENT: SOCIAL ESTEEM: capacity in Published

Shell noun type	Submissions
modal	<p>the State's inability to enforce property rights and contracts or policies (S_Net)</p> <p>the State's actual ability to control a large territory (S_Net)</p> <p>better opportunities to cheat (S_Net)</p> <p>higher potentials of developing trustworthy relationships (S_Net)</p> <p>the low potential of the network to foster norms and to make the system work as a controlling mechanism over defrauding agents (S_Net)</p> <p>the opportunity to demand compensation in excess of the market equilibrium wage (S_CEO)</p> <p>the opportunity of the CEO to demand compensation in excess of the equilibrium wage rate (S_CEO)</p> <p>the CEO's opportunity to increase his earnings (S_CEO)</p> <p>the discretionary power of the CEO in establishing the board (S_CEO)</p> <p>the ability to distinguish between a stock and a flow (S_Econ)</p> <p>the household's ability to manage its general expenses (S_Pov)</p> <p>the household's ability to manage its due responsibilities (S_Pov)</p> <p>a reduced capacity to possess an adequate heating system and to manage general expenses (S_Pov)</p> <p>the impossibility to spend a week away on holiday, to go out in their free time or to invite friends or family for a drink or a meal at least once a month (S_Pov)</p> <p>the opportunity to gain a deeper insight into the deprivation phenomenon (S_Pov)</p>
eventive	<p>Portuguese merchants' difficulties in accessing to and assessing information (S_Net)</p>

Table 6.40 Shell nouns enacting JUDGEMENT: SOCIAL ESTEEM: capacity in Submissions

The corpus contains two instances in which linguistic shell nouns function to invoke JUDGEMENT: SOCIAL ESTEEM: tenacity. The shell nouns **commitment** and **encouragement** (1.57) make it possible to assess how willing a person is or how set their behaviour is. The instance is taken from the text on the debates surrounding the establishment of econometrics as a branch of economics. The lexicalisation of **commitment** and **encouragement** precedes the shell nouns, and the shell nouns interpret how willing Divisia, an economist, was to support the emerging econometrics. Note that the lexicalisation, underlined, is the same for both shell nouns and it is realised by a clause complex plus text reference. As previously argued, in this pattern the shell noun is always interpersonally charged.

6.157 Although Divisia did not feel at ease with the more advanced mathematical methods, he nonetheless believed that they represented the way forward: 'Mathematical economics has very few supporters in France; myself, I don't know much about it; nevertheless, I am to be counted among those who consider that economic phenomena must be studied by methods as precise as those used in the other more advanced sciences' (ibid). This was hardly **an encouragement** and even less **a commitment**, but Frisch only wanted not to be opposed. (S_Econ)

In some cases shell nouns are found to invoke JUDGEMENT: SOCIAL SANCTION: propriety in the corpus. The most commonly used shell noun to do this is **need**, which is modal; however the linguistic shell noun **invitation**, mental shell noun **policy**, and modal shell nouns **requirement** and **criteria** are also found to construe judgements of propriety. Once again, the lexicalisation tends to be realised as Qualifier (6.158), but in one case (6.159), the shell noun is lexicalised cataphorically and spans five clauses.

6.158 The **need** to analyze or summarize multivariate data that are spatially aligned suggests the utility of factor analytic models that can incorporate spatial variation. (P_Pov)

6.159 Fisher formulated the **requirements** for the choice of fellow, as recapitulated by Frisch:

1. The candidate must be an economist acquainted with economic theory.
2. He must have a mathematical foundation.
3. He must have some knowledge of statistics.
4. He must have done some original work.
5. Some of this original work must have been in economic theory.

In these instances the behaviour being evaluated is realised as the lexicalisation of the shell noun. For example in 6.158 analysing or summarising multivariate data that are spatially aligned is perceived as a good thing to do. The shell noun makes it possible to enact this evaluation in an element at clause level.

Table 6.41 lists the instances of judgements of propriety found in the corpus.

Shell noun	Published
modal	the need to analyze or summarize multivariate data that are spatially aligned (P_Pov)
	Submissions
modal	the need to compare the probability distributions over matchings generated by random rules when preferences are ordinal (S_Rand) the urgent need for a code of governance practice to be introduced in this sector (S_CEO) the need to measure poverty from a multi-dimensional perspective (S_Pov) the need to take full advantage of inherent uncertainties (S_Pov) the requirements for the choice of fellow (S_Econ)
linguistic	an invitation to proceed (S_Econ)

Table 6.41 Shell nouns enacting JUDGEMENT: SOCIAL SANCTION: propriety in Published and Submissions

The remaining uses of shell nouns to enact an assessment of judgement involve judgements of veracity. As discussed in Chapter 3, judgements of veracity are related to modalisations of probability. Thus, nominalisations such as **possibility** and

probability can be key ways of enacting judgements of veracity, particularly in empirical studies. The shell noun **probability** makes up the greatest share of the instances of the judgements of veracity, but there are also instances of the eventive shell nouns **alternative** and **choice**, and modal shell noun **chance** functioning as a resource for expressing such a judgement. Examples 6.160 – 6.162 illustrate shell nouns functioning to enact JUDGEMENT: SOCIAL SANCTION: veracity. The shell noun is in bold and its lexicalisation is underlined.

6.160 **Chances** of an agent being blacklisted for opportunistic behaviour were scanty, not only because there were no significant links connecting principals who shared agents, but also because the wide number of actors involved increased the **probability** of establishing relations with other principals. (S_Net)

6.161 If $\varphi(w) = w$, no firm is willing to hire w , so that w has no profitable deviation: his only **alternative** is to remain unmatched. (S_Rand)

6.162 Hence, w faces the **choice** of holding $\varphi(w)$ or being unmatched. (S_Rand)

The lexicalisations of the shell nouns construing judgements of veracity are always esphoric, and with one exception (6.161), realised as Qualifier. The lexicalisation also construes the behaviour being judged, and without exception, this behaviour pertains to the field of the object of study. Thus, in 6.161 and 6.162, from the text on matching workers to firms using random matching mechanisms, it is the worker's behaviour that is evaluated, and in 6.160, from the text on business networks in the 18th century Portuguese-Brazilian trade, it is the truth of the behaviour of the principals in the network.

Construing the judgement by means of a shell noun enables it to be graded with regard to its veracity. For example in 6.160, the evaluation is scaled down; the Attribute *scanty* quantifies the **chances** that a principal would blacklist an agent for opportunistic behaviour. In addition to grading the evaluation with Attributes, the shell nouns **probability** and **possibility** can themselves be seen as graded: downscaled fulfilment. This aspect was covered in the section dealing with shell nouns as an interpersonal resource for GRADUATION (section 6.3.2, p.263).

Shell noun type	Published	Submissions
modal	probability (3), possibility (1)	probability (17), possibility (1), chance (1)
eventive		alternative (1), choice (1)

Table 6.42 Shell nouns enacting JUDGEMENT: SOCIAL SANCTION: veracity in Published and Submissions

The shell nouns found to enact judgements of veracity are listed in table 6.42. The judgements of veracity are present in six texts: S_Rand, P_Rand, S_Net, P_Pov, S_Pov, and P_CEO. Except for S_Rand and P_Rand, they are all empirical studies. S_Rand and P_Rand are more mathematical.

Comparison of the two sub-corpora shows that there are no statistically significant differences in the relative shares of each judgement type. Judgements of veracity and capacity account for over two thirds each of the instances. Judgements of propriety make up a smaller proportion, and judgements of normality and tenacity, which are only found in Submissions, are very minor. The full results of the distribution of selections from the system of JUDGEMENT are given in Appendix G.3.

With regard to the shell nouns used to enact judgements, again the results show that there is no statistically significant difference in their distribution (table 6.43). Modal shell nouns by far account for the greatest share.

Feature	Published		Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
SHELL-NOUN	N=11		N=51					
factual	0.0%	0	2.0%	1	0.00		0.22	
linguistic	9.1%	1	5.9%	3	0.39		0.15	
mental	9.1%	1	2.0%	1	1.21		1.47	
modal	81.8%	9	80.4%	41	0.11		0.01	
eventive	0.0%	0	5.9%	3	0.00		0.68	
circumstantial	0.0%	0	3.9%	2	0.00		0.45	

Table 6.43 Distribution of shell nouns enacting JUDGEMENT in Published and Submissions.

Alongside assessments of Appreciation and Judgement, the corpus contains a few instances of assessments of Affect. The shell nouns used for this function are either nominalised desiderative mental Processes (**desire, hope, intention, and agreement**) or nominalised qualities (**certainty, disinclination**). They are all classed as mental shell nouns, except for **certainty**, which is modal. Examples are given in 6.163-6.165.

- 6.163 he [Mills] was convinced by a battery of reasons, including the **certainty** that the Society would not cultivate mathematical ‘esoterism’ or any ‘kind of separatism’. (S_Econ)
- 6.164 Obviously, weakly responsive preferences are quite restrictive because they (a) almost take us back to the model with only singles and (b) do not capture important aspects such as a couple’s **desire** to live together. (P_Rand)
- 6.165 In his earliest writings such as Moore 1908, he assumed the mantle of a proselytizer within the neoclassical camp for an explicitly statistical economics, praising Cournot, Edgeworth and Pareto, and hinting that economic laws might be initially discovered by empirical statistical methods, and then later rationalized using the principles of “pure economics”. Yet in 1912 Moore’s **hopes** had been rudely discounted by Marshall (P_Econ)

The shell nouns make it possible to construe the feelings in a more objective sounding way. When the lexicalisation construes Qualifier, for example in 6.163 and 6.164, the shell noun projects the stimulus for the feeling, i.e. the Trigger, and the person whose feelings are described, the Emoter, is explicitly present. While the shell noun is not field-specific, both the lexicalisation and the Emoter are. Thus, in 6.163, which comes from the published article on the debates among economists surrounding the development of econometrics, it is the feelings of the economist, Mills, that are described, the shell noun **certainty** construing the subcategory of security. And in 6.164, which is from the published article on random matching, the couple represent one half of the matching market, the shell noun **desire** enacting the subcategory of desire. When the lexicalisation is realised prior to the shell noun, the anaphorically referring shell noun functions to ascribe affect, the lexicalisation, again, functioning as the Trigger. Thus, in 6.165, the shell noun **hopes** encapsulates the preceding information.

Unlike the examples above, two of the instances of AFFECT in the corpus construe an Emoter not from the field of the object of study, but from the field of research. In this case the shell noun might be seen to function concomitantly as a resource for construing dialogic position and Attitude. Both 6.166 and 6.167 come from the introduction section of the research article. In 6.166, the feelings are attributed to other researchers while in 6.167 the shell noun enables the writer to construe his own feelings in a more objective-sounding way. In the examples the Emoter is boxed, the shell noun is in bold and the lexicalisation is underlined.

6.166 The concept rose criticism, not for what it intends to apprehend – the connection between institutions, shared believes and economic performance – rather for its promoters’ intention to include the issue in the range of “capital” theoretical implications given the similarities with other sort of capital, namely human capital. (S_Net)

6.167 Our choice was shaped by the **desire** to use a range of institutions which would cover merchants in different spheres, various stages of their careers, and also by the longevity of extant data in order to be able to measure change over time. (P_Net)

Example 6.167 is quite complex and can be interpreted in multiple ways. In addition to Affect, the instances of Graduation in the lexicalisation, in italics, can invoke a positive Appreciation of the researchers’ desire and ultimately a positive token of Judgement of the researchers themselves. Nominalisation of the desiderative mental Process certainly contributes to the multiple interpretations. What is not ambiguous is that the shell noun enacts an interpersonal function.

Table 6.44 lists the instances of shell nouns functioning to ascribe Affect. In each case, the shell noun is in bold and the lexicalisation is underlined.

Shell noun	Published	Submissions
mental	the desire <u>to use a range of institutions which would cover merchants in different spheres, various stages of their careers</u> (P_Net) a couple’s desire <u>to live together</u> (P_Rand)	in agreement <u>that it is the best stable matching</u> (S_Rand) its promoters’ intention <u>to include the issue in the range of “capital” theoretical implications</u> (S_Net) the legislator’s intentions [anaphoric] (S_Net)
	Moore’s hopes [anaphoric] (P_Econ) an early disinclination <u>to defend the neoclassical theory of pricing and allocation</u> (P_Econ)	the certainty <u>that the Society would not cultivate mathematical ‘esoterism’ or any ‘kind of separatism</u>

Table 6.44 Shell nouns as AFFECT in Published and Submissions

The qualitative analysis of shell nouns as a resource in the system of ATTITUDE has so far focussed on describing how shell nouns contribute to enacting evaluative meanings while the quantitative analysis has compared their use in the Published and Submissions sub-corpora and whether there are differences in use. When the corpus is approached as a whole, it is possible to quantify the tendency for certain shell noun types to enact a particular function. Leaving aside the system of AFFECT, of which there

are only eight instances, table 6.45 shows the distribution of shell nouns across the systems of APPRECIATION and JUDGEMENT for the corpus as a whole.

Feature	Appreciation		Judgement		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
SHELL-NOUN	N=186		N=62					
factual	38.7%	72	1.6%	1	5.91	+++	30.81	+++
linguistic	10.2%	19	6.5%	4	0.88		0.78	
mental	29.0%	54	3.2%	2	4.35	+++	17.71	+++
modal	3.2%	6	80.6%	50	21.04	+++	159.43	+++
eventive	4.3%	8	4.8%	3	0.18		0.03	
circumstantial	14.5%	27	3.2%	2	2.41	+++	5.74	+++

Table 6.45 Distribution of shell nouns in APPRECIATION and JUDGEMENT in whole corpus

The results indicate that there is a strong tendency for factual, mental, modal and circumstantial shell nouns to construe a particular function. While modal shell nouns operate predominantly in the system of JUDGEMENT, factual, mental and circumstantial shell nouns are more likely to be found in APPRECIATION. These results attain high statistical significance. No particular trend is found for linguistic or eventive shell nouns. The modal shell nouns enable the writer to encode judgements at

6.3.4. Modification of shell nouns by means of GRADUATION

Before concluding the descriptive analysis, there is one last way in which shell nouns interact with interpersonal meaning: they may be graded in terms of force or focus by the resources of GRADUATION. Selections from the system of FORCE include resources for intensifying qualities as well as resources for quantifying entities. The results show that in both sub-corpora, the resources for intensifying qualities are almost all employed in nominal groups with cataphorically or esphorically lexicalised shell nouns as Head, some of which may be functioning in a hyper-Theme. An example of intensification fused in the Epithet is given in 6.168, where the resource of GRADUATION is boxed, the shell noun in bold, and the esphoric lexicalisation underlined.

6.168 Under the law, a government grant is “treated as a contribution if its primary **purpose** is to enable the done to provide a service to, or maintain a facility for, the direct benefit of the public”. (P_CEO)

In 6.168 the Epithet ‘primary’ can be considered an intensification of the quality of importance. This interpretation follows Hood (2010). By contrast, 6.169 illustrates a

cataphorically lexicalised shell noun in which the quality of being detailed is scaled. This example also composes a hyper-Theme. It comes from the article on business networks in 18th century Portuguese-Brazilian trade.

6.169 The historical fundaments of these results deserve a more detailed **explanation**. (S_Net)

The instances of FORCE: INTENSIFYING QUALITIES that are found in the nominal groups include upscaling in Epithets such as *key*, *primary* or *fundamental*, as well as downscaling via Adjuncts such as *roughly* or Epithets e.g. *reduced*. Appendix H.1 lists the instances of FORCE: INTENSIFYING QUALITIES that are found in the corpus.

Grading shell nouns by means of quantification is quite frequent in the corpus. Again these instances tend to appear in hyper-Themes, and/or cataphorically lexicalised and esphorically referring nominal groups with shell noun as Head. Grading of a cataphorically lexicalised shell noun in a hyper-Theme is exemplified in 6.170.

6.170 For a variety of **reasons** nonprofits provide an interesting context in which to study the pay of managers in this period. First ... (P_CEO)

The selections from GRADUATION, boxed, couple with the resources of internal conjunction, in italics, to organise the lexicalisation, which is underlined. With esphorically referring shell nouns whose lexicalisation is realised with a fact clause, it is common to find the Numerative 'one' quantifying the shell noun, as in 6.171. As argued in section 6.3.1, such instances also enact dialogic position by means of a framing matrix.

6.171 One **reason** why we would expect managers of nonprofits with more assets to earn more is that they have responsibility over a much larger number of people and resources. (P_CEO)

There are also instances where the entity is quantified with regard to amount rather than number. (6.172) is one of the few examples of an anaphorically lexicalised shell noun coupling with a selection from GRADUATION: FORCE. The instance is taken from the Submissions article on econometrics, and the degree of **success** is upscaled by means of the Epithet *immense*.

6.172 By the end of the first decade of the Society's existence, forty-two Fellows represented the Olympus of the 'Econometric people'. This was, in any case, an immense success. (S_Econ)

The instances of quantifying in GRADUATION: FORCE found in nominal groups with shell noun as Head are detailed in Appendix H.2. Most instances quantify the shell noun. They include Numeratives such as *several*, *some*, *three*. But there are also instances in which the shell noun is graded in terms of mass or presence such as 'this *large effect*' (S_CEO), 'an *alternative strategy*' (S_Rand) or extent such as 'an *immense success*' (S_Econ), 'the *low potential* of the network to foster norms' (S_Net) or 'the *perennial complaint*' (P_Econ).

The nominal group with shell noun as Head can also be graded with respect to the experiential boundaries of the noun. In this case, a modifier functions within the system of FOCUS, either as valeur or fulfilment. Modification of the focus of the shell noun as valeur is found with anaphorically and esphorically referring shell nouns as well as cataphorically lexicalised ones. It is found in both sub-corpora. An example of the authenticity of a shell noun being upscaled and one of the specificity of the shell noun being downscaled are given in 6.173 and 6.174 respectively. The element of focus is boxed, the shell noun is in bold, and the lexicalisation is underlined.

6.173 A classic **example** of how difficult it is to measure performance for managers in the nonprofit sector is the case of a manager of a nursing home. (P_CEO)

6.174 The general **conclusion** is that the CEO's earning in the NPO sector are defined by multiple factors, some of them being individual attributes defined by the market, such as individual characteristics, while others are defined by the performance of the NPO. (S_CEO)

The instances of adjustments of FOCUS as valeur that are found in the corpus are listed in Appendix H.3. Examples of upscaled valeur include 'the *particular case*' (S_Rand), 'A *classic example*' (P_CEO), or 'a *particular remark*' (S_Net), and downscaled valeur '*hardly an encouragement*' (P_Econ), '*just a way*' (S_Net) and 'The *general conclusion*' (S_CEO).

Finally, the meaning of the shell noun can be graded with respect to fulfilment. Earlier, I have described how shell nouns such as **attempt**, **effort**, or **initiative** contribute directly to grading the fulfilment of a proposition. Here I am concerned with

how selections from the system of FULFILMENT can grade the actual shell noun. Because shell nouns encapsulate propositions, they can also be modified with respect to how actualised they are. For example in 6.175 'potential' grades the course of action as not fully actualised.

6.175 One potential course of **action** is therefore to impose that under any play of the sequential game the choices actually made are consistent with some sort of preference ordering. (S_Rand)

There are not many instances of shell nouns being graded in this way in the corpus. Modifiers grading fulfilment include *possible*, *not necessarily*, *perceived*, *expected*, and *potential*. The nominal groups in which they are found are reported in Appendix H.4. It excludes instances such as **possibility** or **propensity** in which it is the shell noun that enacts the interpersonal meaning of fulfilment. As previously mentioned, graduation infused in the shell noun is addressed in section 6.3.2 on pages 263-265.

6.4 . Summary of results and discussion

The first part of this chapter has described the results for the contributions that shell nouns make to ideational, interpersonal and textual meanings at a level beyond the clause. In the main the analysis has been qualitative, although some attempt at a quantitative analysis of the data has been made. The results are also described for the ways in which shell nouns are modified using selections from the textual, ideational and interpersonal systems. I shall now summarise these results, incorporating a discussion of how they relate to those of other studies. The direct contribution of shell nouns to the textual, ideational and interpersonal metafunctions will be presented first, followed by how shell nouns enable coupling of meaning among functions, and in doing so, naturalise particular reading positions.

6.4.1. The contribution of shell nouns to textual meaning

The results of the analysis show that shell nouns contribute directly to textual meaning across clauses by identifying a previous or following clause, or clauses, and encapsulating its, or their, meaning in the noun and the nominal group. This function constitutes an important means by which the text is developed, and it lends weight to the argument that encapsulation and prospection are main means by which a text develops (Álvarez-de-Mon y Rego, 2001; Sinclair, 1993).

A cohesive tie to the preceding or following discourse is often established by means of presuming reference in the nominal group with shell noun as Head or by means of text reference in a relational Process clause in which the nominal group with shell noun as Head takes presenting reference. Although shell nouns also refer esphorically, i.e. the referent clause is realised either as Qualifier or fact clause in a relational Process clause, because this reference is within the clause, it is not discussed here as a textual function at the discourse semantics stratum. Shell nouns of this type are analysed for their textual function at the stratum of the lexicogrammar in sections 5.3.5 and 5.4.5 of chapter 5.

In both sub-corpora anaphorically referring shell nouns are more frequent than cataphoric ones. Most of the anaphorically referring shell nouns are found to refer to the preceding clause or clauses, and the link – established by means of presuming

reference – operates across sentence boundaries. This result is in keeping with those of Gray (2010), whose findings in a corpus of research articles in the fields of Education and Sociology are similar. In addition, in my study there are instances in which there is a clause or more intervening between the nominal group with shell noun as Head and the stretch of text that it refers to. The distribution of shell noun types in anaphorically referring nominal groups differs between the Published and Submissions sub-corpora. In the Published sub-corpus circumstantial shell nouns such as **approach** and **criteria** make up the largest share, followed by factual, mental and linguistic shell nouns in decreasing order of relative share. In the Submissions sub-corpus, factual shell nouns – in particular **result** – account for most instances, followed by mental, linguistic then circumstantial shell nouns. In both sub-corpora most shell nouns are single instances. The only shell nouns to attain multiple counts in both sub-corpora are **result**, **criteria**, **approach** and **case**. The result for the high counts of the shell noun **result** and the single counts for most of the others is similar to those of Dueñas (2003-2004), whose corpus comprised American English and Spanish medium research articles in the field of business management. However, unlike her findings, linguistic shell nouns are found to constitute a sizeable share in both sub-corpora (15% in Published and 19% in Submissions). This difference is accounted for by the high counts obtained in the paired texts on Econometrics.

Unlike anaphorically referring shell nouns, which typically take presuming reference, cataphorically lexicalised shell nouns typically take presenting or neutralised reference, a result that is consistent with other studies (e.g. Tadros, 1994). A great number of these cataphorically lexicalised shell nouns appear in hyper-Themes or macro-Themes. Macro-Themes (headings) such as ‘Results’ or ‘Limitations’, which are composed with neutralised reference, organise large stretches of discourse. Hyper-Themes with a cataphorically lexicalised shell noun can similarly organise information across several clauses. In some cases the expanse of text referred to is limited to a clause or clause complex, in which case the clause containing the shell noun cannot be considered a hyper-Theme although it still predicts the upcoming information. The forward organisational function of shell nouns has been termed signposting (Schmid,

2000), and the results of this study show how shell nouns constitute a systemic option for this function.

A cataphorical lexicalisation that spans several sentences and is predicted by a hyper-Theme is often itself scaffolded by means of resources from internal conjunction, some of which might themselves be composed by esphorically referring shell nouns in relational Process clauses. In such cases, a textual modifier to the shell noun locates its lexicalisation while the shell noun interprets it. The most common way to do this is with the Post-Deictic *following*, but there are also instances of reference to order rather than location (e.g. *next*, *first*). The use of such textual modifiers with cataphorically referring shell nouns is in line with results from Tadros (1994) and Bezerra (2009).

The most frequent shell noun type to be used with a cataphoric reference function is factual. The most common factual shell nouns in descending order of frequency are **result**, **reason**, **proof**, and **example**. These shell nouns attain counts of 10 or more. The other factual shell nouns are used fewer than 5 times. The only other shell nouns to attain counts of above 5 are **remark** and **proposition**, which are linguistic, and **hypothesis**, which is mental. All other shell nouns, which cover all groups, are made up of between 1 and 4 instances.

A comparison between anaphoric and cataphoric reference functions brings to light differences in the distribution of shell nouns between the two functions. Although factual shell nouns account for the greatest share in both functions, they are more likely to compose cataphoric reference than anaphoric. By contrast, circumstantial shell nouns are more likely to compose anaphoric reference than cataphoric. Linguistic shell nouns are found more commonly in cataphoric reference than anaphoric, and for mental shell nouns this trend is reversed. The propensity of factual shell nouns to function cataphorically rather than anaphorically, and the propensity of circumstantial shell nouns to function anaphorically rather than cataphorically is found to be statistically significant. The larger share of factual cataphorically lexicalised shell nouns may be due to their use in hyper-Themes and internal conjunction. The shell noun **reasons** in particular shows multiple instances in this function.

As previously mentioned, some of the instances of cataphorically lexicalised shell nouns in hyper-Themes are in fact instances of internal conjunction. Shell nouns contribute to internal conjunction in several different ways. Circumstances such as ‘in two **ways**’ or ‘for several **reasons**’, which appear in hyper-Themes, organise the discourse and preview the possibility for the rest of the paragraph to be scaffolded through the resources of internal conjunction. The cataphorically lexicalised shell nouns operating in internal conjunction are always identified by a Numerative or presenting reference. Another option for internal conjunction can be realised by an esphorically referring shell nominal group construing Value in a relational Process clause (e.g. The first possible **reason** is that ...). The shell noun **reason** is the most commonly employed in this way. Other instances of this lexicogrammatical pattern include the factual shell noun **advantage**, and linguistic shell nouns **explanation** and **conclusion**. Finally anaphorically referring shell nouns combine with general options from internal and external conjunction that are realised as conjunctive Adjuncts, in which case they can add interpersonal meaning through attitudinal inscription (e.g. Besides these **advantages**) or ideational meaning in terms of field (e.g. In view of this **result**). Unlike the esphorically referring shell nouns, these anaphorically referring shell nouns tend to be mental, with **purpose** accounting for most of the instances (e.g. With this **purpose** in mind). Anaphorically referring shell nouns used in general internal conjunction options include **purpose**, **view**, **result**, and **advantage**. These options typically compose textual Theme, and they function to make the reasoning behind the developing argument explicit.

The contribution of shell nouns to internal conjunction has been noted by Martin (1992), and to hyper-Themes by Martin (1992) and Ravelli (2004). Because their full meaning must be retrieved from the co-text through reference, they scaffold the development of the discourse (Francis, 1994; Martin, 1992; Ravelli, 2004). The results of my study confirm these textual functions of shell nouns.

In addition to functioning as a resource for composing hyper-Themes and internal conjunction, shell nouns are found to contribute to composing hyper-News. Only a few instances of this function are found. The hyper-News can be found at key transitional points in the article, for example the transition from the descriptive report

of others' research to the descriptive report of the writer's own study in the introduction or the transition from methodology to presentation of results. The low number of instances is perhaps surprising, but this use of shell nouns is similar to Shaw's (2000) result that shell nouns often appear in the discussion section of research articles and dissertations, particularly at the transition from the statement of results to the next stage in the argument.

6.4.2. The contribution of shell nouns to ideational meaning

The analysis of shell nouns and experiential meaning in this study focuses on the relation of the shell noun and its lexicalisation to field. The results show that most shell nouns do not contribute specifically to field, and those that do cluster around fields associated with research. The general field of research is construed through the factual shell nouns: **result(s), corollary, finding, evidence, phenomenon, and proof**; the mental shell nouns: **objective, aim, assumption, hypothesis, thesis, premise, lemma, intention, theory, theorem, and analysis**; and linguistic shell nouns: **claim, argument, contribution, and conclusion**. These shell nouns tend to be found in all of the texts of the corpus, regardless of the object of study. The field of methodology is construed through the circumstantial shell nouns: **approach, method, and criteria**; as well as the modal shell noun: **specification**. Shell nouns that construe the object of research understandably vary and tend to be restricted to the parallel texts of the corpus. For example, the two texts on the history of the development of Econometrics as a legitimate field of study within Economics describe the debates among academics, and this is construed mainly through linguistic shell nouns: **observation, statement, assertion, argument, proposal, remark, suggestion, and comment** and mental shell nouns: **interpretation, belief, consideration, idea, opinion, and interpretation**. The pair of texts on random matching construe the decision making process of one side of the matching market through mental shell nouns: **desire, prospect and strategy**, as well as the modal shell noun **condition**. While the shell nouns that construe field as research are found across all of the texts of the corpus, those that construe field as methodology are restricted to five of the texts: P_Pov, P_CEO and P_Rand, S_Pov, and S_CEO. The higher frequency of shell nouns that construe field as research is in line

with results from other studies that focus on academic discourse (Flowerdew, 2003 b; Swales, 2005; Gray, 2010; Gray & Cortes, 2011; Gledhill, 2000).

However, the fact that the shell nouns that construe field as methodology appear in only half the texts is, perhaps, a little surprising. While they may understandably be absent from P_Econ and S_Econ, all the other texts are empirical studies, and in the field of economics, it has been noted that often the methodology and the models of the study are considered more important than the results obtained from the models (Bloor & Bloor, 1993). Furthermore, in the last quarter of the 20th century there has been a growing trend to use “empirical terminology”, i.e. from the fields of research and methodology, in four major economics journals (Goldschmidt & Szmercsanyi, 2007). It may be that in some cases the shell noun was classified as construing the field of research rather than methodology. For example the noun **hypothesis**, which is found in S_Net, might be considered as construing the field of methodology given that the paper presents two hypotheses that are tested and then confirmed or refuted. Although the finding that most shell nouns do not construe a specific field may help to explain why in other studies it has been reported that shell nouns are neither discipline nor register specific (Gray & Cortes, 2011), the results of this study show that field as the object of study discussion can be construed through shell nouns. The prevalence of linguistic shell nouns to construe field in the texts on the development of econometrics, P_Econ and S_Econ, echoes the finding of Charles (2003; 2007), who found that academic discourse in the area of politics uses a higher proportion of metalinguistic nouns, which correspond essentially with linguistic shell nouns (Schmid, 2000), than does the area of material sciences. Politics is concerned with the statements and positions of the politicians just as the development of econometrics in P_Econ and S_Econ is concerned with the statements and positions of the economists involved in establishing econometrics.

Unlike most of the shell nouns in the corpus, their lexicalisations are found to be field specific. Most lexicalisations construe field as the object of study. The lexicalisations are also found to construe field as methodology, field as the context of the object of study, as well as mixed fields, which include the field of research

metaphorically projecting the field of the object of study or methodology projecting the object of study.

When the lexicalisation construes field as the object of study, the most common realisation is to project the lexicalisation as Qualifier to the shell noun. This finding is in line with a result from Gledhill (2000), whose study of introductions to research articles on cancer found that the shell noun 'ability' was typically used to project biochemical processes as Qualifier. The lexicogrammatical pattern corresponds to Schmid's (2000) temporary concept function, which he argues lightens the burden of cognitive retention of the information. A Systemic Functional Linguistics perspective explains the systemic mechanism by which this function is achieved: rank-shifting a clause downwards to function as Qualifier enables it to function within a nominal group that construes a Participant. Because one of the properties of shell nouns is that the meaning becomes specific only when the relevant information of the lexicalisation is retrieved from the co-text, the meaning of the nominal group is instantiated in the text and is not permanent. At the same time, the shell noun creates a discourse entity that can enter into different activity or implication sequences, thus functioning as a concept.

A non-field specific shell noun or even a field-specific one may be modified by a field-specific Classifier or Qualifier realised as a prepositional phrase. Classifiers such as 'endogeneity', 'financial' or 'equilibrium' help specify the shell noun but do not constitute its lexicalisation. Similarly, Qualifiers such as 'with local and confidential networking' or 'of existing measures' also help specify, but the lexicalisation must be retrieved elsewhere. Both the Submissions and Published sub-corpora show instances of such ideational modification, and this finding corroborates those of other studies of academic discourse (Francis, 1994; Moreno, 2004).

6.4.3. The contribution of shell nouns to interpersonal meaning

The results of the study show that shell nouns contribute directly to interpersonal meaning at the stratum of discourse semantics in several ways. First, some shell nouns contribute to creating a heteroglossic text by expanding or contracting the dialogic space between competing voices. Second, some shell nouns

are inscribed with attitudinal value, thus they can serve to assess propositions positively or negatively, judge people and their behaviour, and in some cases ascribe feelings. The third interpersonal function of shell nouns is to grade the value of a proposition. In addition, shell nouns couple with other resources from the systems of ATTITUDE or GRADUATION to inscribe or invoke evaluative readings. This section will summarise the ways in which these evaluative positions are realised in the corpus.

Shell nouns can contribute directly to creating heteroglossic text. Whether the dialogic space accorded to other voices is opened up or shut down depends on both the choice of shell noun and its lexicogrammatical environment. Specific lexicogrammatical patterns can provide a framing matrix for the shell noun which results in a particular dialogic position.

One of the matrices for enacting contracting discourse is when the shell noun – composed with presuming reference – construes Value, and its lexicalisation – realised by a fact clause – construes Token in a relational Process clause (e.g. The **problem** is that ...). Because the nominal group with shell noun as Head is composed with presuming reference, the reader is constrained to interpret the lexicalisation in accordance with the writer's view, and as a result, other views are excluded. This effect is most noticeable with the factual shell nouns found in the corpus (**advantage, fact, irony, point, problem**), but it is also present in the linguistic shell nouns (**question, conclusion**), mental shell nouns (**aim, concept, concern, consideration, function, motivation, objective, purpose, rationale, solution, strategy**), modal shell noun (**step**), and eventive shell noun (**priority**). Moreover, with the non-factual shell nouns, the matrix provides for a more objective sounding text because the grammatical metaphor of the shell noun makes it possible to omit the Sayer/Senser/Actor. The results show that this matrix is more commonly used in the Published sub-corpus than in Submissions. Schmid (2000; 2001) discusses the pattern in terms of topicalising, including within it shell nouns that take presenting reference (e.g. One objective is to). Such a perspective limits interpretation of meaning to the stratum of the clause. When considered at the stratum of discourse semantics, the difference in reference corresponds to a difference in interpersonal meaning. While presuming reference in esphorically referring shell nouns whose lexicalisation is a fact

clause construing Token functions to contract the dialogic space, presenting reference functions to open it up. The latter aspect will be discussed shortly in the summary of the results on expanding discourse.

A second lexicogrammatical pattern that contracts the dialogic space makes use of the factual shell nouns **fact**, **focus** and **accident**, and mental shell noun **intention**, and in the majority of cases couples with negative polarity. When there is a split Subject (e.g. It is a **fact** that...; It is no **accident** that...), the matrix once again constrains interpretation of the fact clause that enacts Subject and constitutes the lexicalisation of the shell noun.

Shell nouns can also lead to a contractive text by projecting the lexicalisation as Qualifier. Again, the shell noun takes presuming reference and the shell nouns are from a restricted group made up of factual shell nouns (**fact**, **evidence**, **proof**, **result**, **factor**), linguistic shell nouns whose meaning conveys a high investment on the part of the writer (**argument**, **assertion**, **statement**), or the mental shell noun **goal**. The shell noun adds an additional layer of meaning in which the embedded clause is removed from negotiation. The shell noun **proof** is found to close the dialogic space in headings as well.

Some shell nouns function to endorse another academic's propositions. They include the factual shell nouns: **corroboration**, **fact** and **result**, and the linguistic shell nouns: **law** and **observation**. These shell nouns are mainly used to refer to the endorsed proposition anaphorically or cataphorically, but there are a couple of instances of esphoric reference where the lexicalisation is construed as Qualifier. In each case, the named source is retrievable in the nearby co-text.

Finally, shell nouns in Relators or in Circumstances of cause can contribute to a heteroglossically contractive text. In the latter use, 'due to the **fact** that' is the most frequently found instance. Relators such as 'with the **aims** of', 'with the **purpose** of', or 'on the **grounds** that' specify the relation between the secondary and primary clause, and in doing so, exclude other perspectives.

Just as some shell nouns function to close the dialogic space, others function to open it up. In particular mental and linguistic shell nouns can clearly mark a proposition as but one among competing voices. The most common linguistic shell nouns used for this function in the corpus are: **proposal**, **conclusion**, **suggestion**, **argument**, **question**, **remark**, and **proposition**. The most common mental shell nouns that are found to expand the dialogic space are: **hypothesis**, **analysis** and **idea**. These linguistic and mental shell nouns can either attribute the proposition to another voice, or present the writer's own. While mental and linguistic shell nouns are the main systemic means of constructing a heteroglossic expanding text, factual shell nouns can also do so when they appear in certain lexicogrammatical environments. A factual shell noun with presenting reference that construes Value in a relational Process clause (e.g. One of the **reasons** is that ...) or when graded as not fully actualised (e.g. The first possible **reason** is that...) construe the text as heteroglossic expanding, although it can be argued in the latter case that it is the modifier *possible* that carries the weight of the function. Factual shell nouns also appear in existential Process clauses (e.g. There is **evidence** that ...). Halliday and Matthiessen (2004) note the tendency for **evidence** to appear in this lexicogrammatical pattern. The final way in which some circumstantial shell nouns – in particular **case** – function to open the dialogic space is in Circumstances (e.g. in this **case**). The Circumstance with an anaphorically referring shell noun neutralises the distinction between *if* and *when*, but continues to acknowledge the existence of other voices.

Comparison between the two sub-corpora shows that there are significant differences in the dialogic positioning construed by shell nouns. The Published sub-corpus has a higher proportion of contractive text than the Submissions, and within dialogic expanding text, the Published sub-corpus shows a greater tendency to attribute the ideas to others than to present them as the writer's own. In particular, the shell noun **fact** is used more frequently to close the dialogic space in the Published than in Submissions. Underuse of **fact** has been found in learner academic discourse (Atkas & Cortes, 2008), but the result contrasts with those of Parkinson (2013), who found the opposite trend, and Hasselgård (2012), who found no difference between Anglophone and EAL writers. The differences in dialogic positioning in my study could

result in quite different responses from the reader; a less contractive text that entertains the writer's voice as but one among others may come across as less assertive while a more contractive text whose acknowledgement of other voices lies in attribution rather than presenting the writer's own may come across as more authoritative. The difference in dialogic position may in part shed light on the reviewer's criticism that the southern European writer was using potential shell nouns in 'vague' meaningless ways (Lillis & Curry, 2010).

The results for the contribution of shell nouns to construing Attitude show that they can carry positive or negative assessments of Things, judgements of people and their behaviour, and feelings, and thus can launch a positive or negative assessment that spreads to the lexicalisation. Factual shell nouns such as **advantage** or **problem** clearly evaluate the lexicalisation as positive or negative, respectively. Negative assessments of Appreciation can also be construed with circumstantial shell nouns such as **limitation** or **shortcoming**. Judgements of veracity can be construed with modal shell nouns such as **probability**, **possibility**, **risk**, or **chance**; Judgements of propriety with modal shell nouns such as **need** and mental shell nouns such as **policy**; Judgements of capacity with modal shell nouns such as **ability**, **inability**, **capacity**, **impossibility**, **opportunity**, **power**, and factual shell nouns such as **difficulty**; Judgements of normality with the modal shell noun **propensity**; and Judgements of tenacity with linguistic shell nouns such as **encouragement** and **commitment**. In the main, the lexicalisation of the shell noun, which construes the behaviour being assessed, is realised in a Qualifier to the shell noun. In a few cases shell nouns are found to enact Affect. The mental shell nouns **desire**, **hope**, **disinclination**, **agreement**, **intention**, and **certainty** are found to ascribe feelings to a Participant, again, the lexicalisation tending to be realised as Qualifier.

There are significant differences between the two sub-corpora in the distribution of shell nouns infused with attitudinal value. While in both sub-corpora Affect makes up only a minor proportion of the instances, the Published sub-corpus shows a relatively higher proportion of Appreciation and lower proportion of Judgement than the Submissions. This difference can mostly be accounted for by the repeated instances of **probability** in the Submissions. Finally, a further difference

between the two sub-corpora is that the Submissions sub-corpus has a higher proportion of positive assessments than the Published. Again this difference can be attributed to the high number of instances of positive judgements of veracity in the Submissions. Yet, even within the system of Appreciation, the Submissions sub-corpus shows a greater tendency to assess semiotic entities positively. This finding may be in line with Moreno *et al.* (2012), who note that Spanish researchers are reluctant to appraise their colleagues' work negatively. Further work, however, is necessary to determine the distribution of positive and negative assessments among other researchers' work, the writer's own work and other entities.

The final way in which shell nouns contribute directly to interpersonal meaning is by grading a proposition, in this case, the lexicalisation of the shell noun. In most cases the lexicalisation is graded in terms of Focus. Modal shell nouns (**probability, possibility, potential, opportunity, risk**) and the mental shell noun **prospect** can grade a proposition in terms of how actualised it is while eventive shell nouns – **effort, initiative, attempt**, and **struggle** – can grade the proposition in terms of completion. The shell noun **specification** can be seen to grade the lexicalisation as upscaled valeur. The Submissions sub-corpus shows a significantly higher proportion of shell nouns used to grade the lexicalisation as not fully actualised. This is again due to the high number of instances of **probability**. There is also one instance in which the shell noun **tendency** functions to grade the force of a proposition in terms of how widespread it is.

The above-mentioned shell nouns make it possible to grade the proposition because they are grammatical metaphors, nominalisations. By nominalising the modal Adjunct or a verb of conation, the degree of actualisation or completion is lost. Thus, 'Other characteristics of managers and organizations are possibly confounding the relationship between firm size and managerial pay in nonprofits' does not mean the same as 'the possibility that other characteristics of managers and organizations are confounding the relationship between firm size and managerial pay in nonprofits' (P_CEO), and 'They attempted to motivate young colleagues' does not mean the same as 'the attempts to motivate young colleagues' (S_Econ). While Halliday (1990) notes that some meaning is lost in the grammatical metaphor, looking at these grammatical

metaphors from a different perspective shows that a different meaning is made and new possibilities arise. The embedded propositions that are not fully actualised or not fully complete double as judgements that may be further graded (e.g. The possibility that other characteristics of managers and organizations are confounding the relationship between firm size and managerial pay in nonprofits is high/low/non-existent.) This capacity makes shell nouns that enact Judgements a valuable systemic resource in research, particularly in empirical studies.

This section has summarised the direct contribution that shell nouns can make to textual, experiential, and interpersonal meaning at the stratum of discourse semantics. The next section will deal with the ways in which meanings are coupled from different systems or ranks and how this coupling construes meaning in the developing text.

6.4.4. Coupling among textual, ideational and interpersonal metafunctions

The coupling among metafunctions with shell nouns can result in a text that naturalises a reader position favourable to the researcher's work and critical of other studies. This burnishing and tarnishing (Hao & Humphrey, 2012; Ventola, 1998) is found in both the Published and Submissions sub-corpora. Coupling may involve all three metafunctions.

Field-specific shell nouns – particularly those from the field of research such as **results** – couple with the textual functions of anaphoric or cataphoric reference. In fact, in both sub-corpora the most common shell nouns to refer anaphorically construe the field of research or research methodology. In addition to contributing to the logogenesis of the text, these field-specific anaphorically referring shell nouns are also found to couple with the system of Attitude, such that the nominal group is inscribed with positive or negative value, which then spreads to the lexicalisation. Thus, the author's **results** can be valued as 'consistent' or 'robust' while another researcher's **results** may be 'not very illuminating' or 'uninformative'; the former burnished, the latter tarnished. Such attitudinal inscription is found to be realised mainly as Attribute, but there are instances in which it is realised as Epithet. Interpersonally oriented

Epithets are found in both Published and Submissions, a result that differs from Moreno (2004), who found that Spanish authors did not use them.

Coupling between interpersonal and textual functions can also be infused in the shell noun construing Head in an anaphoric or cataphoric nominal group. Shell nouns such as **difficulty** or **advantage** are infused with negative or positive value, and as a result of the cohesive tie between the shell noun and its clausal referent, the prosodic value spreads retrospectively or prospectively to the text whose meaning the shell noun encapsulates. In particular, because appraisal of an entity, in this case a shell noun, in terms of the systemic option of valuation is sensitive to field (Martin & White, 2005), using a shell noun like **paradox**, which is found in my study, or **mystery**, which is from Charles (2003), both of which carry positive infused attitudinal value and encapsulate prior information, can help create the justification for writer's study in the research article introduction. Used in this way, the selection of shell nouns that are infused with attitudinal value enables the writer to align the reader with his values.

Previous studies have identified these two systemic resources (infused and isolated attitudinal inscription) for coupling textual (anaphoric) meaning with interpersonal, but they do not provide a comprehensive explanation of how they work. Francis (1994), Flowerdew (2003a), Moreno (2004), and Gray (2010), for example, found similar instances of interpersonal meaning realised as isolated attitudinal inscription in their studies of anaphoric shell nouns; however, their interpretation that the anaphorically referring shell noun functions as a means to introduce and focus on the evaluation realised as Attribute or as a means to slip it in as Epithet "without having to make a special point of it" (Francis, 1994: 97) is limited to analysis at the stratum of the lexicogrammar. Similarly, Charles (2003), and Shaw (2000) note that sentence initial anaphorically referring shell nouns like **mystery** contribute to creating writer stance, but again, they offer no theory-based explanation. The Discourse Semantics approach of Systemic Functional Linguistics that underpins my analysis contributes to a theoretical explanation for how shell nouns constitute a systemic resource for aligning the reader with the values proposed by the author by enabling coupling of interpersonal and textual meanings.

One further way in which interpersonal meaning is coupled with ideational meaning is when an attitudinally inscribed shell noun construes Attribute in a relational Process clause whose Carrier is realised by text reference. Schmid (2000) identifies this pattern as *This-be-N*, and there are instances of it in both sub-corpora. The propensity of this lexicogrammatical pattern to enact an evaluative function has been reported in numerous studies (Flowerdew, 2003c; Kanté, 2010; Schmid, 2000; Yamasaki, 2008). Again, the SFL approach here provides a theoretically motivated explanation. My findings show that in addition to inscribed prosodic value, whether infused in a shell noun like **success**, or realised as Epithet, the evaluative reading may be invoked by interpersonal selections in the surrounding co-text.

When employed in a clause, a cataphorically lexicalised shell noun usually couples with interpersonal and/or textual resources. Resources from the system of GRADUATION serve to quantify the shell noun (e.g. *several reasons*) or to grade the focus of the shell noun (e.g. a *complete surprise*) while resources from the system of ATTITUDE appraise the shell noun and its lexicalisation positively or negatively (e.g. an *unpleasant choice*). Attitudinal value may also be infused in the shell noun (e.g. **advantages, limitations**). In the corpus selections from GRADUATION are found to be more frequently employed with cataphorically lexicalised shell nouns than isolated selections from ATTITUDE. The most common evaluative meaning in terms of Attitude is that of importance, and it is realised through selections such as *key* or *main*. These findings are similar to those in other studies of academic discourse (Tadros, 1994, Hood, 2010).

When a shell noun functions in a hyper-New, it always couples with interpersonal meaning. The interpersonal meaning may draw on the resources of ATTITUDE or those of ENGAGEMENT. In the former system, attitudinal inscription may be infused in the shell noun (e.g. **paradox** [+ve Valuation]), or isolated as Epithet (e.g. these *basic* [-ve Complexity] **findings**). In the latter system, linguistic and mental shell nouns function to enact the dialogic position of the writer's voice (e.g. **inference** [HETEROGLOSSIC: EXPAND: entertain], and factual shell nouns can operate in an esphorically referring matrix to contract the dialogic space (e.g. The **point** is that [HETEROGLOSSIC: CONTRACT: reinforce: pronounce]). Interestingly, the distribution of the

resources of APPRAISAL in hyper-News with shell nouns does not seem to depend so much on whether the shell noun is factual or mental or linguistic. Rather it seems to depend on whether the lexicalisation of the shell noun is anaphoric or not. In a hyper-News, interpersonal meaning in the anaphorically referring shell nouns tends to be enacted through the resources of Attitude while in esphorically referring shell nouns, it is mostly enacted through the resources of Engagement. Notwithstanding, as noted in section 6.1.2, which describes the results for shell nouns in hyper-News, this observation is based on very few instances, and further investigation is clearly needed to ascertain if this distribution is typical or not.

6.4.5. Revisiting Schmid's three functions of shell nouns

Before concluding the summary and discussion of the results on the analysis at the discourse semantics stratum, one further aspect needs to be addressed: how does the SFL analysis relate to Schmid's textual, characterising and cognitive functions of shell nouns?

Clearly, the textual functions of anaphoric and cataphoric reference correspond closely in the two approaches. Standing at the intersection between lexical cohesion and reference, shell nouns enable the text to develop by means of encapsulation and prospection. But they also contribute to textual organisation at a higher level in macro-Themes, hyper-Themes, hyper-News and internal conjunction.

The characterising function, which earlier (chapter 3) I posited as approximating the interpersonal function, is more complex. Schmid (2000) argues that the shell noun characterises its lexicalisation, and indeed, Halliday and Matthiessen (2004) note that a fact noun 'interprets' the embedded clause. In a discourse semantics perspective, the characterising function of shell nouns is accounted for by the ideational metafunction and the interpersonal metafunction. Some shell nouns are, in fact, field-specific although they may also contribute interpersonal meaning in certain matrix constructions. Others are attitudinally charged. Furthermore, as grammatical metaphors, there is clear continuity of meaning between the nominalisation and congruent form. Selection of a shell noun, then, entails weighing up the kinds of meaning that the writer wants to make.

With regard to the cognitive function of temporary concept formation, the coupling of resources from the different strata – rank-shifting a figure to function as Qualifier to a shell noun and the function of the nominal group in the lexicogrammar – constitutes a systemic resource for instantiating in the text a nominal entity that would congruently be realised as a clause. This function depends on the ability of shell nouns to project, and the results of my study seem to corroborate the function of shell nouns for creating temporary concepts. In particular, the fact that Qualifiers to the shell nouns in my study are field-specific suggests that shell nouns are an important systemic resource for the writers to create these temporary concepts. Construed as entities, these concepts lend themselves to academic writing, in particular empirical research, for they can be assessed and measured.

This chapter has presented the results of the analysis of shell nouns at the stratum of discourse semantics. It has found that shell nouns contribute to the logogenesis of a text by functioning as a textual resource that compacts information realised elsewhere, by making it possible to construe field-related discourse entities, and by contributing to writer stance. The chapter also discussed the results with those of other studies, arguing that the SFL approach explicates many of the varied claims made about shell nouns with regard to their functions in discourse. I shall now present the conclusions in Chapter 7.

Chapter 7. Conclusions

The goal of this thesis is to develop an account of shell nouns (Schmid, 2000) grounded in Systemic Functional Linguistics (Halliday and Matthiessen, 2004; Martin, 1992; Martin & White, 2005). The theory is appropriate for the study of shell nouns because it approaches language as a stratified semiotic system network in which meaning is made simultaneously in three metafunctions (the ideational, interpersonal and textual) at different strata (e.g. the lexicogrammar and discourse semantics), and a text is seen as an instantiation of the language system network. The theory thus provides a comprehensive base for the analysis of shell nouns, which, in previous studies, have been shown to contribute evaluative functions (e.g. Schmid, 2000; Charles, 2003; 2007), textual functions across sentence boundaries (e.g. Schmid, 2000; Francis, 1994; Tadros, 1994), and cognitive functions (Schmid, 2000).

The study has taken a corpus-based approach; a parallel corpus comprising five research article submissions written by Portuguese academics from the field of economics and five published research articles on comparable topics constituted the data for the analysis. The analysis was guided by two general research questions:

1. In what ways do shell nouns contribute to ideational, interpersonal and textual meaning?
2. Are there significant differences in their use between published articles and article submissions by Portuguese academics in the field of economics?

The two questions allow for both a qualitative and quantitative analysis of shell nouns. The analysis was carried out at two strata: that of the lexicogrammar and that of discourse semantics. Using version 2.7.4 of Corpus Tool (O'Donnell, 2008), the corpus was tagged according to system networks for the language metafunctions and the two strata. Chapters 5 and 6 explored the metafunctional analysis of shell nouns in detail. I shall now present the conclusions of the study.

The main conclusion of this study is that shell nouns constitute an important systemic resource for the writers of research articles, who need to build an argument, positioning themselves and their study to convince the discourse community that their paper makes a contribution to knowledge in their disciplinary field. One way in which

shell nouns contribute to this purpose is to construe dialogic position. This is achieved predominantly through mental, linguistic and factual shell nouns. While mental and linguistic shell nouns, which are often nominalised projecting Processes tend to expand the dialogic space, factual shell nouns can be used in specific lexicogrammatical patterns to frame a proposition such that the discourse becomes contracting. In the latter case, the shell noun, composed with presuming reference, typically construes Value in an encoding relational Process clause whose Token is realised by a fact clause. This pattern is one of the defining patterns for shell nouns (Schmid, 2000). In addition, because shell nouns can compact clauses through anaphoric or cataphoric reference, the referent information can be appraised positively or negatively with resources from the system of Attitude, thus aligning the reader with particular views and disaligning them from others. The evaluative element can be isolated from or infused in the shell noun. While previous studies have identified these interpersonal functions (e.g. Francis, 1994; Charles, 2003; 2007; Kanté, 2010), the SFL approach in this study provides a theory-based explanation of how the interpersonal meaning is made.

A second major function of shell nouns is textual. Because they are lexicalised by information realised as a clause or more elsewhere in the co-text, they enable the text to unfold by compacting the clause(s) of the lexicalisation, which may appear in the text preceding the shell noun or in text that follows. These anaphoric and cataphoric uses make them a valuable resource for scaffolding a text through macro-Themes, hyper-Themes, hyper-News and internal conjunction at the stratum of discourse semantics. In hyper-Themes and hyper-News, they can couple with interpersonal meaning, and when anaphorically referring shell nouns enter into general options of conjunction, they can add interpersonal or ideational meaning.

At the stratum of the lexicogrammar, anaphorically referring nominal groups with a shell noun as Head often compose Theme, where they constitute a shared point of departure for the clause. When anaphorically referring nominal groups with shell noun as Head construe Token in a decoding relational clause whose Process is realised by a verb such as *reveal*, *confirm*, or *suggest*, they explicitly help build the writer's argument. This function is particularly prevalent with shell nouns that construe the field of research, such as **results** and **findings**.

Another conclusion is that contrary to a generalised belief that shell nouns do not contribute to field, some shell nouns cluster around the field of research and its methodologies. Furthermore, although many shell nouns do not construe a particular field, their lexicalisations do, and a shell noun can be used to project an embedded clause to create field-specific semiotic entities that are instantiated in the discourse. This represents the second defining lexicogrammatical pattern for shell nouns (Schmid, 2000). For example, coupling between the system of JUDGEMENT and a field-specific figure that construes Qualifier to the shell noun enables the academic to construe entities that form the basis of empirical research for they can be tested, quantified, or manipulated. Thus it becomes possible to set a value for “the **probability** that *w* is matched to a partner at least as good as *v*” (S_Rand), or measure “the household’s **ability** to manage its general expenses” (S_Pov).

As a nominal group, the shell noun is flexible to enter into different ranks and functions of the lexicogrammar. They can construe a wide range of Participant functions that a clause could not, and are found to do so in this study. Although they appear most frequently in relational Process clauses, they also construe Goal and indirect Participants in Circumstances quite frequently. As a ranking element of the clause, they are then open to the potential of the systems of Thematic structure and Information structure.

The capacity of shell nouns to function as described above stems from their status as semiotic abstractions, which have the ability to refer to text as fact or report. This ability, in turn, derives from their status as grammatical metaphors, more precisely, nominalisations of projecting Processes or qualities, nominalised modality, expansion nouns, fact nouns and nominalised verbs of conation. In other words, shell nouns come together as a class under the semogenic process of nominalisation. Grammatical metaphor neutralises some of the distinctions in the grammar not only between projection and expansion as argued by Halliday & Matthiessen (1999; 2004), but also between modality and projection and modality and expansion. The intersecting nature of the relation is illustrated in figure 7.1.

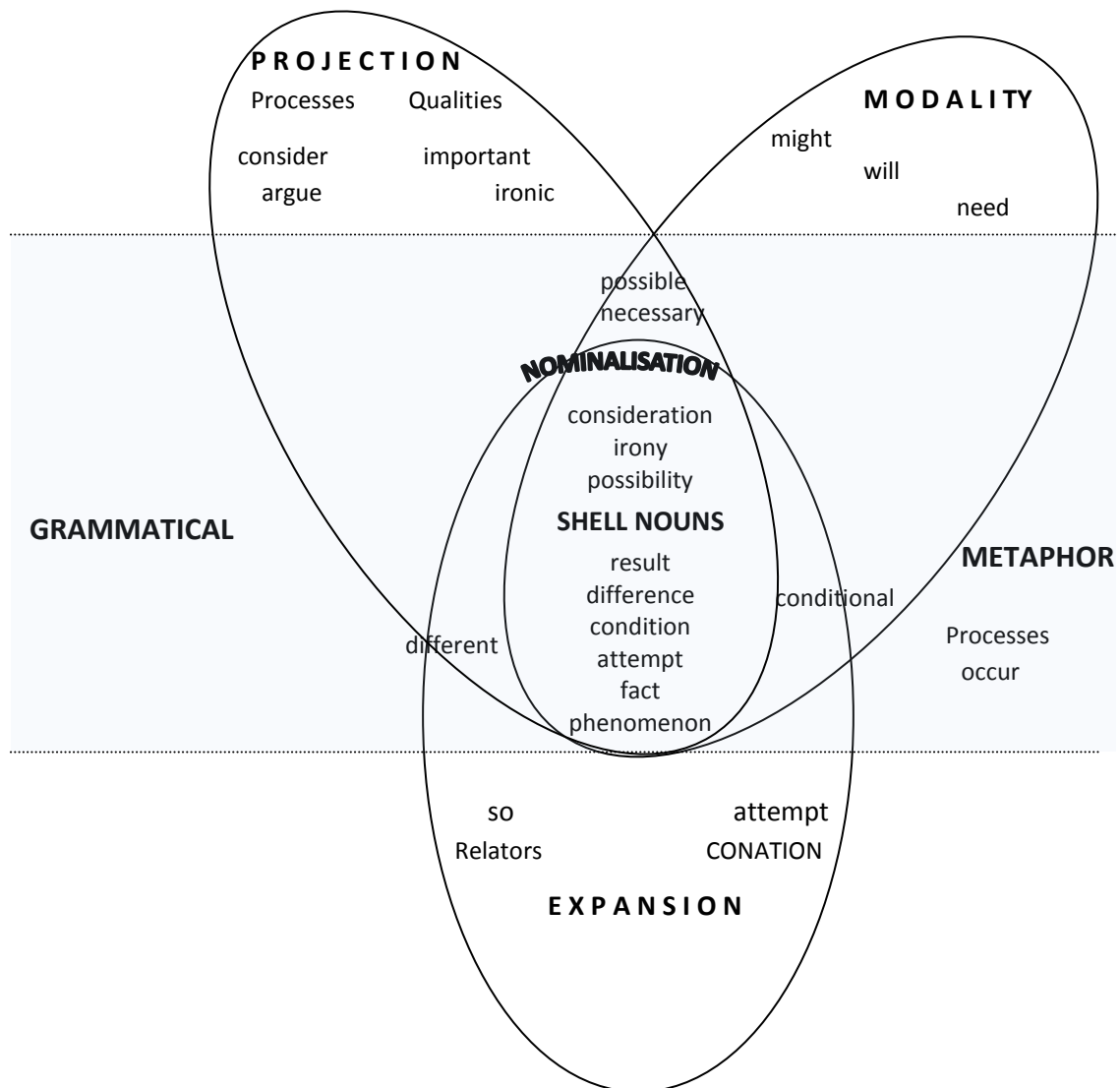


Figure 7.1 The class of shell nouns in a Systemic Functional Linguistics perspective

The study of shell nouns also aimed to determine whether there are differences in their use between the Published and Submissions sub-corpora. Although the Submissions sub-corpus uses shell nouns more frequently than the Published, both make use of the different functions that shell nouns can construe at the strata of the lexicogrammar and discourse semantics. One difference, however, stands out: dialogic position construed through shell nouns. The writers in the Submissions sub-corpus use shell nouns predominantly to expand the dialogic space, placing their claims as but one among many voices. By contrast, the writers of the Published sub-corpus use shell nouns to contract the dialogic space, and when they acknowledge other voices, it is to

attribute the propositions to others. In addition, the Submissions sub-corpus shows a higher proportion of selections from the system of Engagement than Attitude in comparison with the Published sub-corpus, which reveals the reverse trend. A further difference in the instances of Attitude is that the Published sub-corpus shows a greater tendency for negative appraisal of shell nouns than does the Submissions. When these results are taken together, the resultant texts of the Published sub-corpus could come across as more authoritative and those of the Submissions as less assertive, more tentative.

These differences in interpersonal selections affect the reading position, which could have consequences for the reviewing process. Recall the criticism levelled at a southern European academic by a peer reviewer: that the use of 'vague' 'weasel words' contributed nothing to the text and functioned only "to fill a void in the author's mind" (Lillis & Curry, 2010:150). While this study has not investigated the use of abstract words such as 'phenomena', 'approach', 'consideration' in non-shell-like uses, it has shown quite clearly that when they are used as shell nouns, they are not devoid of meaning. Ravelli (2004) has argued that non-shell-like uses of these nouns contribute to the abstraction of the text, but only further investigation can determine if and how they are related to use as a shell noun.

The study of shell nouns undertaken in this thesis opens up other possibilities for research. Given that the findings of this study are restricted to texts from economics, it remains to be seen if the uses of shell nouns in these texts are similar to those in other fields and disciplines. Moreover, further qualitative studies are necessary to identify the range of resources other than shell nouns that are used and how they interact or complement each other. This knowledge could then be used to better inform novice or EAL scholars who wish to write for academic publication. Another avenue of research would be to investigate how shell nouns are used in other registers and other modes of discourse.

This thesis has been all about shell nouns, and it has explored their functions in a Systemic Functional Linguistics perspective. It has found that they can contribute to the logogenetic development of a text by compacting propositions or proposals in a

nominal group that construes a Participant or indirect Participant. In addition to this textual function, shell nouns often enact interpersonal functions, helping to construe the stance of the writer. To conclude the thesis, it therefore seems fitting to return to an instance cited in the opening chapter and re-examine it through the SFL lens. The instance is from the film *Chisum*, and it is reproduced here for convenience. The shell nouns are in bold.

Chisum: And if one of your men cross my land or even touch one of my cows, or do anything to that store, I'm not going to the sheriff, the governor or the president of the United States. I'm coming to see you.

Murphy: Mr Chisum, that sounds like a **threat**. [Chisum punches Murphy and knocks him down]

Chisum: Wrong word. **Fact**.

At the stratum of the lexicogrammar, Murphy reconstrues Chisum's utterance as a nominal group, realised by text reference, that construes Carrier in an attributive relational Process clause, whose Attribute is realised by the shell noun **threat**. At the stratum of discourse semantics, Murphy construes the discourse as heteroglossically expanding and appraises Chisum's utterance negatively using resources from the system of ATTITUDE, the shell noun **threat** being inscribed with negative prosodic value. Chisum's response is not in words but in action: he punches Murphy and knocks him down. But this is not the end of the scene. After evaluating Murphy's assessment negatively as *wrong*, he reappraises his own utterance using a selection from the system of HETEROGLOSSIA: CONTRACT. In definitively shutting down the dialogic space with the shell noun **fact**, Chisum closes both the conversation and the scene.

While I neither pretend nor intend to be as assertive as John Wayne, aka Chisum, it is on this note that I end my thesis.

Bibliography

- Aktas, R.N. & V. Cortes 2008. Shell nouns as cohesive devices in published and ESL student writing. *Journal of English for Academic Purposes*, 7 (1): 3-14 doi: 10.1016/j.jeap.2008.02.002
- Álvarez-de-Mon y Rego, I. 2001. Encapsulation and prospection in written scientific English. *Estudios Ingleses de la Universidad Complutense*, 9: 81-101.
- Banks, D. 2005. On the historical origins of nominalized processes in scientific text. *English for Specific Purposes*, 24: 347-357 doi: 10.1016/j.esp.2004.08.002
- Barbin, B.J. 2008. Scholarly marketing publication: the American advantage? *European Business Review*, 20 (5) 370-383.
- Barros, R. 2014. Portuguese academics' attitudes to English as the Academic Lingua Franca: a case study. In K. Bennet (ed.) *The Semi-periphery of Academic writing Discourses, communities and practices*. Palgrave Macmillan 105-120
- Bennet, K. 2012. *English Academic Discourse: Hegemonic Status and Implications for Translation (with particular reference to Portuguese)*. Saarbrücken: Lambert Academic Publishing.
- Berkenkotter, C. & T.N. Huckin 1995. *Genre knowledge in disciplinary communication: cognition/culture/power*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Bertucci, R. 2006. Anáforas encapsuladoras: Uma análise em textos de opinião. *Revista Letras*, 70: 207-201. <http://repositorio.utfpr.edu.br/jspui/handle/1/545> Retrieved 12/3/2012
- Bezerra, G. 2009. O uso de rótulos em livros didáticos de história do Brasil em diferentes épocas. *Cadernos de CNLF* 18 (4): 2384-2394 http://www.filologia.org.br/xiiicnlf/XIII_CNLF_04/tomo_2/o_uso_de_rotulos_e_m_livros_didaticos_de_GABRIELI.pdf Retrieved 12/3/2012
- Biber, D. 1991. Using computer-based text corpora to analyze the referential strategies of spoken and written texts. In J. Svartvik (ed.) *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991*. Berlin, New York: Mouton de Gruyter. pp. 213-249.
- Biber, D. & S. Conrad 2009. *Register, genre, and style*. Cambridge: Cambridge University Press.
- Biber, D., S. Conrad, & R. Reppen 1998. *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Biber, D., S. Johansson, G. Leech, S. Conrad & E. Finegan 1999. *Longman grammar of spoken and written English*. Harlow: Pearson Education Limited.
- Bloor, M. & T. Bloor 1993. How Economists modify propositions. In W. Henderson, T. Dudley-Evans & R. Backhouse (eds) *Economics and Language*. London: Routledge. pp. 153-169.
- Bloor, T. & M. Bloor 1995. *The functional analysis of English: A Hallidayan approach*. London: Arnold.

- Brett, P. 1994. A genre analysis of the results section of sociology articles. *English for Specific Purposes*, 13 (1): 37-50. [doi:10.1016/0889-4906\(94\)90024-8](https://doi.org/10.1016/0889-4906(94)90024-8)
- Brock-Utne, B. 2001. The growth of English for academic communication in the Nordic countries. *International Review of Education*, 47 (3/4): 221-233. <http://link.springer.com/article/10.1023/A:1017941523866> Retrieved 9/11/2011
- Casanave, C.P. 2008. The stigmatizing effect of Goffman's stigma label: a response to John Flowerdew. *Journal of English for Academic Purposes*, 7 (4): 264-267. [doi:10.1016/j.ieap.2008.10.013](https://doi.org/10.1016/j.ieap.2008.10.013)
- Charles, M. 2003. 'This mystery ...': a corpus-based study of the use of nouns to construct stance in theses from two contrasting disciplines. *Journal of English for Academic Purposes*, 2 (4): 313-326. [doi:10.1016/S1475-1585\(03\)00048-1](https://doi.org/10.1016/S1475-1585(03)00048-1)
- Charles, M. 2007. Argument or evidence? Disciplinary variation in the use of the Noun that pattern in stance construction. *English for Specific Purposes*, 26 (2): 203-218. [doi:10.1016/j.esp.2006.08.004](https://doi.org/10.1016/j.esp.2006.08.004)
- Chevalier, G. & R. Hudson 2001. The use of intentional language in scientific articles in finance. *Journal of Economic Methodology*, 8 (2): 203-228.
- Christie, F. & C. Cléirigh 2008. On the importance of showing. In C. Wu, C.M.I.M. Matthiessen & M. Herke (eds.) *Proceedings of ISFC35: Voices around the world*. Sydney, 2008. pp. 13-18. http://www.cricyt.edu.ar/institutos/incihusa/ul/webhelp/ISFC2008_Proceedings.pdf#page=27 Retrieved 14/5/2012
- Coffin, C. & J.P. Donohue 2012. Academic literacies and Systemic Functional Linguistics: How do they relate? *Journal of English for Academic Purposes*, 11 (1): 64-75. doi: 10.1016/j.eap.2011.11.004
- Conrad, S. & D. Biber 1999. Adverbial Marking of Stance in speech and writing. In S. Hunston & G. Thompson (eds.) *Evaluation in Text: Authorial stance and the construction of discourse*. Oxford: Oxford University Press. pp. 56-73.
- Corbett, J. 2009. Theme, field and genre: thematic realisations in academic articles and their popularisations. In G. Forey & G. Thompson (eds.) *Text Type and Texture*. London: Equinox. pp. 70-93.
- Cortes, V. 2004. Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for specific Purposes*, 23 (4): 397-423. [doi:10.1016/j.esp.2003.12.001](https://doi.org/10.1016/j.esp.2003.12.001)
- Coupé, T. 2003. Revealed performances: Worldwide Rankings of economists and economics departments 1990-2000. *Journal of the European Economic Association*, 1 (6): 1309-1345. doi: 10.1162/154247603322752557
- Cummings, M. 2005. The role of Theme and Rheme in contrasting methods of organization in texts. In C. Butler, M. de los Ángeles Gómez-González & S.M. Doval-Suárez (eds.) *The Dynamics of Language Use*. Amsterdam/Philadelphia: John Benjamins Publishing Co. pp. 129-154.

- Curry, M.J. & T. Lillis 2004. Multilingual scholars and the imperative to publish in English: Negotiating interests, demands and rewards. *TESOL Quarterly*, 38 (4): 663-688. <http://www.jstor.org/stable/3588284> Retrieved 14/5/2012
- Dahl, T. 2004. Textual metadiscourse in research articles: a marker of national culture or of academic discipline? *Journal of Pragmatics*, 36: 1807-1825. doi: 10.1016/j.pragma.2004.05.004
- Davis, W. 2007. Economists' opinions of economists' work. *The American Journal of Economics and Sociology*, 66 (2): 267-288.
- Diamond, A.M. (Jr) 1989. The Core journals of economics. *Journal Citation Studies*, 50 (1). Reprinted in *Essays of an Information Scientist: Creativity, Delayed Recognition, and other Essays* Vol.12. pp. 3-11.
- Dudley-Evans, T. 1994. Genre analysis: an approach to text analysis for ESP. In M. Coulthard (ed.) *Advances in Written Text Analysis*. London, New York: Routledge. pp. 219-228.
- Dueñas, M.P. 2003-2004. Analysing stance in American and Spanish business management RAs : the case of sentence-initial 'retrospective labels'. *Journal of English Studies*, 4: 137-154. <https://publicaciones.unirioja.es/ojs/index.php/jes/article/view/92> Retrieved 21/7/2013
- Dueñas, M.P. 2012. Getting research published internationally in English: An ethnographic account of a team of Finance Spanish scholars' struggles. *Ibérica: Revista de la Asociación Europea de Lenguas para Fines Específicos*, 24: 139-155 http://www.aelfe.org/documents/12_24_Mur.pdf Retrieved 10/2/2014
- Durán, J. 2012. An SFL analysis of existential clauses in research articles. In J.S. Knox (ed.) *To boldly proceed. Papers from the 39th International Systemic Functional Linguistics*. pp. 111- 116.
- Engwall, G. 1991. Comments on 'Modern Swedish text corpora'. In J. Svartvik (ed.) *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991*. Berlin, New York: Mouton de Gruyter. pp. 164-169.
- Ferguson, G., C. Perez-Llantada & R. Plo 2011. English as a language of international scientific publication: a study of attitudes. *World Englishes*, 30 (1): 41-59.
- Flowerdew, J. 2000. Discourse Community, Legitimate Peripheral Participation, and the Nonnative-English-Speaking Scholar. *TESOL Quarterly*, 34 (1): 127-150. <http://www.jstor.org/stable/3588099> Retrieved 10/11/2011
- Flowerdew, J. 2001. Attitudes of journal editors to nonnative speaker contributions. *TESOL Quarterly*, 35 (1): 121-150. doi: 10.2307/3587862
- Flowerdew, J. 2003a. Signalling nouns in discourse. *English for Specific Purposes*, 22 (4): 329-346. doi:10.1016/S0889-4906(02)00017-0
- Flowerdew, J. 2003b. Register-specificity of signalling nouns in discourse. In P. Leistyna & C.F. Meyer (eds.) *Corpus analysis: Language structure and language use*. Amsterdam, New York: Rodopi. pp. 35-46.

- Flowerdew, J. 2008. Scholarly writers who use English as an Additional Language: What can Goffman's "Stigma" tell us? *Journal of English for Academic Purposes*, 7: 77-86. doi:10.1016/j.jeap.2008.03.002
- Flowerdew, J. 2009. Goffman's stigma and EAL writers: The author responds to Casanave. *Journal of English for Academic Purposes*, 8: 69-72. doi:10.1016/j.jeap.2009.01.001
- Flowerdew, J. 2009. Use of signalling nouns in a learner corpus. In J. Flowerdew & M. Mahlberg (eds.) *Lexical cohesion and corpus linguistics*. Amsterdam/Philadelphia: John Benjamins Pub. Company. pp 85-102. First published in *International Journal of Corpus Linguistics*, 11: 3 (2006).
- Flowerdew, L. 2003c. A Combined corpus and systemic-functional analysis of the problem-solution pattern in a student and professional corpus of technical writing. *TESOL Quarterly*, 37 (3): 489-511. doi/10.2307/3588401/pdf
- Flowerdew, L. 2004. An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP: countering criticisms against corpus-based methodologies. *English for Specific Purposes*, 24 (3): 321-322 doi:10.1016/j.esp.2004.09.002
- Fontaine, L. & Y. Kodratoff 2003. The role of thematic and concept texture in scientific text: comparing native and non-native writers of English. The English version of *Comparison du rôle de la progression thématique et de la texture conceptuelle chez des scientifiques anglophones et francophones s'exprimant en Anglais*. *Special Issue Asp*, GERAS. <http://www.cardiff.ac.uk/encap/fontaine/thematic-concept.pdf> Retrieved 12/2/2011
- Francis, G. 1994. Labelling discourse: an aspect of nominal-group lexical cohesion. In M. Coulthard (ed.) *Advances in Written Text Analysis*. London, New York: Routledge. pp. 83-101.
- Freedman, A. & P. Medway (eds.) 1994 *Genre and the new rhetoric*. London: Taylor & Francis.
- Frey, B.S. & R. Eichenberger 1997. Economists: First semester, High Flyers and UFOs. In P. van Bergeijk et al. (eds.) *Economic Science and Practice: the roles of academic economists and policy makers*. Cheltenham, UK: Edward Elgar Publishing. pp. 15-48.
- Fries, P.H. 1994. On Theme, Rheme and discourse goals. In M. Coulthard (ed.) *Advances in written text analysis*. London, New York: Routledge. pp. 229-249.
- Fries, P.H. 1995a Themes, Methods of Development, and Texts. In R. Hasan & P.H. Fries (eds.) *On Subject and Theme: A discourse functional perspective*. Amsterdam: John Benjamins Publishing Company. pp. 317-360.
- Fries, P.H. 1995b A personal view of Theme. In M. Ghadessy (ed.) *Thematic Development in English Texts*. London & New York: Pinter Publishers. pp. 1-19.
- Fries, P.H. 2009. The textual metafunction as a site for a discussion of the goals of linguistics and techniques of linguistic analysis. In G. Forey and G. Thompson (eds.) *Text Type and Texture*. London: Equinox. pp. 8-44.

- Gazzola, M. 2010. The linguistic effects of academic performance indicators: the case of Italy. *ELF Working Paper #2* Geneva: Observatoire Économie Langues Formation.
- Gellerstam, M. 1991. Modern Swedish text corpora. In J. Svartvik (ed.) *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991*. Berlin, New York: Mouton de Gruyter. pp. 149-163.
- Ghadessey, M. 1995. Thematic development and its relationship to registers and genres. In M. Ghadessy (ed.) *Thematic development in English texts*. London & New York: Pinter Publishers. pp. 129-146.
- Gledhill, C. 2000. The discourse function of collocation in research article introductions. *English for Specific Purposes*, 19 (2): 115-135. doi.org/10.1016/S0889-4906(98)00015-05
- Godsen, H. 2009. Thematic content in peer reviews of scientific papers. In G. Forey & G. Thompson (eds.) *Text Type and Texture*. London: Equinox. pp. 94-107.
- Goldschmidt, N. & B. Szmrecsanyi 2007. What do Economists Talk About? A Linguistic Analysis of Published Writing in Economic Journals. *The American Journal of Economics and Sociology*, 66 (2): 335-378.
- Gómez-González, M.A. 2001. *The Theme-Topic Interface: Evidence from English* Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Gray, B. 2010. On the use of demonstrative pronouns and determiners as cohesive devices: A focus on sentence-initial *this/these* in academic prose. *Journal of English for Academic Purposes*, 9 (3): 167-183. doi:10.1016/j.jeap.2009.11.003
- Gray, B. & V. Cortes 2011. Perception vs. Evidence: An analysis of *this* and *these* in academic prose. *Journal of English for Academic Purposes*, 30 (1): 31-43. doi:10.1016/j.esp.2010.06.004
- Guillén Galve, I. 1998. The textual interplay of grammatical metaphor on the nominalization occurring in written medical English. *Journal of Pragmatics*, 30 (3): 363-385. doi:10.1016/S0378-2166(98)00002-2
- Halliday, M.A.K. 1993b. The model. In Halliday, M.A.K. & J.R. Martin (eds.) *Reading Science*. London: The Falmer Press. pp. 22-50.
- Halliday, M.A.K. 1988/1993. On the language of physical science. In M.A.K. Halliday & J.R. Martin (eds.) *Writing Science: Literacy and discursive power*. London: The Falmer Press Pp 54-68 Reprinted from M. Ghadessy (ed.) 1988. *Registers of written English*. London/New York: Pinter Publishers. Chapter 10
- Halliday, M.A.K. 1989/1993a. Some grammatical problems in scientific English. In M.A.K. Halliday & J.R. Martin (eds.) *Reading Science*. London: The Falmer Press. pp. 69-85. Reprinted from *Australian review of Applied Linguistics: Genre and Systemic Functional Studies*, 5 (6).
- Halliday, M.A.K. 1990/2002. The construction of knowledge and value in the grammar of scientific discourse: with reference to Charles Darwin's *The Origin of Species*. In J. Webster (ed.) *Linguistic studies of text and discourse: The collected works of M.A.K. Halliday. Vol. 2*. London: Continuum. pp. 168-192.

- Halliday, M.A.K. 1991. Language as system and language as instance: The corpus as a theoretical construct. In J. Svartvik (ed.) *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991*. Berlin, New York: Mouton de Gruyter. pp. 61-77.
- Halliday, M.A.K. 1998. Things and relations: Regrammaticising experience as technical knowledge. In J.R. Martin & R. Veel (eds.) *Reading science: Critical and functional perspectives on discourses of science*. London & New York: Routledge. pp. 185-235.
- Halliday, M.A.K. & R. Hasan 1976. *Cohesion in English*. London, New York: Longman.
- Halliday, M.A.K & J.R. Martin 1993. *Writing science: Literacy and discursive power*. London: The Falmer Press.
- Halliday, M.A.K. 1977/2002. Text as semantic choice in social contexts. In J. Webster (ed.) *Linguistic studies of text and discourse: The collected works of M.A.K. Halliday. Vol. 2*. London: Continuum. pp. 22-81.
- Halliday, M.A.K. & Z.L. James 1993. A quantitative study of polarity and primary tense in the English finite clause. In J. M. Sinclair, M. Hoey and G. Fox (eds.) *Techniques of description: Spoken and written discourse*. New York: Routledge. pp. 32-66.
- Halliday, M.A.K. & C.M.I.M. Matthiessen 1999. *Construing experience through meaning: A language-based approach to cognition*. London, New York: Continuum
- Halliday, M.A.K. and C.M.I.M. Matthiessen 2004. *An Introduction to Functional Grammar* 3rd edition. London: Arnold.
- Hamel, R.E. 2007. The dominance of English in the international scientific periodical literature and the future of language use in science. *AILA Review*, 20: 53-71. doi: 10.1075/aila.20.06ham
- Hao, J. & S.L. Humphrey 2012. Burnishing and tarnishing in academic literacy. In J.S. Knox (ed.) *To boldly proceed. Papers from the 39th International Systemic Functional Linguistics Congress*. pp 15-20.
- Hasan, R. 1985. Meaning, Context and Text - Fifty Years after Malinowski. In J.D. Benson & Greaves, William S. (eds.) *Systemic perspectives on discourse. Vol. 1*. Norwood, N.J.: Ablex Publishing Corporation. pp. 16-49.
- Hasselgård, H. 2012. Fact, ideas, questions, problems, and issues in advanced learners' English. *Nordic Journal of English Studies*, 11 (1): 22-54. <http://ojs.ub.gu.se/ojs/index.php/njes/article/view/1036> Retrieved 9/6/2013
- Henshall, A.C. 2012. English as an international language and language policies in economics journals. *Anglo-Saxónica, Series III. (4)*: 131-162.
- Hirsh, D. 2010. *Academic vocabulary in context*. Bern, Switzerland: Peter Lang AG
- Hoey, M. 1993. A common signal in discourse: How the word *reason* is used in texts. In J.M. Sinclair, M. Hoey & G. Fox (eds.) *Techniques of Description: Spoken and written discourse*. New York: Routledge. pp. 67-82.

- Hoey, M. 1999. Persuasive rhetoric in linguistics: A stylistic study of some features of the language of Noam Chomsky. In S. Hunston & G. Thompson (eds.) *Evaluation in text: Authorial stance and the construction of discourse*. Oxford: Oxford University Press. pp. 28-37.
- Holmes, R. 1997. Genre analysis and the social sciences: An investigation of the structure of research article discussion sections in three disciplines. *English for Specific Purposes*, 16 (4): 321-327. [doi:10.1016/S0889-4906\(96\)0038-5](https://doi.org/10.1016/S0889-4906(96)0038-5)
- Hood, S. 2006. The persuasive power of prosodies: Radiating values in academic writing. *Journal of English for Academic Purposes*, 5 (1): 37-49. [doi:10.1016/j.jeap.2005.11.001](https://doi.org/10.1016/j.jeap.2005.11.001)
- Hood, S. 2008. Summary writing in academic contexts: Implicating meaning in processes of change. *Linguistics and Education*, 19: 351-365. doi: 10.1016/j.linged.2008.06.003
- Hood, S. 2010. *Appraising Research: Evaluation in academic writing*. Houndsmills: Palgrave Macmillan.
- Hopkins, A. & T. Dudley-Evans 1988. A genre-based investigation of the discussion sections in articles and dissertations. *English for Specific Purposes*, 7 (2): 113-121. [doi:10.1016/0889-4906\(88\)90029-4](https://doi.org/10.1016/0889-4906(88)90029-4)
- Hunston, S. 1999. Evaluation and the planes of discourse: Status and value in persuasive texts. In S. Hunston & G. Thompson (eds.) *Evaluation in text: Authorial stance and the construction of discourse*. Oxford: Oxford University Press. pp. 176-207.
- Hunston, S. 2011. *Corpus approaches to evaluation: Phraseology and evaluative language*. New York, London: Routledge.
- Hunston, S. & G. Francis 2000. *Pattern grammar: A corpus-driven approach to the lexical grammar of English*. Amsterdam, Philadelphia: John Benjamins B.V.
- Hunston, S. & J. Sinclair 1999. A local grammar of evaluation. In S. Hunston & G. Thompson (eds.) *Evaluation in Text: Authorial stance and the construction of discourse*. Oxford: Oxford University Press. pp 74-101 .
- Hyland, K. 2000. *Disciplinary discourses: Social interactions in academic writing*. Harlow, England: Pearson Education Limited.
- Hyon, S. 1996. Genre in three traditions: Implications for ESL. *TESOL Quarterly*, 30 (4): 693-722 <http://www.jstor.org/stable/3587930>. Retrieved 4/3/2010
- Ivanic, R. 1991. Nouns in search of a context: A study of nouns with both open- and closed-system characteristics. *International Review of applied Linguistics in language Teaching*, 29 (2): 93-114. doi: 10.1515/iral.1991.29.2.93
- Jordan, M.P. 1985. Non-Thematic re-entry: An introduction to and extension of the system of nominal group reference/substitution in everyday English use. In J.D. Benson & W.S. Greaves (eds.) *Systemic perspectives on discourse. Vol. 1*. Norwood, N.J.: Ablex Publishing Corporation. pp. 322-332.
- Kanté, I. 2010. Head nouns as modal stance markers –academic texts vs. legal texts. *Lidil*, 41. <http://lidil.revues.org/index3022.html>. Retrieved 5/7/2012

- Klamer, A. & P. Hendrik 2002. Attention and the art of scientific publishing. *Journal of Economic Methodology*, 9 (3): 289-316.
- Klamer, A. & D.N. McCloskey 1988. Economics in the human conversation. In A. Klamer, D.N. McCloskey & R.M. Solow (eds.) *The consequences of economic rhetoric*. Cambridge: Cambridge University Press. pp 3-20.
- Koch, I.G. 2006. Rotulação: uma estratégia textual de construção do sentido. Labeling: a textual strategy of meaning construction. *Calidescópico*, 4 (2): 85-89. <http://revistas.unisinos.br/index.php/calidoscopio/article/view/5990> Retrieved 12/3/2013
- Labarde, D.N., R.D. Tollison & G. Karahan 2002. Quality control in Economics. *Kyklos*, 55 (3): 315-333.
- Leech, G. 1991. Corpora and theories of linguistic performance. In J. Svartvik (ed.) *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991*. Berlin, New York: Mouton de Gruyter. pp. 105-122.
- Lemke, J.L. 1985. Ideology, Intertextuality, and the Notion of Register. In J.D. Benson, & W.S. Greaves (eds.) *Systemic perspectives on discourse*. Vol. 1. Norwood, N.J.: Ablex Publishing Corporation. pp 275-294.
- Lillis, T & M.J. Curry 2006. Professional academic writing by multilingual scholars: Interactions with literacy brokers in the production of English medium texts. *Written Communication*, 23 (1): 3-35. 10.1177/0741088305283754 <http://wcx.sagepub.com> Retrieved 14/5/2012
- Lillis, T & M.J. Curry 2010. *Academic Writing in a Global Context: The Politics and Practices of Publishing in English*. London: Routledge.
- Lim, J. 2006. Method sections of management research articles: A pedagogically motivated qualitative study. *English for Specific Purposes*, 25: 282-389. doi: 10.1016/j.esp.2005.07.001
- Lin, L. & S. Evans 2012. Structural patterns in empirical research articles: A cross-disciplinary study. *English for Specific Purposes*, 31 (3): 150-160. [doi:10.1016/j.esp.2011.10.002](https://doi.org/10.1016/j.esp.2011.10.002)
- Lyytinen, K., R. Baskervill, J. Iavari, & D. Te'eni 2007. Why the old world cannot publish? *European Journal of Information Systems*, 16 (4): 317-326.
- Mair, C. 1991. Comments on Chafe's The importance of corpus linguistics to understanding the nature of language. In J. Svartvik (ed.) *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991*. Berlin, New York: Mouton de Gruyter. pp. 98-103.
- Martin, J. R. 1985. Process and text: Two aspects of human semiosis. In J.D. Benson & W.S. Greaves (eds.) *Systemic perspectives on discourse*. Vol. 1, *Selected theoretical papers from the 9th International systemic Workshop*. Norward, New Jersey: Ablex Publishing Corporation. pp. 248-274.

- Martin, J.R. 1991. Nominalization in science and humanities: Distilling knowledge and scaffolding text. In E. Ventola (ed.) *Functional and systemic linguistics: approaches and uses*. Berlin, New York: Mouton de Gruyter. pp. 307-337.
- Martin, J.R. 1992. *English text: System and structure*. Philadelphia: John Benjamins Publishing Company.
- Martin, J.R. 1995a. More than what the message is about: English Theme. In M. Ghadessy (ed.) *Thematic development in English texts*. London & New York: Pinter Publishers. pp. 223-258.
- Martin, J. R. 1995b. Text and clause: Fractal resonance. *Text*, 15 (1): 5-42.
- Martin, J.R. 2000. Beyond exchange: APPRAISAL systems in English. In S. Hunston & G. Thompson *Evaluation in text* (eds.) Oxford, Oxford University Press. pp. 142-175.
- Martin, J.R. 2006. Genre and field: Social processes and knowledge structures in systemic functional semiotics. *Proceedings from the 33rd International Systemic Functional Congress*.
- Martin, J.R. 2008. Tenderness: Realisation and instantiation in a Botswanan town. In N. Norgaard (ed.) *Systemic Functional Linguistics in Use. Odense Working Papers in Language and Communication*, 29: 30-62.
- Martin, J. R. 2010. Semantic variation - Modelling realisation, instantiation and individuation in social semiosis. In M. Bednarek & J.R. Martin (eds.) *New discourse on language: Functional perspectives on multimodality, identity and affiliation*. London, New York: Continuum. pp. 1-34.
- Martin, J.R. & D. Rose 2003. *Working with discourse: Meaning beyond the clause*. London, New York: Continuum.
- Martin, J.R. & D. Rose 2008. *Genre relations: Mapping culture*. London/Oakville: Equinox Publishing Ltd.
- Martin, J. R. & R. Veel (eds.) 1998. *Reading science: Critical and functional perspectives on discourses of science*. London & New York: Routledge.
- Martin, J.R. & P.R.R. White 2005. *The language of evaluation: Appraisal in English*. Houndmills & New York: Palgrave Macmillan.
- Martín, P. 2003. A genre analysis of English and Spanish research paper abstracts in experimental social sciences. *English for Specific Purposes*, 22 (1): 25-43. [http://dx.doi.org/10.1016/S0889.4906\(01\)00033-3](http://dx.doi.org/10.1016/S0889.4906(01)00033-3)
- Martínez, I. 2003. Aspects of theme in the method and discussion sections of biology journal articles in English. *Journal of English for Specific Purposes*, 2: 103-123. doi: 10.1016/S1475-1585(03)00003-1
- Martínez, I. 2001. Impersonality in the research article as revealed by analysis of the transitivity structure. *Journal of English for Specific Purposes*, 20 (3): 227-247 [doi:10.1016/S0889-4906\(00\)00013-2](http://dx.doi.org/10.1016/S0889-4906(00)00013-2)

- Martins, A.P. 2009. O uso de sintagmas nominais como rótulos em entrevistas jornalísticas. *Cadernos de CNLF*, 18 (4): 2395-2402. http://www.filologia.org.br/xiiicnlf/XIII_CNLF_04/tomo_2/o_uso_de_sintagmas_nominais_como_rotulos_em_ANA.pdf Retrieved 12/3/2013
- Martins, A.P. & V.L. Silva 2008. O uso de SN's definidos vs. Demonstrativos como rótulos em entrevistas jornalísticas. *Revista Linguística*, 4 (1) http://www.letras.ufrj.br/poslinguistica/revistalinguistica/wp-content/uploads/2012/09/Artigo_Paredes.pdf Retrieved 12/3/2013
- Mattheoudakis, M. & A.-M. Hatzitheodorou 2009. Discourse organising nouns in Greek and American university students' argumentative essays: A contrastive study' Paper presented at 19th International Symposium on Theoretical and Applied Linguistics, 3-5 April 2009. Aristotle University of Thessaloniki. Thessaloniki, Greece. http://www.enl.auth.gr/symposium19/19thpapers/032_Mattheoudakis_Hatzitheodorou.pdf
- Matthiessen, C. 1995. THEME as an enabling resource in ideational 'knowledge' construction. In M. Ghadessy (ed.) *Thematic development in English texts*. London & New York: Pinter Publishers. pp 20-54.
- McAfee, R. P. 2010. Edifying Editing. *The American Economist*, 55 (1): 1-8. <http://www.mcafee.cc/Papers/PDF/EditorExperiences.pdf> Retrieved 11/12/2011
- McLaglen, A. 1970. Chisum. <http://www.imdb.com/title/tt0065547/quotes> Retrieved 21/9/2014
- Moore, N. 2011. Modelling the flow of discourse in a corpus of written academic English. <http://www.birmingham.ac.uk/documents/college-artslaw/corpus/conference-archives/2011/Paper-11.pdf> Retrieved 1/27/2012
- Moreno, A.I. 1997. Genre constraints across languages: causal metatext in Spanish and English RAs. *English for Specific Purposes*, 16 (3): 161-179. [doi:10.1016/S0889-4906\(96\)00023-3](https://doi.org/10.1016/S0889-4906(96)00023-3)
- Moreno, A.I. 2004. Retrospective labelling in premise-conclusion metatext: an English-Spanish contrastive study of research articles on business and economics. *Journal of English for Academic Purposes*, 3 (4): 321-339. [doi:10.1016/j.jeap.2004.07.005](https://doi.org/10.1016/j.jeap.2004.07.005)
- Moreno, A.I., J. Rey-Rocha, S. Burgess, I. López-Navarro & I. Sachdev 2012. Spanish researchers' perceived difficulty writing research articles for English-medium journals: the impact of proficiency in English vs publication experience. *Iberica*, 24: 157-184. http://www.aelfe.org/documents/13_24_Moreno.pdf Retrieved 9/9/2014
- Morgan, M.S. 2001. Models, stories and the economic world. *Journal of Economic Methodology*, 8 (3): 361-384.
- Myers, G. 1994. Narratives of science and nature in popularizing molecular genetics. In M. Coulthard (ed.) *Advances in written text analysis*. London, New York: Routledge. pp. 179-190.

- Nkemleke, D. 2010. Camaroonian and foreign scholars' discourse: The rhetoric of conference abstracts. *World Englishes*, 29 (2): 173-191.
- O'Donnell, M. 2008. Demonstration of the UAM CorpusTool for text and image annotation. *Proceedings of the ACL-08: HLT Demo Session (Companion Volume)*, Columbus, Ohio, June 2008. Association for Computational Linguistics. pp. 13–16.
- Parada, F. 2009. Psychology publishing in Portugal: Hardship and disappointment follows expectations. *Psychology Science Quarterly*, 51 (1): 42-54. http://journaldatabase.info/articles/psychology_publishing_portugal.html Retrieved 4/3/2011
- Parkinson, J. 2013. Adopting academic values: Student use of that-complement clauses in academic writing. *System*. doi.org/10.1016/j.system.2013.03.002
- Parks, R.P. 2002. The Faustian grip of academic publishing. *Journal of Economic Methodology*, 9 (2): 317-335.
- Perez-Llantada, C., R. Plo & G. Ferguson 2011. “You don’t say what you know, only what you can”: The perceptions and practices of senior Spanish academics regarding research dissemination in English. *English for Specific Purposes*, 30.1: 18-30. doi:10.1016/j.esp.2010.05.001
- Pettinari, C. 1985. Comparison of the production of surgical reports by native and non-native speaking surgeons. In J.D. Benson & W.S. Greaves (eds.) *Systemic perspectives on discourse. Vol. 2*. Norwood, N.J.: Ablex Publishing Corporation. pp. 187-203.
- Price, C. & T. King 2008. Context inducing nouns. *Proceedings of the workshop on Knowledge and Reasoning for Answering questions*. Manchester. <http://dl.acm.org/citation.cfm?doid=1642004.1642006> Retrieved 8/4/2012
- Ravelli, L.J. 1995. A Dynamic perspective: Implications for metafunctional interaction and an understanding of Theme. In R. Hasan & P. Fries (eds.) *On Subject and Theme: A discourse functional perspective*. Amsterdam: John Benjamins Publishing Company. pp 187-234.
- Ravelli, L. 2004. Familiar territory, shifting grounds: Aspects of organization in written texts. In L. Ravelli & R. Ellis (eds.) *Analysing academic writing: Contextualised frameworks*. London: Continuum. pp. 104-130.
- Ribeira, J. 2007. Text deixis in narrative sequences. *International Journal of English Studies*, 7 (1): 149-168. doi.org/10.6018/48971
- Rose, D. 1998. Science discourse and industrial hierarchy. in J.R. Martin & R. Veel (eds.) *Reading science: Critical and functional perspectives on discourses of science*. London & New York: Routledge. pp. 236-265.
- Samraj, B. 2005. An exploration of a genre set: Research article abstracts and introductions in two disciplines. *English for Specific Purposes*, 24 (2): 141-156. doi.org/10.1016/j.esp.2002.10.001
- Schmid, H.-J. 2000. *English Abstract Nouns as Conceptual Shell: From corpus to cognition*. Berlin/New York: Mouton.

- Schmid, H.-J. 2001. Presupposition can be a bluff: How abstract nouns can be used as presupposition triggers. *Journal of Pragmatics*, 33 (10): 1529-1552 [doi:10.1016/S0378-2166\(01\)00027-3](https://doi.org/10.1016/S0378-2166(01)00027-3)
- Shaw, P. 2000. Towards classifying the arguments in research genres. In A. Trosborg (ed.) *Analysing Professional Genres*. Amsterdam: John Benjamins Publishing Co. pp. 41-56.
- Sinclair, J.M. 1993. Written discourse structure. In J.M. Sinclair, M. Hoey & G. Fox (eds.) *Techniques of description: Spoken and written discourse*. New York: Routledge. pp. 6-31.
- Smith, E.L. 1985. Functional Types of Scientific Prose. In J.D. Benson & W.S. Greaves, (eds.) *Systemic perspectives on discourse. Vol. 2*. Norwood, N.J.: Ablex Publishing Corporation. pp. 241-257.
- Smyers, P. & B. Levering 2000. Educational research: Language and content. Lessons in publication Policies from the Low Countries. *British Journal of Educational Studies*, 48 (1): 70-81.
- Stewart, D.M. 2008. Academic publishing in marketing: best and worst practices. *European Business Review*, 20 (5): 421-433.
- Straker, J. 2010. The Use of academic vocabulary in the coursework of international students: Some pedagogic considerations. In G. Blue (ed.) *Developing Academic Literacy*. Oxford: Peter Lang. pp. 53-69.
- Svartvik, J. (ed) 1991. *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991*. Berlin, New York: Mouton de Gruyter
- Svensson, G. & T. Helgerson 2008. Scientific identity of "top" research journals in the broader discipline of marketing. *European Business Review*, 20 (5): 384-400.
- Swales, J. 1990. *Genre Analysis*. Cambridge: Cambridge University Press.
- Swales, J. 1997. English as Tyrannosaurus rex. *World Englishes*, 16 (3): 373-382.
- Swales, J. 2005. Attended and unattended "this" in academic writing: A long and unfinished story. *ESP Malaysia*, 11: 1-15.
- Tadros, A. 1994. Predictive categories in expository text. In M. Coulthard (ed.) *Advances in written text analysis*. London, New York: Routledge. pp. 69-82.
- Teubert, W. & A. Čermáková 2004/2007. *Corpus linguistics: A short introduction* London: Continuum
- Thompson, G. & J. Zhou 1999. Evaluation and organization in text: the structuring role of evaluative disjuncts. In S. Hunston & G. Thompson (eds.) *Evaluation in text: Authorial stance and the construction of discourse*. Oxford: Oxford University Press. pp. 121- 141.
- Thompson, G. & S. Hunston 1999. Evaluation: An introduction. In S. Hunston & G. Thompson (eds.) *Evaluation in text: Authorial stance and the construction of discourse*. Oxford: Oxford University Press. pp. 1-27.

- Thompson, G. & S. Thompson 2009. Theme, Subject and the unfolding of text. In G. Forey and G. Thompson (eds.) *Text type and texture*. London: Equinox. pp. 45-69.
- Trosborg, A. 2000. Introduction to *Analysing Professional Genres*. Amsterdam: John Benjamins Publishing Co. pp. vii-xvi.
- Turner, J. 2010. Supporting academic literacy: Issues of proofreading and language proficiency. In G. Blue (ed.) *Developing Academic Literacy*. Oxford: Peter Lang. pp. 39-51.
- Valero-Garcés, C. 1996. Contrastive ESP rhetoric: metatext in Spanish-English economics texts. *English for Specific Purposes*, 15 (4): 279-294. [doi:10.1016/S0889-4906\(96\)00013-0](https://doi.org/10.1016/S0889-4906(96)00013-0)
- Van Bergeijk, P.A.G., A.L. Bovenberg, E.E.C. van Damme, & J. van Sinderen 1997. Introduction. In P.A.G. van Bergeijk et al. (eds.) *Economic science and practice: The roles of academic economists and policy makers*. Cheltenham, UK: Edward Elgar Publishing. pp. 1-12.
- Veel, R. 1998. The greening of school science: Ecogenesis in secondary classrooms. In J.R. Martin & R. Veel (eds.) *Reading science: Critical and functional perspectives on discourses of science*. London & New York: Routledge. pp. 114-151.
- Ventola, E. 1995. Thematic development and translation. In M. Ghadessy (ed.) *Thematic development in English texts*. London & New York: Pinter Publishers. pp. 85-104.
- Ventola, E. 1998. Interpersonal choices in academic work. In A. Sánchez-Macarro, & R. Carter (eds.) *Linguistic choices across genres*. Amsterdam/Philadelphia: John Benjamins Publishing Company. pp. 117-136.
- Ventola, E. & A. Mauranen 1991. Non-native writing and native revising of scientific articles. In E. Ventola (ed.) *Functional and systemic linguistics: Approaches and uses*. Berlin, New York: Mouton de Gruyter. pp. 457-492.
- White, P.R.R. 1998. Extended reality: proto-nouns and the vernacular: Distinguishing the technological from the scientific. In J.R. Martin & R. Veel (eds) *Reading science: Critical and functional perspectives on discourses of science*. London & New York: Routledge. pp 266-296.
- Whittaker, R. 1995. Thematic development in academic and non-academic texts. In M. Ghadessy (ed.) *Thematic development in English texts*. London & New York: Pinter Publishers. pp. 105-128.
- Willis, D. 1993. Grammar and lexis: some pedagogical implications. In J.M. Sinclair, M. Hoey & G. Fox (eds.) *Techniques of description: Spoken and written discourse*. New York: Routledge. pp. 83-93.
- Winter, E.O. 1977. A Clause-relational approach to English texts: A Study of some predictive lexical items in written discourse. *Instructional Science*, 6: 1-92. http://download.springer.com/static/pdf/479/art%253A10.1007%252FBF00125597.pdf?auth66=1425471287_83b356db8c910985d95ca66ae9f0c681&ext=.pdf Retrieved 10/2/2011

- Winter, E.O. 1992. The notion of unspecific vs specific as one way of analysing the information of a fund-raising letter. In W.C. Mann & S.A. Thomopson (eds.) *Discourse descriptions: Diverse linguistic analyses of a fund-raising text*. Amsterdam: John Benjamins Publishing Company. pp 131-170.
- Yamasaki, N. 2008. Collocations and colligations associated with discourse functions of unspecific anaphoric nouns. *International Journal of Corpus Linguistics*, 13 (1): 75-98. doi: 10.1075/ijcl.13.1.05yam
- Yang, R.& D. Allison 2003. Research articles in applied linguistics: Moving from results to conclusions. *English for Specific Purposes*, 22 (4): 365-385. doi: 10.1016/S0889-4906(02)00026-1

Appendices

List of Appendix tables

Appendix A	General analysis at the stratum of the lexico grammar.....	329
A.1	Distribution of features tagged with ideational function as type and rank in Published and Submissions.....	329
A.2	Distribution of features tagged with a transitivity function at clause rank in Published and Submissions.....	329
A.3	Distribution of shell nouns tagged with interpersonal function at clause rank in Published and Submissions	330
A.4	Distribution of features tagged with textual function at clause rank in Published and Submissions.....	331
A.5	Distribution of Participants in Published and Submissions.....	332
A.6	Distribution of Process types in Published and Submissions	333
A.7	Distribution of clause types containing shell noun in Published and Submissions.....	335
Appendix B	Textual and Interpersonal functions of Participants realised with shell nouns	336
B.1	Interpersonal functions of shell nouns construing Actor in Published and Submissions 336	
B.2	Textual functions of shell nouns construing Actor in Published and Submissions	336
B.3	Interpersonal functions of shell nouns construing Goal in Published and Submissions 337	
B.4	Textual functions of shell nouns construing Goal in Published and Submissions.....	337
B.5	Interpersonal functions of shell nouns construing Verbiage in Published and Submissions.....	338
B.6	Textual functions of shell nouns construing Verbiage in Published and Submissions.	338
B.7	Interpersonal functions of shell nouns construing Phenomenon in Published and Submissions.....	339
B.8	Textual functions of shell nouns construing Phenomenon in Published and Submissions 339	
B.9	Interpersonal functions of shell nouns construing Existent in Published and Submissions.....	340
B.10	Textual functions of shell nouns construing Existent in Published and Submissions ..	340
B.11	Interpersonal functions of sShell nouns construing Token in Published and Submissions 341	
B.12	Textual functions of shell nouns construing Token in Published and Submissions	341
B.13	Interpersonal function s of shell nouns construing Value in Published and Submissions 341	
B.14	Textual functions of shell nouns construing Value in Published and Submissions.....	341

B.15	Interpersonal functions of shell nouns construing Carrier in Published and Submissions	342
B.16	Textual functions of shell nouns construing Carrier in Published and Submissions	342
B.17	Interpersonal functions of shell nouns construing Attribute in Published and Submissions.....	342
B.18	Textual functions of shell nouns construing Attribute in Published and Submissions	342
B.19	Interpersonal function of shell nouns construing Medium in Published and Submissions	343
B.20	Textual function of shell nouns construing Medium in Published and Submissions ...	343
Appendix C	Shell nouns as Participants across texts.....	344
C.1	Shell nouns construing Actor in Published.....	344
C.2	Shell nouns construing Actor in Submissions.....	344
C.3	Shell nouns construing Goal in Published	345
C.4	Shell nouns construing Goal in Submissions	346
C.5	Shell nouns construing Scope in Published	347
C.6	Shell nouns construing Scope in Submissions	347
C.7	Shell nouns construing Sayer in Published and Submissions	348
C.8	Shell nouns construing Verbiage in Published	348
C.9	Shell nouns construing Verbiage in Submissions	348
C.10	Shell nouns construing Senser in Submissions.....	349
C.11	Shell nouns construing Phenomenon in Published.....	349
C.12	Shell nouns construing Phenomenon in Submissions.....	349
C.13	Shell nouns construing Existent in Published.....	350
C.14	Shell nouns construing Existent in Submissions.....	350
C.15	Shell nouns construing Token in Published.....	351
C.16	Shell nouns construing Token in Submissions.....	352
C.17	Shell nouns construing Value in Published	353
C.18	Shell nouns construing Value in Submissions	354
C.19	Shell nouns construing Carrier in Published.....	355
C.20	Shell nouns construing Carrier in Submissions.....	356
C.21	Shell nouns construing Attribute in Published.....	357
C.22	Shell nouns construing Attribute in Submissions.....	357
C.23	Shell nouns construing Medium in Published	358
C.24	Shell nouns construing Medium in Submissions.....	358
Appendix D	Modification of shell nouns construing Participant	359

D.1	Distribution of experiential elements at group rank in Participants realised with shell nouns in Published and Submissions	359
D.2	Shell nouns construing Participant modified by Post-Deictic in Published and Submissions.....	360
D.3	Shell nouns construing Participant modified by Numerative in Published and Submissions.....	361
D.4	Shell nouns construing Participant modified by Epithet in Published and Submissions 362	
D.5	Shell nouns construing Participant modified by Classifier in Published and Submissions 363	
D.6	Participant functions of shell nouns with lexicalisation projected as Qualifier in Published and Submissions	364
D.7	Textual functions of shell nouns with lexicalisation projected as Qualifier in Published and Submissions.....	365
D.8	Interpersonal functions of shell nouns with lexicalisation projected as Qualifier in Published and Submissions	365
D.9	Shell nouns construing Participant modified by Qualifier (expanding clause) in Published.....	366
D.10	Shell nouns construing Participant modified by Qualifier (expanding clause) in Submissions.....	367
D.11	Shell nouns construing Participant modified by Qualifier (prepositional phrase) in Published.....	368
D.12	Shell nouns construing Participant modified by Qualifier (prespositional phrase) in Submissions.....	369
Appendix E	Circumstances realised with shell nouns	370
E.1	Distribution of Circumstances in Published and Submissions.....	370
E.2	Textual functions of Circumstances realised with shell noun	371
E.3	Distribution of elements tagged at group rank in Circumstances in Published and Submissions.....	372
Appendix F	Textual functions at the discourse semantics stratum	374
F.1	Reference in Published and Submissions.....	374
F.2	Distance of lexicalisation from shell noun in Published and Submissions	374
F.3	Textual functions of anaphorically referring shell nouns in Published and Submissions 375	
F.4	Coupling of anaphorically referring shell nouns with interpersonal meaning in Published and Submissions	377
F.5	Textual functions of cataphorically lexicalised shell nouns in major clauses in Published and Submissions.....	380

F.6	Cataphorically lexicalised shell nouns in major clauses in Published and Submissions	382
F.7	Distribution of shell noun types in hyper-Themes in Published and Submissions.....	383
F.8	Distribution of interpersonal functions in shell nouns projecting field as object of study in Qualifier.....	383
Appendix G	Interpersonal functions of shell nouns at the discourse semantics stratum	384
G.1	Distribution of the contribution of shell nouns as a resource in ENGAGEMENT in Published and Submissions	384
G.2	Distribution of interpersonal selections from APPRECIATION with nominal groups with shell noun as Head in Published and Submissions.....	385
G.3	Distribution of interpersonal functions for shell nouns construing JUDGEMENT	385
Appendix H	Modification of Head in shell nominal groups	387
H.1	GRADUATION as FORCE: intensifying in nominal groups with shell noun as Head in Published and Submissions	387
H.2	GRADUATION as FORCE: quantity in nominal groups with shell noun as Head in Published and Submissions.....	388
H.3	FOCUS: valeur in nominal groups with shell noun as Head	390
H.4	FOCUS: Fulfilment: actualisation in nominal groups with shell noun as Head	390

Appendix A General analysis at the stratum of the lexico grammar

A.1 Distribution of features tagged with ideational function as type and rank in Published and Submissions

Feature	Published		Submissions					
	Percent	N	Percent	N	T Stat	Signif.	ChiSqu	Signif.
IDEATION	N=1268		N=1708					
ideational-function	98.2%	1245	98.0%	1674	0.35		0.12	
no-ideational-function	1.8%	23	2.0%	34	0.35		0.12	
IDEATION-TYPE	N=1245		N=16764					
experiential-function	99.9%	1244	99.6%	1668	1.52		2.31	
logical	0.1%	1	0.4%	6	1.52		2.31	
EXPERIENTIAL-FUNCTION	N=1244		N=1670					
group	61.8%	769	61.8%	1031	0.00		0.00	
clause	38.2%	475	38.2%	637	0.00		0.00	
LOGICAL	N=1		N=6					
relator	100.0%	1	100.0%	6	0.00		0.00	
NO-IDEATIONAL-FUNCTION	N=23		N=34					
conjunctive-adjunct	8.7%	2	14.7%	5	0.67		0.46	
minor-clause	91.3%	21	85.3%	29	0.67		0.46	
MINOR-CLAUSE	N=21		N=29					
heading	100.0%	21	100.0%	29	0.00		0.00	
other	0.0%	0	0.0%	0	0.00		0.00	

A.2 Distribution of features tagged with a transitivity function at clause rank in Published and Submissions

Feature	Published		Submissions					
	Percent	N	Percent	N	T Stat	Signif.	ChiSqu	Signif.
TRANSITIVITY	N=475		N=637					
process	47.6%	226	46.9%	299	0.21		0.04	
participant	44.4%	211	42.9%	273	0.52		0.27	
circumstance	8.0%	38	10.2%	65	1.25		1.57	

A.3 Distribution of shell nouns tagged with interpersonal function at clause rank in Published and Submissions

Feature	Published		Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
INTERPERSONAL-TYPE	N=286		N=405					
interpersonal-function	87.1%	249	82.7%	335	1.56		2.42	
no-interpersonal-function	12.9%	37	17.3%	70	1.56		2.42	
INTERPERSONAL-FUNCTION-TYPE	N=249		N=335					
subject	47.0%	117	38.2%	128	2.13	++	4.52	++
complement	33.7%	84	40.0%	134	1.55		2.40	
adjunct	19.3%	48	21.8%	73	0.74		0.55	
NO-INTERPERSONAL-FUNCTION-TYPE	N=37		N=70					
conjunction)	2.7%	1	8.6%	6	1.16		1.36	
(conjunctive-adjunct)	5.4%	2	7.1%	5	0.34		0.12	
absolute	56.8%	21	41.4%	29	1.51		2.28	
(qualifier)	35.1%	13	42.9%	30	0.77		0.60	

A.4 Distribution of features tagged with textual function at clause rank in Published and Submissions

Feature	Published		Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
TEXTUAL-TYPE	N=293		N=415					
textual-function	88.7%	260	85.8%	356	1.15		1.33	
no-textual-function	11.3%	33	14.2%	59	1.15		1.33	
TEXTUAL-FUNCTION-TYPE	N=260		N=356					
theme-selection	50.8%	132	49.2%	175	0.39		0.16	
n-rheme	39.6%	103	42.1%	150	0.63		0.39	
(other)	9.6%	25	8.7%	31	0.39		0.15	
THEME-SELECTION	N=132		N=175					
clause-theme-selection	93.9%	124	93.7%	164	0.08		0.01	
clausal-theme	6.1%	8	6.3%	11	0.08		0.01	
CLAUSE-THEME-SELECTION-TYPE	N=124		N=164					
topical-theme	97.6%	121	92.1%	151	2.03	++	4.08	++
textual-theme	2.4%	3	6.7%	11	1.68	+	2.81	+
predicated-theme	0.0%	0	1.2%	2	0.00		1.52	
TOPICAL-THEME-TYPE	N=121		N=151					
unmarked-theme	93.4%	113	81.5%	123	2.92	+++	8.33	+++
marked-theme	6.6%	8	18.5%	28	2.92	+++	8.33	+++
TEXTUAL-THEME-TYPE	N=3		N=11					
conjunction-theme	33.3%	1	54.5%	6	0.61		0.42	
conjunctive-adjunct-theme	66.7%	2	45.5%	5	0.61		0.42	
N-RHEME-TYPE	N=103		N=150					
not-internal	61.2%	63	57.3%	86	0.61		0.37	
clause-complex-internal	38.8%	40	42.7%	64	0.61		0.37	
NO-TEXTUAL-FUNCTION-TYPE	N=33		N=59					
(heading)	63.6%	21	49.2%	29	1.34		1.79	
((qualifier))	36.4%	12	50.8%	30	1.34		1.79	

A.5 Distribution of Participants in Published and Submissions

Feature	Published		Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
PARTICIPANT-TYPE	N=211		N=273					
actor	3.8%	8	1.8%	5	1.32		1.75	
goal	12.8%	27	17.9%	49	1.55		2.39	
token	10.4%	22	12.8%	35	0.81		0.66	
value	28.4%	60	23.4%	64	1.25		1.56	
carrier	13.7%	29	13.2%	36	0.18		0.03	
attributor	0.0%	0	1.5%	4	0.00		3.12	+
senser	0.0%	0	1.5%	4	0.00		3.12	+
phenomenon	2.8%	6	3.7%	10	0.50		0.25	
existent	3.3%	7	1.5%	4	1.36		1.84	
medium	6.6%	14	7.7%	21	0.44		0.20	
agent	1.4%	3	1.5%	4	0.04		0.00	
initiator	0.0%	0	0.0%	0	0.00		0.00	
sayer	0.5%	1	0.4%	1	0.18		0.03	
scope	0.9%	2	0.7%	2	0.26		0.07	
verbiage	4.7%	10	3.3%	9	0.81		0.66	
assigner	0.0%	0	0.7%	2	0.00		1.55	
beneficiary	0.0%	0	0.4%	1	0.00		0.77	
attribute	10.4%	22	8.1%	22	0.90		0.81	

A.6 Distribution of Process types in Published and Submissions

Feature	Published		Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
PROCESS-TYPE	N=226		N=299					
relational	58.0%	131	53.5%	160	1.02		1.03	
material	27.4%	62	32.8%	98	1.32		1.73	
verbal	8.0%	18	5.7%	17	1.04		1.07	
mental-pr	3.1%	7	6.7%	20	1.85	+	3.40	+
existential	3.5%	8	1.3%	4	1.67	+	2.79	+
MARKEDNESS	N=131		N=160					
unmarked-pr	67.9%	89	53.1%	85	2.58	+++	6.57	+++
marked-pr	32.1%	42	46.9%	75	2.58	+++	6.57	+++
MODE-OF-RELATION	N=131		N=160					
identifying	58.8%	77	59.4%	95	0.10		0.01	
attributive	41.2%	54	40.6%	65	0.10		0.01	
SINGLE-AGENCY-TYPE	N=67		N=74					
exemplifying	29.9%	20	18.9%	14	1.52		2.30	
exhausting	70.1%	47	81.1%	60	1.52		2.30	
EXHAUSTING-TYPE	N=47		N=60					
neutral	66.0%	31	56.7%	34	0.97		0.95	
symbol	0.0%	0	0.0%	0	0.00		0.00	
significance	2.1%	1	11.7%	7	1.88	+	3.47	+
sign	31.9%	15	31.7%	19	0.03		0.00	
naming	0.0%	0	0.0%	0	0.00		0.00	
defining	0.0%	0	0.0%	0	0.00		0.00	
kind/part	0.0%	0	0.0%	0	0.00		0.00	
SIGN-TYPE	N=15		N=19					
high	6.7%	1	36.8%	7	2.14	++	4.24	++
median	26.7%	4	36.8%	7	0.61		0.40	
low	66.7%	10	26.3%	5	2.49	+++	5.54	+++
TYPE-OF-RELATION	N=131		N=160					
intensive	80.9%	106	73.1%	117	1.56		2.44	
possessive	8.4%	11	11.9%	19	0.97		0.94	
circumstantial	10.7%	14	15.0%	24	1.08		1.18	
AGENCY	N=106		N=117					
single-agency	99.1%	105	92.3%	108	2.45	+++	5.91	+++
double-agency	0.9%	1	7.7%	9	2.45	+++	5.91	+++
ASSIGNMENT	N=1		N=9					
assignment-neutral	100.0%	1	88.9%	8	0.32		0.12	
assignment-elaborating	0.0%	0	0.0%	0	0.00		0.00	
assignment-projection	0.0%	0	11.1%	1	0.00		0.12	
CIRCUMSTANTIAL--REL-PRO-TYPE	N=14		N=24					
spatio-temporal	42.9%	6	37.5%	9	0.32		0.11	

comparative	7.1%	1	4.2%	1	0.39		0.16	
causal:-reason	35.7%	5	25.0%	6	0.69		0.49	
causal:-concession	7.1%	1	8.3%	2	0.13		0.02	
causal:-condition	0.0%	0	4.2%	1	0.00		0.60	
matter-circ-rel	7.1%	1	20.8%	5	1.10		1.25	

A.7 Distribution of clause types containing shell noun in Published and Submissions

CLAUSE-TYPE	N=261		N=362					
rankshifted	3.4%	9	4.4%	16	0.61		0.37	
bound	12.6%	33	19.9%	72	2.39	+++	5.68	+++
free	76.2%	199	67.7%	245	2.34	++	5.43	+++
(minor-clause)	7.7%	20	8.0%	29	0.16		0.03	
RANKSHIFTED-TYPE	N=9		N=16					
fact-cl	66.7%	6	37.5%	6	1.40		1.96	
embedded-elaborating-clause	33.3%	3	62.5%	10	1.40		1.96	
LOGICO-SEMANTIC-TYPE	N=131		N=203					
projection	5.3%	7	6.9%	14	0.57		0.33	
expansion	94.7%	124	93.1%	189	0.57		0.33	
PROJECTION-TYPE	N=7		N=14					
locution	57.1%	4	57.1%	8	0.00		0.00	
idea	42.9%	3	42.9%	6	0.00		0.00	
EXPANSION-TYPE	N=124		N=189					
elaborating	33.1%	41	33.3%	63	0.05		0.00	
extending	18.5%	23	25.4%	48	1.42		2.00	
enhancing	48.4%	60	41.3%	78	1.24		1.54	
TAXIS	N=131		N=203					
hypotaxis	71.0%	93	70.9%	144	0.01		0.00	
parataxis	29.0%	38	29.1%	59	0.01		0.00	
FREE-TYPE	N=199		N=245					
clause-complex	49.2%	98	53.5%	131	0.88		0.78	
simple	50.8%	101	46.5%	114	0.88		0.78	

Appendix B Textual and Interpersonal functions of Participants realised with shell nouns

B.1 Interpersonal functions of shell nouns construing Actor in Published and Submissions

Feature	Actor in Published		Actor in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
INTERPERSONAL-FUNCTION-TYPE	N=8		N=5					
Subject	100.00	8	100.00	5	0.000		0.000	
Complement	0.00	0	0.00	0	0.000		0.000	
Adjunct	0.00	0	0.00	0	0.000		0.000	

B.2 Textual functions of shell nouns construing Actor in Published and Submissions

Feature	Actor in Published		Actor in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
TEXTUAL-FUNCTION-TYPE	N=8		N=5					
theme-selection	100.00	8	100.00	100.00	0.000		0.000	
n-rheme	0.00	0	0.00	0.00	0.000		0.000	
(other)	0.00	0	0.00	0.00	0.000		0.000	

B.3 Interpersonal functions of shell nouns construing Goal in Published and Submissions

Feature	Goal in Published		Goal in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
INTERPERSONAL-FUNCTION-TYPE	N=27		N=49					
subject	3.7	1	0.0	0	0.00		1.84	
complement	96.3	26	100.0	49	1.35		1.84	
subject	3.7	1	0.0	0	0.00		1.84	

Note: In one instance in each of the sub-corpora, the shell noun construing Goal appears in a Qualifier to another shell noun. As detailed in chapter 4, Qualifiers are tagged as having no interpersonal function, so cross classification of Goal with interpersonal function results in a count that is lower than the counts for Goal, where shell nouns construing Head in the nominal group and shell nouns in Qualifiers are counted.

B.4 Textual functions of shell nouns construing Goal in Published and Submissions

Feature	Goal in Published		Goal in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
TEXTUAL-FUNCTION-TYPE	N=27		N=49					
theme-selection	3.7	1	0.0	0	0.00		1.84	
n-rheme	85.2	23	81.6	40	0.39		0.15	
(other)	11.1	3	18.4	9	0.82		0.69	

Note: In one instance in each of the sub-corpora, the shell noun construing Goal appears in a Qualifier to another shell noun. As detailed in chapter 4, Qualifiers are tagged as having no textual function, so cross classification of Goal with textual function results in a count that is lower than the counts for Goal, where shell nouns construing Head in the nominal group and shell nouns in Qualifiers are counted.

B.5 Interpersonal functions of shell nouns construing Verbiage in Published and Submissions

Feature	Verbiage in Published		Verbiage in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
INTERPERSONAL-FUNCTION-TYPE	N= 10		N= 10					
Subject	0.0	0	10.0	1	0.00		1.05	
Complement	100.0	10	90.0	9	1.00		1.05	
Adjunct	0.0	0	0.0	0	0.00		0.00	

B.6 Textual functions of shell nouns construing Verbiage in Published and Submissions

Feature	Verbiage in Published		Verbiage in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
TEXTUAL-FUNCTION-TYPE	N = 10		N = 10					
Theme-selection	0.0	0	10.0	1	0.00		1.05	
N-Rheme	80.0	8	80.0	8	0.00		0.00	
(Other)	20.0	2	10.0	1	0.60		0.39	

B.7 Interpersonal functions of shell nouns construing Phenomenon in Published and Submissions

Feature	Phenomenon in Published		Phenomenon in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
INTERPERSONAL-FUNCTION-TYPE	N=6		N=10					
Subject	66.7	4	0.0	0	0.00		8.89	+++
Complement	33.3	2	100.0	10	4.18	+++	8.89	+++
Adjunct	0.0	0	0.0	0	0.00		0.00	

B.8 Textual functions of shell nouns construing Phenomenon in Published and Submissions

Feature	Phenomenon in Published		Phenomenon in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
TEXTUAL-FUNCTION-TYPE	N=6		N=10					
theme-selection	50.0	3	0.0	0	0.00		6.15	+++
n-rheme	33.3	2	60.0	6	1.00		1.07	
(other)	16.7	1	40.0	4	0.94		0.95	

B.9 Interpersonal functions of shell nouns construing Existent in Published and Submissions

Feature	Existent in Published		Existent in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
INTERPERSONAL-FUNCTION-TYPE	N=7		N=4					
Subject	0.00	0	0.00	0	0.00		0.00	
Complement	100.00	7	100.00	4	0.00		0.00	
Adjunct	0.00	0	0.00	0	0.00		0.00	

B.10 Textual functions of shell nouns construing Existent in Published and Submissions

Feature	Existent in Published		Existent in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
TEXTUAL-FUNCTION-TYPE	N=7		N=4					
Theme-selection	0.00	0	0.00	0	0.00		0.00	
N-Rheme	100.00	7	100.00	4	0.00		0.00	
(Other)	0.00	0	0.00	0	0.00		0.00	

B.11 Interpersonal functions of sShell nouns construing Token in Published and Submissions

Feature	Token in Published		Token in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
INTERPERSONAL-FUNCTION-TYPE	N=22		N=35					
subject	81.8%	18	82.9%	29	0.10		0.01	
complement	13.6%	3	14.3%	5	0.07		0.00	
adjunct	4.5%	1	2.9%	1	0.33		0.11	

B.12 Textual functions of shell nouns construing Token in Published and Submissions

Feature	Token in Published		Token in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
TEXTUAL-FUNCTION-TYPE	N=22		N=35					
theme-selection	77.3%	17	88.6%	31	1.13		1.30	
n-rheme	18.2%	4	11.4%	4	0.71		0.51	
(other)	4.5%	1	0.0%	0	0.00		1.62	

B.13 Interpersonal function s of shell nouns construing Value in Published and Submissions

Feature	Value in Published		Value in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
INTERPERSONAL-FUNCTION-TYPE	N=61		N=64					
subject	68.9%	42	43.8%	28	2.90	+++	7.99	+++
complement	31.1%	19	51.6%	33	2.35	++	5.36	++
adjunct	0.0%	0	4.7%	3	0.00		2.93	+

B.14 Textual functions of shell nouns construing Value in Published and Submissions

Feature	Value in Published		Value in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
TEXTUAL-FUNCTION-TYPE	N=61		N=64					
theme-selection	67.2%	41	39.1%	25	3.26	+++	9.93	+++
n-rheme	21.3%	13	53.1%	34	3.85	+++	13.47	+++
(other)	11.5%	7	7.8%	5	0.69		0.48	

B.15 Interpersonal functions of shell nouns construing Carrier in Published and Submissions

Feature	Carrier in Published		Carrier in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
INTERPERSONAL-FUNCTION-TYPE	N= 29		N= 36					
subject	100.0%	29	86.1%	31	2.13	++	4.36	++
complement	0.0%	0	13.9%	5	0.00		4.36	++
adjunct	0.0%	0	0.0%	0	0.00		0.00	

B.16 Textual functions of shell nouns construing Carrier in Published and Submissions

Feature	Carrier in Published		Carrier in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
TEXTUAL-FUNCTION-TYPE	N= 29		N= 36					
selection	93.1%	27	80.6%	29	1.46		2.12	
n-rheme	0.0%	0	13.9%	5	0.00		4.36	++
(other)	6.9%	2	5.6%	2	0.22		0.05	

B.17 Interpersonal functions of shell nouns construing Attribute in Published and Submissions

Feature	Attribute in Published		Attribute in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
INTERPERSONAL-FUNCTION-TYPE	N=22		N=22					
subject	0.0%	0	0.0%	0	0.00		0.00	
complement	72.7%	16	95.5%	21	2.12	++	4.25	++
adjunct	27.3%	6	4.5%	1	2.12	++	4.25	++

B.18 Textual functions of shell nouns construing Attribute in Published and Submissions

Feature	Attribute in Published		Attribute in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
TEXTUAL-FUNCTION-TYPE	N=22		N=22					
theme-selection	4.5%	1	4.5%	1	0.00		0.00	
n-rheme	77.3%	17	77.3%	17	0.00		0.00	
(other)	18.2%	4	18.2%	4	0.00		0.00	

B.19 Interpersonal function of shell nouns construing Medium in Published and Submissions

Feature	Medium in Published		Medium in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
INTERPERSONAL-FUNCTION-TYPE	N=14		N=21					
Subject	100.0%	14	100.0%	21	0.00		0.00	
Complement	0.0%	0	0.0%	0	0.00		0.00	
Adjunct	0.0%	0	0.0%	0	0.00		0.00	

B.20 Textual function of shell nouns construing Medium in Published and Submissions

Qualifier	Medium in Published		Medium in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
TEXTUAL-FUNCTION-TYPE	N= 14		N= 21					
Theme-selection	100.0%	14	95.2%	20	0.81		0.69	
N-Rheme	0.0%	0	0.0%	0	0.00		0.00	
(Other)	0.0%	0	4.8%	1	0.00		0.69	

Appendix C Shell nouns as Participants across texts

C.1 Shell nouns construing Actor in Published

	P_Rand		P_Net		P_CEO		P_Pov		P_Econ	
Shell noun										
Factual										
Linguistic									appeals	1
Mental									assumption	1
Modal										
Eventive			success	1						
Circumstantial					criteria	3	method	1	practice	1

C.2 Shell nouns construing Actor in Submissions

	S_Rand		S_Net		S_CEO		S_Pov		S_Econ	
Shell noun										
Factual			fact	1						
Linguistic	proposition	1							proposal	1
Mental			analysis	1						
Modal										
Eventive										
Circumstantial									restriction	1

C.3 Shell nouns construing Goal in Published

	P_Rand		P_Net		P_CEO		P_Pov		P_Econ	
Shell noun										
Factual			evidence	1	reason	1				
Linguistic	question	1	argument	2	laws	1				
	answer	1	arguments	2						
			assertion	1						
Mental	desire	1			idea	2	concerns	1	thesis	2
	idea	1							impression	1
Modal			ability	2						
Eventive				1			improvements	1		
Circumstantial				1	situation	1	approach	1		
							approaches	1		
							shortcoming	1		
							shortcomings	1		
							ways	1		
							method	1		

C.4 Shell nouns construing Goal in Submissions

	S_Rand		S_Net		S_CEO		S_Pov		S_Econ	
Shell noun										
Factual	results proof problem example	2 2 1 1	result	1	problem	2				
Linguistic	statement contribution	1 1	statement	1					proposal insinuation suggestion	2 1 1
Mental	strategy goal	1 1	hypothesis incentives solution solutions intentions	4 2 2 1 1	hypothesis	2			view goal	1 1
Modal	probability possibility	1 1	probability risk	4 2	opportunity	1	opportunity	1	requirements	1
Eventive	refinement	1	efforts	1					effort	1
Circumstantial					approach	1	criteria	1	criteria	1

C.5 Shell nouns construing Scope in Published

	P_Rand		P_Net		P_CEO		P_Pov		P_Econ	
Shell noun										
Factual										
Linguistic										
Mental							assumption	1		
Modal			trend	1						
Eventive										
Circumstantial										

C.6 Shell nouns construing Scope in Submissions

	S_Rand		S_Net		S_CEO		S_Pov		S_Econ	
Shell noun										
Factual										
Linguistic										
Mental										
Modal									challenge task	1 1
Eventive										
Circumstantial										

C.7 Shell nouns construing Sayer in Published and Submissions

	P_CEO		S_Rand	
Factual	findings	1	example	1

C.8 Shell nouns construing Verbiage in Published

	P_Rand		P_Net		P_CEO		P_Pov		P_Econ	
Shell noun										
Factual	result	2			reason	1			finding	1
Linguistic	question	1								
	answer	1								
Mental	theorem	1								
Modal							specification	1		
Eventive										
Circumstantial							ways	1		
							approach	1		

C.9 Shell nouns construing Verbiage in Submissions

	S_Rand		S_Net		S_CEO		S_Pov		S_Econ	
Shell noun										
Factual					results	2	findings	1	reason	1
Linguistic							conclusions	1	argument	1
									suggestion	1
									proposal	1
Mental	assumption	1							ideas	1
Modal										
Eventive										
Circumstantial										

C.10 Shell nouns construing Senser in Submissions

	S_Rand		S_Net		S_CEO		S_Pov		S_Econ	
Shell noun										
Factual										
Linguistic									proposal	1
Mental			hypothesis	2					interpretation	1
Modal										
Eventive										
Circumstantial										

C.11 Shell nouns construing Phenomenon in Published

	P_Rand		P_Net		P_CEO		P_Pov		P_Econ	
Shell noun										
Factual									fact problem	1 1
Linguistic									observations statement	1 1
Mental									hopes discinclination	1 1
Modal										
Eventive										
Circumstantial	process	1								

C.12 Shell nouns construing Phenomenon in Submissions

	S_Rand		S_Net		S_CEO		S_Pov		S_Econ	
Shell noun										
Factual	example	2					findings	1		
Linguistic									proposal suggestion	2 1
Mental	strategy	2	hypothesis	1			implications	1		
Modal							probabilities	1		
Eventive										
Circumstantial										

C.13 Shell nouns construing Existent in Published

	P_Rand		P_Net		P_CEO		P_Pov		P_Econ	
Shell noun										
Factual					features evidence	1 2				
Linguistic										
Mental							motivations	1		
Modal	need	1	tendency	1						
Eventive										
Circumstantial					ways	1				

C.14 Shell nouns construing Existent in Submissions

	S_Rand		S_Net		S_CEO		S_Pov		S_Econ	
Shell noun										
Factual										
Linguistic										
Mental			consensus intention	1						
Modal			probabilities	1						
Eventive							attempts	1		
Circumstantial										

C.15 Shell nouns construing Token in Published

	P_Rand		P_Net		P_CEO		P_Pov		P_Econ	
Shell noun										
Factual			fact results	3 1	fact case effect results	1 1 1 2			phenomenon	1
Linguistic									descriptions	1
Mental			analysis	3	ideas	1	analyses	1	view	1
Modal							need probability	1 1		
Eventive										
Circumstantial					constraint	2	way approach	1 1		

C.16 Shell nouns construing Token in Submissions

	S_Rand		S_Net		S_CEO		S_Pov		S_Econ	
Shell noun										
Factual	fact	1	fact	2	result	1	advantage	1		
	result	1	results	1	findings	1				
	proof	1	aspect	1	elements	1				
	results	2	advantage	1	profile	1				
	corollary	1	feature	1						
Linguistic	proposition	2	questions	1					suggestion	2
	definition	1	remarks	1					comment	1
Mental	strategy	1	solution	1	strategy	1	perspective	1		
			hypothesis	1	conceptualization	1				
					issues	1				
Modal	need	1	probability	1			possibility	1	ability	1
Eventive					development	1			efforts	1
Circumstantial										

C.17 Shell nouns construing Value in Published

	P_Rand		P_Net		P_CEO		P_Pov		P_Econ	
Shell noun										
Factual	reasons problem advantage similarity	1 1 1 1	case evidence problem fact focus	2 1 1 1 1	reason problem fact case example	2 1 1 1 1	advantage advantages	2 1	point problem thing aspects irony advantage dilemma	2 1 1 1 1 1 1
Linguistic	conclusion	1	conclusion	1	explanation	1	suggestion	1	proposal	1
Mental	aim intuition analysis	1 1 1	picture intention	1 1	purpose idea	1 1	objective motivation rationale consideration concern	2 2 1 1 1	objective purpose concern resolution interpretation	1 1 1 1 1
Modal	step	1			possibility	1			steps	1
Eventive	attempt	1	attempt	1					choice task	1 1
Circumstantial	conditions approach	1 1			constraint	1	way limitation limitations criterion	1 1 1 1		

C.18 Shell nouns construing Value in Submissions

	S_Rand		S_Net		S_CEO		S_Pov		S_Econ	
Shell noun										
Factual	case	1	factor	1	results	1	case	1	case	1
	example	1	case	1	implication	1	advantage	1	result	1
	difficulties	1	importance	1	implications	1			fact	1
			matter	1						
			reason	1						
Linguistic	proposition	1			conclusion	4			motto	1
Mental	assumption	2	aim	2	hypothesis	3	aim	2	concept	1
	aim	1	issues	1	function	1	strategy	1	solution	1
	purpose lessons	1	strategy	1	inference	1				
		1								
Modal	probability	2	probability	3	opportunity	2	ability	2		
	condition	1	inability	1			probability	1		
	requirement	1	potential	1			impossibility	1		
							capacity	1		
Eventive	alternative	1	difficulties	1			improvement		priority	1
	action	1					modification			
	choice	1								
Circumstantial	way	1	way	1						

C.19 Shell nouns construing Carrier in Published

	P_Rand		P_Net		P_CEO		P_Pov		P_Econ	
Shell noun										
Factual	proof results	3 1			result results pattern	4 3 1	feature	1	fact	2
Linguistic									claim appeals explanation metaphor point	1 1 1 1 1
Mental	assumption theorem	2 1			idea	1				
Modal	probability	1	trend	1						
Eventive										
Circumstantial			position occurrence	1 1			criterion	1		

C.20 Shell nouns construing Carrier in Submissions

	S_Rand		S_Net		S_CEO		S_Pov		S_Econ	
Shell noun										
Factual	result problem	2 1	fact alternatives cases	1 1 1	problem result	1 1	feature	1	result paradox	1 1
Linguistic	remarks observation	3 1	question argument	1 1					answers suggestion	1 1
Mental	lemma	1	consensus hypothesis	1 1	concerns	1	strategy	1	idea	1
Modal	probability	2	probability probabilities chances	1 1 1	power need	1 1			task agenda	1 1
Eventive										
Circumstantial					procedure approaches situation	1 1 1			criteria situation	1 1

C.21 Shell nouns construing Attribute in Published

	P_Rand		P_Net		P_CEO		P_Pov		P_Econ	
Shell noun										
Factual			fact problems	1 1	fact findings benefit	3 1 1			fact evidence accident	2 1 1
Linguistic							recommendations	1	tidings dictum metaphor	1 1 1
Mental							assumptions	1	view theory belief	1 1 1
Modal			trend	1						
Eventive									choices	1
Circumstantial									method	1

C.22 Shell nouns construing Attribute in Submissions

	S_Rand		S_Net		S_CEO		S_Pov		S_Econ	
Shell noun										
Factual	result property implication	1 1 1	problem advantage benefits	1 1 1						
Linguistic	propositions	1	remark explanation	1 1					encouragement commitment insinuation	1 1 1
Mental	agreement	1	solution subject	1 1			concern	1		
Modal	condition	1	opportunities potentials	1 1					certainty	1
Eventive									success	1
Circumstantial			way means	1 1			approach	1	limitation	1

C.23 Shell nouns construing Medium in Published

	P_Rand		P_Net		P_CEO		P_Pov		P_Econ	
Shell noun										
Factual	proof	1			reason results features	2 1 1	difficulty	1		
Linguistic							question	1		
Mental									expectation	1
Modal			trend	3						
Eventive									struggle	1
Circumstantial			position	1					situation	1

C.24 Shell nouns construing Medium in Submissions

	S_Rand		S_Net		S_CEO		S_Pov		S_Econ	
Shell noun										
Factual	result	2	results	1	corroboration	1	results	1	difficulty miracle	1 1
Linguistic	proposition	1							proposal remarks comment argument	1 1 1 1
Mental			hypothesis strategy	1 1	hypothesis policy	1 1				
Modal					steps	1	need	1	steps	1
Eventive									initiative	1
Circumstantial					approach	1	criterion	1		

Appendix D Modification of shell nouns construing Participant

D.1 Distribution of experiential elements at group rank in Participants realised with shell nouns in Published and Submissions

Feature	Published		Submissions					
	Percent	N	Percent	N	T Stat	Signif.	ChiSqu	Signif.
GROUP-EXP-FUNCTION	N=611		N=768					
deictic	31.4%	192	32.3%	248	0.34		0.12	
post-deictic	4.1%	25	3.9%	30	0.17		0.03	
numerative	3.1%	19	2.7%	21	0.41		0.17	
epithet	4.7%	29	4.8%	37	0.06		0.00	
classifier	3.3%	20	2.3%	18	1.05		1.10	
thing	35.5%	217	37.0%	284	0.56		0.32	
qualifier	17.8%	109	16.9%	130	0.44		0.20	
DEIXIS	N=192		N=248					
specific	81.8%	157	83.5%	207	0.47		0.22	
non-specific	18.2%	35	16.5%	41	0.47		0.22	
SPECIFIC-TYPE	N=157		N=207					
personal-possession	10.8%	17	17.4%	36	1.76	+	3.09	+
demonstrative-proximity	89.2%	140	82.6%	171	1.76	+	3.09	+
PERSONAL-POSSESSION-TYPE	N=17		N=36					
pronominal	70.6%	12	52.8%	19	1.22		1.51	
embedded-nominal-group	29.4%	5	47.2%	17	1.22		1.51	
DEMONSTRATIVE-PROXIMITY-TYPE	N=140		N=171					
non-selective	65.7%	92	77.2%	132	2.25	++	5.03	++
selective	34.3%	48	22.8%	39	2.25	++	5.03	++
NUMBER	N=48		N=39					
singular	77.1%	37	74.4%	29	0.29		0.09	
plural	22.9%	11	25.6%	10	0.29		0.09	
DISTANCE	N=48		N=39					
near	97.9%	47	97.4%	38	0.15		0.02	
far	2.1%	1	2.6%	1	0.15		0.02	
SHELL-NOUN-TYPE	N=217		N=284					
factual	37.8%	82	26.8%	76	2.65	+++	6.93	+++
linguistic	13.8%	30	17.6%	50	1.14		1.31	
mental	22.1%	48	24.3%	69	0.57		0.33	
modal	7.8%	17	18.0%	51	3.31	+++	10.75	+++
eventive	3.7%	8	6.0%	17	1.17		1.37	
circumstantial-shell-noun	14.7%	32	7.4%	21	2.66	+++	7.03	+++
QUALIFIER-TYPE	N=109		N=130					
projected	54.1%	59	59.2%	77	0.79		0.63	
expanded	45.9%	50	40.8%	53	0.79		0.63	

PROJECTED-TYPE	N=59		N=77					
finite	69.5%	41	27.3%	21	5.36	+++	24.00	+++
non-finite	30.5%	18	72.7%	56	5.36	+++	24.00	+++
NON-FINITE-TYPE	N=18		N=56					
perfective	88.9%	16	50.0%	28	3.07	+++	8.55	+++
of-+-imperfective	11.1%	2	50.0%	28	3.07	+++	8.55	+++
EXPANDED-TYPE	N=50		N=53					
embedded-clause	30.0%	15	17.0%	9	1.57		2.44	
prepositional-phrase	70.0%	35	83.0%	44	1.57		2.44	
EMBEDDED-CLAUSE-TYPE	N=15		N=9					
non-finite-cl	60.0%	9	55.6%	5	0.20		0.05	
finite-cl	40.0%	6	44.4%	4	0.20		0.05	
NON-FINITE-CL-TYPE	N=9		N=5					
imperfective-cl	0.0%	0	20.0%	1	0.00		1.94	
neutral-cl	22.2%	2	60.0%	3	1.41		2.00	
prep-+-imperfective-cl	77.8%	7	20.0%	1	2.34	++	4.38	++
perfective-cl	0.0%	0	0.0%	0	0.00		0.00	

D.2 Shell nouns construing Participant modified by Post-Deictic in Published and Submissions

	Published			Submissions		
	Post Deictic	Shell noun	N	Post-Deictic	Shell noun	N
Factual	last, possible, other, classic, same, differing, basic, next, mere, little-known	reason results feature example findings fact thing benefit	3 2 1 1 1 2 1 1	following, above, next, additional, further, present	result(s) example proof advantage corroboration problem	6 3 1 1 1 1
linguistic	further	tidings	1	final	conclusion	1
mental	primary, customary, following, same	motivation assumption view theorem idea	1 1 1 1 1	alternative, following, same, initial, described, own, dominant, most usual	strategy lemma assumption ideas concept view	5 2 2 1 1 1
modal	expected, following, next	step(s) trend	2 1	above, different	requirement condition	1 1
eventive				only, next, possible, potential	alternative initiative modification action	1 1 1 1
circumstantial	final, other, potential	limitation ways	1 1	other, alternative	way approaches	1 1

D.3 Shell nouns construing Participant modified by Numerative in Published and Submissions

Shell noun	Published			Submissions		
	Numerative	Shell noun	Nº	Numerative	Shell noun	Nº
Factual	several, first, third, one, two, one of the	reason(s) advantage (s) feature aspect	4 2 1 1	first, second, two, three, four, one of the, some of the, a battery of	advantage aspect s elements findings reasons difficulties alternatives	2 1 1 1 1 1 1
linguistic	none of those	appeals	1	two, a couple of, first	remarks observation	2 1
mental	one two a number of	objective motivations assumptions	1 1 1	four, one of the, second, several,	hypothesis lessons solutions issues	3 1 1 1
modal				several first	steps task	1 1
eventive	several first two	improvements attempt choices	1 1 1	one, many	action attempts	1 1
circumstantial	two, both, a host of three	ways approaches conditions	2 1 1	one	way	1

D.4 Shell nouns construing Participant modified by Epithet in Published and Submissions

Shell noun	Published			Submissions		
	Epithet	Shell noun	N	Epithet	Shell noun	N
Factual	distinct, large, serious, most curious, brief, main, important	advantage	1	main	aspects	1
		effect	1	predominant	findings	1
		dilemma	1	inevitable	result	2
		aspects	1	illuminating	implication	1
		result	1			
		problem	1			
		similarity findings	1			
		evidence	1			
			1			
linguistic	important bad specific	question	1	general, main,	conclusion	3
		news	1	particular,	argument	1
		recommendations,	1	detailed,	remark	1
		tidings	1	unfair	explanation insinuation	1 1
mental	main, major, primary, complicated, ideal, biased, prime, early, important	objective	2	clearest, main,	lessons	1
		assumptions	1	complex,	goal	1
		concerns	1	rational,	solution	1
		objective	2	fundamental,	subject	1
		consideration	1	well enough	strategy	1
		purpose	1	discussed	concerns	2
		picture	1		function	1
		resolution	1			
		view	1			
disinclination	1					
modal				high, higher,	probability	4
				low,	potential	2
				discretionary,	power	1
				urgent,	need	1
				decisive,	steps	1
				strong, better,	condition	1
				reduced	opportunities capacity	1 1
eventive	serious, unpleasant	attempt	2	intellectual,	efforts	2
		choice	1	strenuous,	improvement	1
				substantial immense	success	1
circumstantial	general, effective, strange, radically wrong	approach	1	dangerous,	limitation	1
		way	1	more	approach	1
		occurrence	1	extensive		
		method	1			

D.5 Shell nouns construing Participant modified by Classifier in Published and Submissions

Shell noun	Published			Submissions		
	Classifier	Shell noun	N	Classifier	Shell noun	N
Factual	institutional, key, empirical, biographical	features	2	college admissions, endogeneity, converse	problem	3
		findings	1		result	1
		evidence	1			
linguistic				concluding, converse	remarks statement	1 1
mental	exploratory, cluster, existence, structural, scientific	analysis(es)	2	empirical managerial, constitutional	strategy	1
		theorem	1		policy	1
		assumptions	1		goal	1
		theory	1			
modal	posterior	probability	1	average sufficient	probability condition	4 2
eventive						
circumstantial	model choice, model-based, non-distribution, financial, compensation, selection, responsiveness	constraint	4	microeconomic	approach(es)	2
		method	3			
		conditions	1			
			1			
			1			

D.6 Participant functions of shell nouns with lexicalisation projected as Qualifier in Published and Submissions

Feature	Participant with projected Qualifier in Published		Participant with projected Qualifier in Submissions					
	Percent	N	Percent	N	T Stat	Signif.	ChiSqu	Signif.
PARTICIPANT-TYPE	N=60		N=78					
actor	1.7%	1	1.3%	1	0.19		0.04	
goal	23.3%	14	19.2%	15	0.58		0.34	
token	11.7%	7	11.5%	9	0.02		0.00	
value	15.0%	9	26.9%	21	1.69	+	2.83	+
carrier	13.3%	8	12.8%	10	0.09		0.01	
attributor	0.0%	0	0.0%	0	0.00		0.00	
senser	0.0%	0	0.0%	0	0.00		0.00	
phenomenon	3.3%	2	1.3%	1	0.82		0.67	
existent	8.3%	5	3.8%	3	1.12		1.25	
medium	5.0%	3	3.8%	3	0.33		0.11	
agent	3.3%	2	2.6%	2	0.27		0.07	
initiator	0.0%	0	0.0%	0	0.00		0.00	
sayer	0.0%	0	0.0%	0	0.00		0.00	
scope	1.7%	1	0.0%	0	0.00		1.31	
verbiage	3.3%	2	1.3%	1	0.82		0.67	
assigner	0.0%	0	0.0%	0	0.00		0.00	
beneficiary	0.0%	0	1.3%	1	0.00		0.77	
attribute	10.0%	6	14.1%	11	0.72		0.53	

D.7 Textual functions of shell nouns with lexicalisation projected as Qualifier in Published and Submissions

Feature	Participant with projected Qualifier in Published		Participant with projected Qualifier in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
TEXTUAL-FUNCTION-TYPE	N=60		N=76					
theme-selection	35.0%	21	26.3%	20	1.09		1.20	
n-rheme	63.3%	38	72.4%	55	1.12		1.27	
(other)	1.7%	1	1.3%	1	0.17		0.03	

D.8 Interpersonal functions of shell nouns with lexicalisation projected as Qualifier in Published and Submissions

Feature	Participant with projected Qualifier in Published		Participant with projected Qualifier in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
INTERPERSONAL-FUNCTION-TYPE	N=60		N=78					
subject	36.7%	22	25.6%	20	1.40		1.95	
complement	53.3%	32	65.4%	51	1.43		2.05	
adjunct	10.0%	6	9.0%	7	0.20		0.04	

D.9 Shell nouns construing Participant modified by Qualifier (expanding clause) in Published

		Published
Shell noun	Shell noun	Qualifier realised as expanding clause
factual	An advantage The next thing One of the reasons at least two key features A classic example	of allowing different weights for each variable (P_Pov) which impresses the modern reader (P_Econ) for centralizing this market, and later for reorganizing it (P_Rand) of nonprofits that distinguish them from for-profit organizations (P_CEO) of how difficult it is to measure performance for managers in the nonprofit sector (P_CEO)
linguistic	a dictum	seconded by Koopmans (P_Econ)
mental	at least two motivations The motivation The rationale in the prejudiced point of view The biased point of view	for incorporating possible spatial dependencies in the model (P_Pov) for constraining θ to have 1's along the diagonal (P_Pov) for setting $\theta = D \cdot 1$ (P_Pov) from which economists regarded the possibilities of science (P_Econ) implied in these descriptions (P_Econ)
modal		
eventive	a trend	that continued until 1810 (P_Net)
circumstantial	a general approach this method in the radically wrong method	to incorporating spatial correlation in a factor-analytic model (P_Pov) of reporting (P_Pov) which they pursued (P_Econ)

D.10 Shell nouns construing Participant modified by Qualifier (expanding clause) in Submissions

Submissions		
Shell noun	Shell noun	Qualifier realised by expanding clause
factual	some results The results One of the difficulties the first alternatives the reason for mixed agency, the reason	that extend to deterministic mechanisms (S_Rand) that follow (S_Rand) that arises in attempting to apply the common game theoretical tools (S_Rand) mentioned above(S_Net) integrated in category B, becoming the pattern that assured the best economic performance (S_Net) for supporting his suggestion (S_Econ)
linguistic		
mental	the four solutions	previously pointed out in section 2 (S_Net)
modal	the risk	implied in Category A (S_Net)
eventive	The task	this small group was setting itself (S_Econ)
circumstantial		

D.11 Shell nouns construing Participant modified by Qualifier (prepositional phrase) in Published

Shell noun	Published	
	Shell noun	Qualifier realised as prepositional phrase
factual	<p>The problem at least two key features</p> <p>The first possible reason</p> <p>Another possible reason</p> <p>the case</p> <p>the results</p> <p>The differing results</p> <p>The problem</p> <p>the focus</p> <p>The results</p> <p>The problem</p> <p>One of the most curious aspects</p> <p>One advantage</p> <p>The advantage</p> <p>The main problem</p> <p>the proof</p> <p>An important similarity</p>	<p>with this legislation (P_CEO)</p> <p>of nonprofits that distinguish them from for-profit organizations (P_CEO)</p> <p>for differences in pay between employees of nonprofit organizations and for-profit firms (P_CEO)</p> <p>for differences in pay (P_CEO)</p> <p>of a manager of a nursing home (P_CEO)</p> <p>in this section (P_CEO)</p> <p>for the unpaid/paid members (P_CEO)</p> <p>with local and confidential networking (P_Net)</p> <p>of this analysis (P_Net)</p> <p>of this research (P_Net)</p> <p>with this interpretation (P_Econ)</p> <p>of the rise of neoclassical theory (P_Econ)</p> <p>of a familiarity with Hermetian matrices (P_Econ)</p> <p>of this approach (P_Rand) (P_Rand)</p> <p>with any attempt to apply the RV-algorithm directly to weakly responsive couples markets (P_Rand)</p> <p>of convergence (P_Rand)</p> <p>between roommate problems and couples markets (P_Rand)</p>
linguistic	<p>some conclusions</p> <p>these appeals</p> <p>the answer</p> <p>the answer</p> <p>one conclusion</p>	<p>about Liverpool's metropolitan business networks, and also about networks more generally. (P_Net)</p> <p>to science (P_Econ)</p> <p>to this question (P_Rand)</p> <p>to the corresponding question for roommate problems (P_Rand)</p> <p>from our analysis (P_Rand)</p>
mental	<p>one objective</p> <p>The primary motivation</p> <p>Exploratory analyses</p> <p>The main objective</p> <p>A major consideration</p> <p>a complicated picture</p> <p>the intention</p> <p>the purpose</p>	<p>in this area of research (P_Pov)</p> <p>for our proposed methodology (P_Pov)</p> <p>of the Rhode Island census data (P_Pov)</p> <p>of our analysis (P_Pov)</p> <p>in both policy and research (P_Pov)</p> <p>of metropolitan business networks (P_Net)</p> <p>of this paper (P_Net)</p> <p>of this paper (P_Econ)</p>
modal	the next step	in our analysis (P_Rand)
eventive	whose task	in its early years (P_Econ)
circumstantial	<p>A final limitation</p> <p>The model choice</p> <p>criterion</p> <p>Limitations</p>	<p>of empirical indices (P_Pov)</p> <p>for model q (P_Pov)</p> <p>of the Townsend index (P_Pov)</p>

D.12 Shell nouns construing Participant modified by Qualifier (prepositional phrase) in Submissions

Shell noun	Submissions	
	Shell noun	Qualifier realised as prepositional phrase
factual	one of the clearest lessons a refinement The proof an implication an example the proof a factor the case the benefits results the results the implications the predominant implication the result The paradox the implications The first advantage	from the study of deterministic procedures (S_Rand) of the Nash equilibrium concept (S_Rand) of the above result (S_Rand) of Propositions 6 and 7 (S_Rand) of such kind of strategies (S_Rand) for a worker w (S_Rand) for Portuguese merchants' difficulties in accessing to and assessing information (S_Net) under study (S_Net) of solution 2) and 3) (S_Net) from this strategy (S_Net) of the model to test the second hypothesis (S_Net) of these results of the present research of the election of Fellows (S_Econ) of the situation (S_Econ) of our findings (S_Pov) of this instrument (S_Pov)
linguistic	The definition The general conclusion The answers The answers the motto This proposal His comment	of Q_f (S_Rand) of the present paper from the 'econometric people' (S_Econ) The answers against Black (S_Econ) of the Commission (S_Econ) for selection criteria (S_Econ) to Morgenstern's suggestion (S_Econ)
mental	The notion an alternative strategy an alternative strategy the described strategy the aim The purpose a consensus the aim The analysis The fundamental concerns his own ideas an alternative view The solution The aim	of ordinal Nash equilibrium (S_Rand) for f (S_Rand) for w (S_Rand) in the example above (S_Rand) of the last section (S_Rand) of this paper (S_Rand) about the effects of social change on trust (S_Net) of this paper (S_Net) on networks (S_Net) of corporate governance on these topics (S_Econ) on the role of theory, of empiric models and of the economic research (S_Econ) of the divergence (S_Econ) of this paper (S_Pov)
modal	a sufficient condition the discretionary power the decisive steps the requirements the proposed criteria	for an ordinal Nash equilibrium in the game $(P, \tilde{?}, P)$ (S_Rand) of the CEO in establishing the Board in the creation of the econometric movement (S_Econ) for the choice of a Fellow (S_Econ) for selection (S_Econ)
eventive	A possible modification	in the BLCM (S_Pov)

Appendix E Circumstances realised with shell nouns

E.1 Distribution of Circumstances in Published and Submissions

Feature	Published		Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
CIRCUMSTANCE-TYPE	N=38		N=65					
enhancing-circ	84.2%	32	80.0%	52	0.53		0.28	
extending-circ-(accompaniment)	5.3%	2	1.5%	1	1.08		1.18	
elaborating-circ-(role)	0.0%	0	6.2%	4	0.00		2.43	
projection-circ	10.5%	4	12.3%	8	0.27		0.07	
ENHANCING-CIRC-TYPE	N=32		N=52					
extent	0.0%	0	0.0%	0	0.00		0.00	
location	40.6%	13	28.8%	15	1.11		1.24	
manner	12.5%	4	28.8%	15	1.75	+	3.02	+
cause	31.2%	10	17.3%	9	1.49		2.20	
contingency	15.6%	5	25.0%	13	1.01		1.03	
CAUSE-TYPE	N=10		N=9					
reason	80.0%	8	88.9%	8	0.51		0.28	
purpose	20.0%	2	11.1%	1	0.51		0.28	
behalf	0.0%	0	0.0%	0	0.00		0.00	
PROJECTION-CIRC-TYPE	N=4		N=8					
matter	75.0%	3	87.5%	7	0.51		0.30	
angle	25.0%	1	12.5%	1	0.51		0.30	

E.2 Textual functions of Circumstances realised with shell noun

Feature	Circumstance in Published		Circumstance in Submissions					
	Percent	N	Percent	N	T Stat	Signif.	ChiSqu	Signif.
TEXTUAL-FUNCTION-TYPE	N=38		N=65					
theme-selection	21.1%	8	43.1%	28	2.30	++	5.12	++
n-rheme	68.4%	26	49.2%	32	1.91	+	3.59	+
(other)	10.5%	4	7.7%	5	0.49		0.24	
THEME-SELECTION	N=8		N=28					
topical-theme	100.0%	8	100.0%	28	0.00		0.00	
textual-theme	0.0%	0	0.0%	0	0.00		0.00	
predicated-theme	0.0%	0	0.0%	0	0.00		0.00	
TOPICAL-THEME-TYPE	N=8		N=28					
unmarked-theme	0.0%	0	0.0%	0	0.00		0.00	
marked-theme	100.0%	8	100.0%	28	0.00		0.00	
N-RHEME-TYPE	N=26		N=32					
not-internal	57.7%	15	56.2%	18	0.11		0.01	
clause-complex-internal	42.3%	11	43.8%	14	0.11		0.01	

E.3 Distribution of elements tagged at group rank in Circumstances in Published and Submissions

Feature	Circumstance in Published		Circumstance in Submissions					
	Percent	N	Percent	N	T Stat	Signif.	ChiSqu	Signif.
GROUP-EXP-FUNCTION	N=107		N=165					
deictic	30.8%	33	28.5%	47	0.42		0.17	
post-deictic	3.7%	4	7.3%	12	1.21		1.46	
numerative	5.6%	6	4.8%	8	0.28		0.08	
epithet	2.8%	3	3.6%	6	0.37		0.14	
classifier	1.9%	2	3.6%	6	0.84		0.71	
thing	35.5%	38	40.6%	67	0.84		0.71	
qualifier	19.6%	21	11.5%	19	1.85	+	3.40	+
DEIXIS	N=33		N=47					
specific	93.9%	31	87.2%	41	0.98		0.97	
non-specific	6.1%	2	12.8%	6	0.98		0.97	
SPECIFIC-TYPE	N=31		N=41					
personal-possession	12.9%	4	9.8%	4	0.42		0.18	
demonstrative-proximity	87.1%	27	90.2%	37	0.42		0.18	
PERSONAL-POSSESSION-TYPE	N=4		N=4					
pronominal	100.0%	4	0.0%	0	0.00		8.00	+++
embedded-nominal-group	0.0%	0	100.0%	4	0.00		8.00	+++
DEMONSTRATIVE-PROXIMITY-TYPE	N=27		N=37					
non-selective	77.8%	21	48.6%	18	2.43	+++	5.56	+++
selective	22.2%	6	51.4%	19	2.43	+++	5.56	+++
NUMBER	N=6		N=19					
singular	83.3%	5	73.7%	14	0.46		0.23	
plural	16.7%	1	26.3%	5	0.46		0.23	
DISTANCE	N=6		N=19					
near	83.3%	5	100.0%	19	1.87	+	3.30	+
far	16.7%	1	0.0%	0	0.00		3.30	+
QUALIFIER-TYPE	N=21		N=19					
projected	81.0%	17	78.9%	15	0.15		0.03	
expanded	19.0%	4	21.1%	4	0.15		0.03	
PROJECTED-TYPE	N=17		N=15					
finite	58.8%	10	40.0%	6	1.05		1.13	
non-finite	41.2%	7	60.0%	9	1.05		1.13	
NON-FINITE-TYPE	N=7		N=9					
perfective	28.6%	2	66.7%	6	1.53		2.29	
of+-imperfective	71.4%	5	33.3%	3	1.53		2.29	
EXPANDED-TYPE	N=4		N=4					

embedded-clause	50.0%	2	50.0%	2	0.00		0.00	
prepositional-phrase	50.0%	2	50.0%	2	0.00		0.00	
EMBEDDED-CLAUSE-TYPE	N=2		N=2					
non-finite-cl	100.0%	2	100.0%	2	0.00		0.00	
finite-cl	0.0%	0	0.0%	0	0.00		0.00	
NON-FINITE-CL-TYPE	N=2		N=2					
imperfective-cl	0.0%	0	50.0%	1	0.00		1.33	
neutral-cl	100.0%	2	50.0%	1	1.00		1.33	
prep+-imperfective-cl	0.0%	0	0.0%	0	0.00		0.00	
perfective-cl	0.0%	0	0.0%	0	0.00		0.00	

Appendix F Textual functions at the discourse semantics stratum

F.1 Reference in Published and Submissions

Feature	Published		Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
REFERENCE	N=284		N=406					
neutralised	7.7%	22	12.6%	51	2.03	++	4.10	++
effected	92.3%	262	87.4%	355	2.03	++	4.10	++
EFFECTED-TYPE	N=262		N=355					
presenting	24.0%	63	23.1%	82	0.27		0.08	
presuming	76.0%	199	76.9%	273	0.27		0.08	

F.2 Distance of lexicalisation from shell noun in Published and Submissions

Feature	Published		Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	Feature				
DISTANCE	N=284		N=406					
esphoric: within-group	26.8%	76	24.9%	101	0.56		0.31	
esphoric: within-clause	19.7%	56	9.6%	39	3.83	+++	14.39	+++
anaphoric/cataphoric: adjoining-clause	42.3%	120	41.9%	170	0.10		0.01	
anaphoric/cataphoric: >clause-away	11.3%	32	23.6%	96	4.16	+++	16.94	+++

F.3 Textual functions of anaphorically referring shell nouns in Published and Submissions

Feature	Anaphoric nominal groups in Published		Anaphoric nominal groups in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
RERERENCE: EFFECTED-TYPE	N=88		N=155					
presenting	4.5%	4	7.7%	12	0.96		0.93	
presuming	95.5%	84	92.3%	143	0.96		0.93	
PRESUMING-TYPE	N=84		N=143					
directed	66.7%	56	44.8%	64	3.25	+++	10.20	+++
undirected	27.4%	23	39.9%	57	1.91	+	3.61	+
possessive-det	6.0%	5	15.4%	22	2.13	++	4.49	++
POSSESSIVE-DET- TYPE	N=5		N=22					
researcher	80.0%	4	86.4%	19	0.35		0.13	
nonresearcher	20.0%	1	13.6%	3	0.35		0.13	
SUPERSET	N=84		N=143					
+superset	1.2%	1	11.9%	17	2.92	+++	8.29	+++
-superset	98.8%	83	88.1%	126	2.92	+++	8.29	+++
ORDER	N=1		N=17					
ordinal	0.0%	0	29.4%	5	0.00		0.41	
positional	100.0%	1	70.6%	12	0.61		0.41	
LEXICALISATION	N=90		N=172					
lexicalisation	100.0%	90	100.0%	172	0.00		0.00	
DIRECTION	N=90		N=172					
esphoric	0.0%	0	0.0%	0	0.00		0.00	
anaphoric	100.0%	90	100.0%	172	0.00		0.00	
catphoric	0.0%	0	0.0%	0	0.00		0.00	
DISTANCE	N=90		N=172					
within-group	0.0%	0	0.0%	0	0.00		0.00	
within-clause	4.4%	4	4.7%	8	0.08		0.01	
adjoining-clause	73.3%	66	53.5%	92	3.17	+++	9.72	+++
>clause-away	22.2%	20	41.9%	72	3.21	+++	10.00	+++
ADJOINING-CLAUSE- TYPE	N=66		N=92					
in-clause-complex	13.6%	9	13.0%	12	0.11		0.01	
in-adjoining-sentence	86.4%	57	87.0%	80	0.11		0.01	
(REALISATION)	N=90		N=172					
shell-nominal-group	8.9%	8	12.2%	21	0.81		0.66	
as-clause	88.9%	80	86.0%	148	0.65		0.42	
equation	2.2%	2	0.0%	0	0.00		3.85	++
(other)	0.0%	0	1.7%	3	0.00		1.59	

AS-CLAUSE-TYPE	N=37		N=46					
projected-clause	0.0%	0	0.0%	0	0.00		0.00	
fact-clause	2.7%	1	2.2%	1	0.15		0.02	
clause	86.5%	32	78.3%	36	0.96		0.94	
>clause	0.0%	0	0.0%	0	0.00		0.00	
dem-ref+clause	10.8%	4	19.6%	9	1.09		1.19	
>CLAUSE-TYPE	N=43		N=102					
clause-complex	39.5%	17	29.4%	30	1.19		1.41	
across-sentences	60.5%	26	70.6%	72	1.19		1.41	

F.4 Coupling of anaphorically referring shell nouns with interpersonal meaning in Published and Submissions

Feature	Anaphorically referring shell nouns in Published		Anaphorically referring shell nouns in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
INTERPERSONAL-TYPE	N=101		N=177					
appraisal	80.2%	81	78.0%	138	0.44		0.19	
unaccounted	19.8%	20	22.0%	39	0.44		0.19	
APPRAISAL	N=81		N=138					
attitude	49.4%	40	34.1%	47	2.25	++	5.01	++
graduation	9.9%	8	10.9%	15	0.23		0.05	
engagement	40.7%	33	55.1%	76	2.06	++	4.19	++
ATTITUDE-TYPE	N=40		N=47					
affect	2.5%	1	2.1%	1	0.11		0.01	
judgement	2.5%	1	6.4%	3	0.86		0.74	
appreciation	95.0%	38	91.5%	43	0.64		0.41	
JUDGEMENT-TYPE	N=1		N=3					
social-esteem	0.0%	0	66.7%	2	0.00		1.33	
social-sanction	100.0%	1	33.3%	1	1.00		1.33	
SOCIAL-ESTEEM-TYPE	N=0		N=2					
normality	0.0%	0	0.0%	0	0.00		0.00	
capacity	0.0%	0	0.0%	0	0.00		0.00	
tenacity	0.0%	0	100.0%	2	0.00		0.00	
SOCIAL-SANCTION-TYPE	N=1		N=1					
veracity-(truth)	0.0%	0	0.0%	0	0.00		0.00	
propriety-(ethics)	100.0%	1	100.0%	1	0.00		0.00	
APPRECIATION-TYPE	N=38		N=43					
valuation	34.2%	13	51.2%	22	1.54		2.36	
complexity	63.2%	24	46.5%	20	1.50		2.25	
reaction	2.6%	1	2.3%	1	0.09		0.01	
EXPLICITNESS	N=40		N=47					
inscribe	65.0%	26	66.0%	31	0.09		0.01	
invoke	35.0%	14	34.0%	16	0.09		0.01	
INSCRIBE-TYPE	N=26		N=31					
infused	26.9%	7	61.3%	19	2.71	+++	6.73	+++
isolated	73.1%	19	38.7%	12	2.71	+++	6.73	+++
ATTITUDE-POLARITY	N=40		N=47					
positive-attitude	42.5%	17	70.2%	33	2.68	+++	6.79	+++

negative-attitude	57.5%	23	29.8%	14	2.68	+++	6.79	+++
GRADUATION-TYPE	N=8		N=15					
force	50.0%	4	53.3%	8	0.15		0.02	
focus	50.0%	4	46.7%	7	0.15		0.02	
FORCE-TYPE	N=4		N=8					
intensifying	50.0%	2	25.0%	2	0.82		0.75	
quantifying	50.0%	2	75.0%	6	0.82		0.75	
QUANTIFYING-TYPE	N=2		N=6					
a-thing	100.0%	2	33.3%	2	1.73		2.67	
a-process	0.0%	0	66.7%	4	0.00		2.67	
A-THING-TYPE	N=2		N=2					
number	50.0%	1	0.0%	0	0.00		1.33	
mass/presence	50.0%	1	100.0%	2	1.00		1.33	
A-PROCESS-TYPE	N=0		N=4					
extent	0.0%	0	100.0%	4	0.00		0.00	
frequency	0.0%	0	0.0%	0	0.00		0.00	
EXTENT-TYPE	N=0		N=4					
distance	0.0%	0	75.0%	3	0.00		0.00	
scope	0.0%	0	25.0%	1	0.00		0.00	
FOCUS-TYPE	N=4		N=7					
valeur	75.0%	3	14.3%	1	2.29	++	4.05	++
fulfilment	25.0%	1	85.7%	6	2.29	++	4.05	++
VALEUR-TYPE	N=3		N=1					
authenticity	0.0%	0	0.0%	0	0.00		0.00	
specificity	100.0%	3	100.0%	1	0.00		0.00	
FULFILMENT-TYPE	N=1		N=6					
completion	0.0%	0	66.7%	4	0.00		1.56	
actualisation	100.0%	1	33.3%	2	1.20		1.56	
GRADUATION-TYPE2	N=8		N=15					
up-scale	62.5%	5	46.7%	7	0.70		0.52	
down-scale	37.5%	3	53.3%	8	0.70		0.52	
ENGAGEMENT-TYPE	N=33		N=76					
monogloss	0.0%	0	2.6%	2	0.00		0.88	
heterogloss	100.0%	33	97.4%	74	0.94		0.88	
MONOGLOSS-TYPE	N=0		N=2					
presume	0.0%	0	100.0%	2	0.00		0.00	
assert	0.0%	0	0.0%	0	0.00		0.00	
HETEROGLOSS-TYPE	N=33		N=74					
contract	9.1%	3	10.8%	8	0.27		0.07	
expand	90.9%	30	89.2%	66	0.27		0.07	

CONTRACT-TYPE	N=3		N=8					
disclaim	0.0%	0	12.5%	1	0.00		0.41	
proclaim	100.0%	3	87.5%	7	0.59		0.41	
DISCLAIM-TYPE	N=0		N=1					
deny	0.0%	0	100.0%	1	0.00		0.00	
counter	0.0%	0	0.0%	0	0.00		0.00	
PROCLAIM-TYPE	N=3		N=7					
concur	0.0%	0	0.0%	0	0.00		0.00	
reinforce	33.3%	1	71.4%	5	1.08		1.27	
endorse	66.7%	2	28.6%	2	1.08		1.27	
CONCUR-TYPE	N=0		N=0					
affirm	0.0%	0	0.0%	0	0.00		0.00	
concede	0.0%	0	0.0%	0	0.00		0.00	
REINFORCE-TYPE	N=1		N=5					
pronounce	100.0%	1	40.0%	2	1.00		1.20	
justify	0.0%	0	60.0%	3	0.00		1.20	
EXPAND-TYPE	N=30		N=66					
entertain	43.3%	13	57.6%	38	1.29		1.68	
attribute	56.7%	17	42.4%	28	1.29		1.68	
ATTRIBUTE-TYPE	N=17		N=28					
acknowledge	76.5%	13	89.3%	25	1.14		1.32	
distance2	23.5%	4	10.7%	3	1.14		1.32	

F.5 Textual functions of cataphorically lexicalised shell nouns in major clauses in Published and Submissions

Feature	Cataphorically lexicalised shell nouns in Published		Cataphorically lexicalised shell nouns in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
REFERENCE: EFFECTED-TYPE	N=45		N=68					
presenting	46.7%	21	50.0%	34	0.34		0.12	
presuming	53.3%	24	50.0%	34	0.34		0.12	
PRESUMING-TYPE	N=24		N=34					
directed	16.7%	4	2.9%	1	1.86	+	3.36	+
undirected	58.3%	14	76.5%	26	1.47		2.16	
possessive-det	25.0%	6	20.6%	7	0.39		0.16	
POSSESSIVE-DET-TYPE	N=6		N=7					
researcher	83.3%	5	71.4%	5	0.47		0.26	
nonresearcher	16.7%	1	28.6%	2	0.47		0.26	
SUPERSET	N=24		N=34					
+superset	20.8%	5	35.3%	12	1.19		1.42	
-superset	79.2%	19	64.7%	22	1.19		1.42	
ORDER	N=5		N=12					
ordinal	20.0%	1	25.0%	3	0.21		0.05	
positional	80.0%	4	75.0%	9	0.21		0.05	
LEXICALISATION	N=45		N=68					
lexicalisation	100.0%	45	100.0%	68	0.00		0.00	
DIRECTION	N=45		N=68					
esphoric	0.0%	0	0.0%	0	0.00		0.00	
anaphoric	0.0%	0	2.9%	2	0.00		1.35	
catphoric	100.0%	45	97.1%	66	1.16		1.35	
DISTANCE	N=45		N=68					
adjoining-clause	80.0%	36	75.0%	51	0.61		0.38	
>clause-away	20.0%	9	25.0%	17	0.61		0.38	
ADJOINING- CLAUSE-TYPE	N=36		N=51					
in-clause-complex	22.2%	8	45.1%	23	2.23	++	4.81	++
in-adjoining-sentence	77.8%	28	54.9%	28	2.23	++	4.81	++
(REALISATION)	N=45		N=68					
shell-nominal-group	0.0%	0	1.5%	1	0.00		0.67	
as-clause	100.0%	45	98.5%	67	0.81		0.67	
equation	0.0%	0	0.0%	0	0.00		0.00	
(other)	0.0%	0	0.0%	0	0.00		0.00	
AS-CLAUSE-TYPE	N=3		N=12					
projected-clause	0.0%	0	0.0%	0	0.00		0.00	

fact-clause	0.0%	0	8.3%	1	0.00		0.27	
clause	100.0%	3	91.7%	11	0.49		0.27	
>CLAUSE-TYPE	N=42		N=55					
clause-complex	23.8%	10	21.8%	12	0.23		0.05	
across-sentences	76.2%	32	78.2%	43	0.23		0.05	

F.6 Cataphorically lexicalised shell nouns in major clauses in Published and Submissions

	Published		Submissions	
	Shell noun	N	Shell noun	N
factual	reason result example feature case advantage problem dilemma evidence corollary proof	7 3 4 3 3 1 1 1 1 1 1	proof result example implication reason finding problem corollary aspect corroboration element paradox miracle	10 9 4 4 3 2 1 1 1 1 1 1 1
linguistic	remark answer explanation comment argument conclusion recommendation news tidings manifesto dictum claim	5 2 1 1 1 1 1 1 1 1 1 1	proposition remark suggestion definition statement observation explanation answer comment argument proposal conclusion	12 5 2 1 1 1 1 1 1 1 1 1
mental	analysis theorem view motivation hypothesis idea assumption objective	3 3 2 2 1 1 1 1	hypothesis strategy lemma analysis concern issue idea surprise view	6 3 2 1 1 1 1 1 1
modal	step	1	step task condition requirement	2 2 1 1
eventive	choice improvement	2 1	refinement initiative	1 1
	way limitation constraint method	1 1 1 1	way criterion approach field situation	2 2 1 1 1

F.7 Distribution of shell noun types in hyper-Themes in Published and Submissions

Feature	Hyper-Themes in Published		Hyper-Themes in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
SHELL-NOUN	N=15		N=16					
factual	60.0%	9	37.5%	6	1.24		1.57	
linguistic	6.7%	1	18.8%	3	0.99		1.01	
mental	13.3%	2	18.8%	3	0.40		0.17	
modal	0.0%	0	6.2%	1	0.00		0.97	
eventive	13.3%	2	0.0%	0	0.00		2.28	
circumstantial	6.7%	1	18.8%	3	0.99		1.01	

F.8 Distribution of interpersonal functions in shell nouns projecting field as object of study in Qualifier

	Shell nouns projecting field as object of study in Published		Shell nouns projecting field as object of study in Submissions	
	N	%	N	%
INTERPERSONAL FUNCTIONS	N=66		N=120	
ATTITUDE	14	21.2	46	38.3
ENGAGEMENT	29	43.9	21	17.5
GRADUATION	23	34.9	53	44.2
ATTITUDE	N=14		N=46	
AFFECT	2	14.2	1	2.2
JUDGEMENT	9	64.2	35	76.1
APPRECIATION	3	21.4	10	21.7
ENGAGEMENT	N=29		N=21	
CONTRACTION	16	55.2	8	38.1
EXPAND: entertain	4	13.8	13	61.9
EXPAND: attribute	9	31.0	0	0.0
GRADUATION	N=23		N=53	
Force	9	39.1	19	35.8
Focus: fulfilment	14	60.9	34	64.2
NO INTERPERSONAL FUNCTION	N=1		N=7	

Due to tagging procedures, I was unable to use Corpus Tool to test the statistical significance of these figures.

Appendix G Interpersonal functions of shell nouns at the discourse semantics stratum

G.1 Distribution of the contribution of shell nouns as a resource in ENGAGEMENT in Published and Submissions

Feature	Shell noun in Published		Shell noun in Submissions					
	Percent	N	Percent	N	T Stat	Signif.	ChiSqu	Signif.
ENGAGEMENT-TYPE	N=165		N=190					
monogloss	0.0%	0	1.6%	3	0.00		2.63	
heterogloss	100.0%	165	98.4%	187	1.62		2.63	
MONOGLOSS-TYPE	N=0		N=3					
presume	0.0%	0	100.0%	3	0.00		0.00	
assert	0.0%	0	0.0%	0	0.00		0.00	
HETEROGLOSS-TYPE	N=165		N=187					
contract	40.0%	66	28.3%	53	2.32	++	5.32	++
expand	60.0%	99	71.7%	134	2.32	++	5.32	++
CONTRACT-TYPE	N=66		N=53					
disclaim	10.6%	7	7.5%	4	0.57		0.33	
proclaim	89.4%	59	92.5%	49	0.57		0.33	
DISCLAIM-TYPE	N=7		N=4					
deny	100.0%	7	100.0%	4	0.00		0.00	
counter	0.0%	0	0.0%	0	0.00		0.00	
PROCLAIM-TYPE	N=59		N=49					
concur	0.0%	0	0.0%	0	0.00		0.00	
reinforce	88.1%	52	93.9%	46	1.02		1.05	
endorse	11.9%	7	6.1%	3	1.02		1.05	
CONCUR-TYPE	N=0		N=0					
affirm	0.0%	0	0.0%	0	0.00		0.00	
concede	0.0%	0	0.0%	0	0.00		0.00	
REINFORCE-TYPE	N=52		N=46					
pronounce	71.2%	37	56.5%	26	1.51		2.28	
justify	28.8%	15	43.5%	20	1.51		2.28	
EXPAND-TYPE	N=99		N=134					
entertain	54.5%	54	69.4%	93	2.34	++	5.40	++
attribute	45.5%	45	30.6%	41	2.34	++	5.40	++
ATTRIBUTE-TYPE	N=45		N=41					
acknowledge	84.4%	38	90.2%	37	0.80		0.65	
distance2	15.6%	7	9.8%	4	0.80		0.65	

G.2 Distribution of interpersonal selections from APPRECIATION with nominal groups with shell noun as Head in Published and Submissions

Feature	Shell noun in Published		Shell noun in Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
APPRECIATION-TYPE	N=96		N=90					
valuation	53.1%	51	53.3%	48	0.03		0.00	
complexity	38.5%	37	41.1%	37	0.36		0.13	
reaction	8.3%	8	5.6%	5	0.74		0.55	
EXPLICITNESS	N=96		N=91					
inscribe	72.9%	70	60.4%	55	1.82	+	3.28	+
invoke	27.1%	26	39.6%	36	1.82	+	3.28	+
INSCRIBE-TYPE	N=70		N=55					
infused	28.6%	20	43.6%	24	1.76	+	3.06	+
isolated	71.4%	50	56.4%	31	1.76	+	3.06	+
ATTITUDE-POLARITY	N=96		N=91					
positive-attitude	57.3%	55	76.9%	70	2.90	+++	8.12	+++
negative-attitude	42.7%	41	23.1%	21	2.90	+++	8.12	+++

G.3 Distribution of interpersonal functions for shell nouns construing JUDGEMENT

Feature	Published		Submissions		T Stat	Signif.	ChiSqu	Signif.
	Percent	N	Percent	N				
JUDGEMENT-TYPE	N=11		N=51					
social-esteem	36.4%	4	39.2%	20	0.17		0.03	
social-sanction	63.6%	7	60.8%	31	0.17		0.03	
SOCIAL-ESTEEM-TYPE	N=4		N=20					
normality	0.0%	0	5.0%	1	0.00		0.21	
capacity	100.0%	4	85.0%	17	0.80		0.69	
tenacity	0.0%	0	10.0%	2	0.00		0.44	
SOCIAL-SANCTION-TYPE	N=7		N=31					
veracity-(truth)	57.1%	4	71.0%	22	0.70		0.51	
propriety-(ethics)	42.9%	3	29.0%	9	0.70		0.51	
EXPLICITNESS	N=11		N=51					
inscribe	100.0%	11	98.0%	50	0.46		0.22	
invoke	0.0%	0	2.0%	1	0.00		0.22	
INSCRIBE-TYPE	N=11		N=50					
infused	100.0%	11	98.0%	50	0.47		0.22	
isolated	0.0%	0	2.0%	1	0.00		0.22	

ATTITUDE-POLARITY	N=11		N=51					
positive-attitude	90.9%	10	92.2%	47	0.14		0.02	
negative-attitude	9.1%	1	7.8%	4	0.14		0.02	

Appendix H Modification of Head in shell nominal groups

H.1 GRADUATION as FORCE: intensifying in nominal groups with shell noun as Head in Published and Submissions

Shell noun	Published	Submissions
factual	at least two <i>key</i> features of non-profits (P_CEO) <i>Roughly</i> the same results (P_CEO) Among the <i>main</i> empirical findings (P_CEO) our <i>main</i> result (P_Rand) the <i>main</i> problem (P_Rand)	Three <i>main</i> aspects (S_Net) some of the <i>main</i> findings (S_Pov)
linguistic		a <i>more detailed</i> explanation (S_Net) The <i>main</i> conclusion (S_CEO)
mental	The <i>primary</i> motivation (P_Pov) The <i>main</i> objective of our analysis (P_Pov) its <i>primary</i> purpose (P_Pov)	our <i>main</i> goal (S_Rand) The <i>fundamental</i> concerns of corporate governance (P_CEO) their <i>main</i> function (P_CEO) the <i>central</i> aim of determining their socio-economic characteristics (P_CEO) The <i>dominant</i> concept (S_Econ)
modal		<i>higher</i> potentials of developing trustworthy relationships (S_Net) the <i>urgent</i> need for a code of governance practice to be introduced in this sector (S_CEO) a <i>reduced</i> capacity to possess an adequate heating system and to provide general expenses (S_Pov)
eventive		the <i>strenuous</i> efforts to combine historical memory with the promotion of technical expertise (S_Econ)
circumstantial		an <i>extremely</i> strong condition to fulfil (P_S_Rand) a <i>more extensive</i> approach to identify different dimensions of well-being needs (S_Pov)

H.2 GRADUATION as FORCE: quantity in nominal groups with shell noun as Head in Published and Submissions

Shell noun	Published	Submissions
factual	<p><i>several</i> advantages (P_CEO) <i>a variety of</i> reasons(P_CEO) <i>at least two</i> reasons (P_CEO) the <i>added</i> benefit (P_CEO) <i>at least two</i> key features (P_CEO) <i>several</i> reasons (P_CEO) <i>One</i> reason (P_CEO) This <i>large</i> effect (P_CEO) <i>some</i> evidence (P_CEO) <i>several</i> advantages (P_Pov) <i>many</i> reasons (P_Net) <i>more</i> evidence (P_Net) <i>One of</i> the most curious aspects (P_Econ) <i>a little-known</i> fact (P_Econ) <i>some brief</i> biographical evidence (P_Econ) <i>One of</i> the reasons (P_Rand)</p>	<p><i>some</i> results (S_Rand) this <i>additional</i> advantage (S_Net) <i>two</i> reasons (S_Net) <i>Three</i> main aspects (S_Net) <i>Further</i> corroboration (S_Net) <i>Four</i> elements (S_Net) the <i>present</i> result (S_Net) <i>a battery of</i> reasons (S_Econ) The <i>first</i> advantage (S_Pov) The <i>second</i> advantage (S_Pov) <i>some of</i> the main findings (S_Pov)</p>
linguistic	<p><i>some</i> specific recommendations (P_Pov) <i>some</i> conclusions (P_Net) <i>some</i> bad news (P_Econ) the <i>perennial</i> complaint (P_Econ) the <i>further</i> bad tidings (P_Econ) <i>none of</i> those appeals (P_Econ) the <i>accompanying</i> assertions (P_Econ) <i>one</i> conclusion (P_Rand)</p>	<p><i>some</i> concluding remarks (S_Rand) <i>Two</i> remarks (S_Rand) <i>a couple of</i> remarks (S_Rand) <i>some</i> conclusions (S_Pov)</p>
mental	<p><i>one</i> objective (P_Pov) <i>a number of</i> important structural assumptions (P_Pov) the <i>customary</i> assumption (P_Pov) <i>at least two</i> motivations (P_Pov) <i>a number of</i> points of view (P_Econ) an <i>early</i> disinclination (P_Econ)</p>	<p><i>one</i> of the clearest lessons (S_Rand) an <i>alternative</i> strategy (S_Rand) the <i>shortage of</i> incentives (S_Net) a <i>well-enough</i> discussed subject (S_Net) the <i>present</i> purpose (S_Net) <i>several</i> issues (S_CEO) these <i>three</i> goals (S_Econ) an <i>alternative</i> view (S_Econ) a <i>growing</i> concern (S_Pov) <i>several</i> strategies (S_Pov) The <i>most usual</i> strategy (S_Pov) An <i>alternative</i> strategy (S_Pov)</p>
modal	<p><i>one</i> possibility (P_CEO) a <i>tendency</i> (P_Net)</p>	<p><i>higher</i> potentials (S_Net) The <i>high</i> average probability (S_Net) a <i>low</i> average probability (S_Net) the <i>low</i> potential (S_Net) <i>high</i> probabilities (S_Net) the <i>higher</i> probability (S_Net) <i>Several</i> steps (S_CEO)</p>

eventive	<p>several improvements (P_Pov) the <i>first</i> serious attempt (P_Net) two choices (P_Econ)</p>	<p>his <i>only</i> alternative (S_Rand) <i>One</i> potential course of action (S_Rand) these <i>early</i> efforts (S_Econ) an <i>immense</i> success (S_Econ) the strenuous efforts (S_Econ) <i>many</i> attempts (S_Pov) A possible modification (S_Pov)</p>
circumstantial	<p>two ways (P_Pov) <i>both</i> approaches (P_Pov) <i>a host of other</i> potential ways (P_CEO) This seemingly strange occurrence (P_Net) the radically wrong method (P_Econ) the three weak responsiveness conditions (P_Rand)</p>	<p>two ways (S_Rand) <i>one</i> way (S_Rand) <i>another</i> way (S_Net) <i>Alternative</i> approaches (S_CEO) <i>three</i> different fields (S_Econ) two criteria (S_Econ) Morgenstern's proposed restriction (S_Econ) the proposed criteria (S_Econ) a more <i>extensive</i> approach (S_Pov)</p>

H.3 FOCUS: valeur in nominal groups with shell noun as Head

Shell noun	Published	Submission
factual	the <i>distinct</i> advantage that uncertainty about factor scores can be accurately summarized ... (P_Pov) <i>Distinct</i> institutional features (P_CEO) A <i>classic</i> example (P_CEO) the <i>mere</i> fact (P_Econ) <i>just</i> this phenomenon (P_Econ)	the <i>particular</i> case (S_Rand) <i>Each of</i> these cases (S_Net)
linguistic	some <i>specific</i> recommendations (P_Pov)	<i>hardly</i> an encouragement (P_Econ) <i>even less</i> a commitment (P_Econ) a <i>particular</i> remark (S_Net) The <i>general</i> conclusion (S_CEO)
mental		his <i>own</i> ideas (S_Econ)
modal		the State's <i>actual</i> ability (S_Net)
eventive		
circumstantial	a <i>general</i> approach (P_Pov) <i>each of</i> these shortcomings (P_Pov)	This <i>very</i> situation (S_Net) <i>just</i> a way (S_Net)

H.4 FOCUS: Fulfilment: actualisation in nominal groups with shell noun as Head

Shell noun	Published	Submissions
factual	The first <i>possible</i> reason (P_CEO) Another <i>possible</i> reason (P_CEO) not <i>necessarily</i> the case (P_Net)	
linguistic		
mental		
modal	the <i>perceived</i> ability of familial, ethnic and religious networks to reduce moral hazard (P_Net)	
eventive	this <i>expected</i> trend (P_Net)	One <i>potential</i> course of action (S_Rand) A <i>possible</i> modification in the BLCM (S_Pov)
circumstantial	a host of other <i>potential</i> ways (P_CEO)	

Note: Appendix H.4 excludes instances such as **possibility** or **propensity** in which it is the shell noun that enacts the interpersonal meaning of fulfilment.

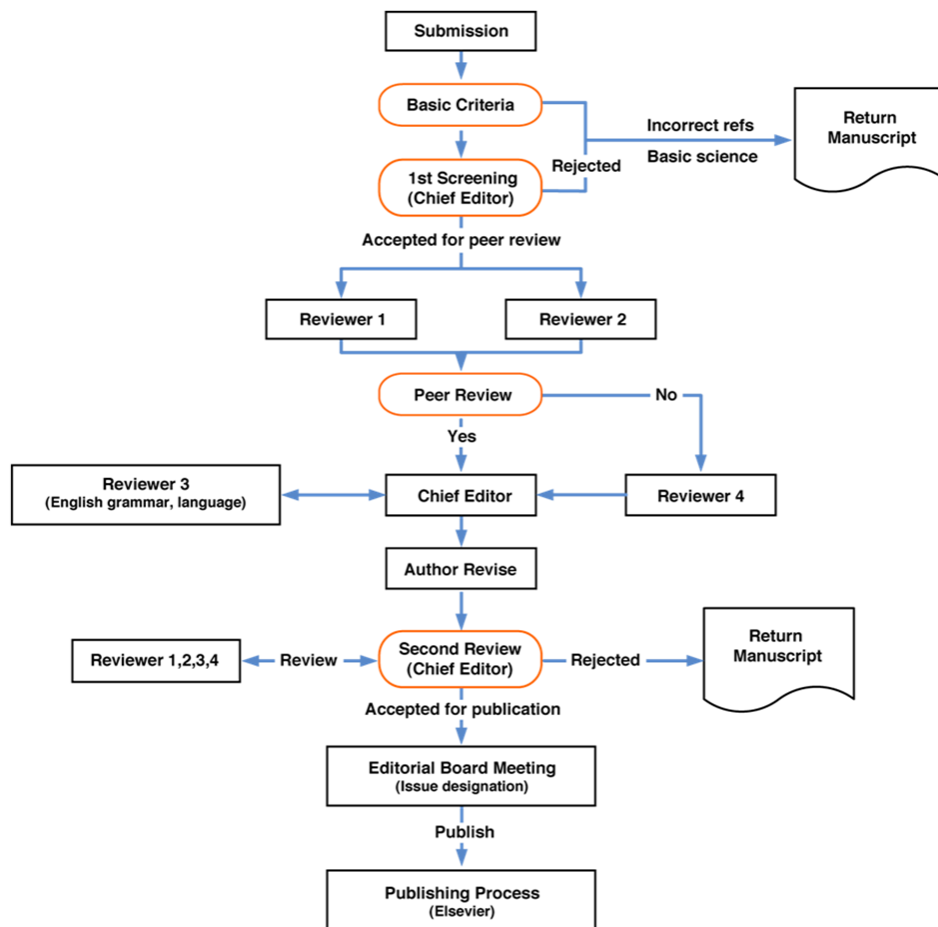
Annexes

List of Annexes

Annex A	Publication process for Elsevier	394
A.1.	Peer-review process flowchart.....	394
Annex B	The corpus	395
B.1	P_Rand	395
B.2	P_Net.....	407
B.3	P_Econ.....	429
B.4	P_Pov.....	446
B.5	P_CEO	459
B.6	S_Rand.....	474
B.7	S_Net	492
B.8	S_Econ	511
B.9	S_Pov	528
B.10	S_CEO	539

Annex A Publication process for Elsevier

A.1. Peer-review process flowchart



Source: <http://www.elsevier.com/reviewers/reviewer-guidelines>

Annex B The corpus

B.1 P_Rand

Klaus, Bettina & Klijn, Flip:

Paths to stability for matching markets with couples

1. Introduction

We consider entrance level labor markets that can be modeled as two-sided matching problems: one side consisting of students, graduates, or workers, the other side consisting of residencies, jobs, or firms. Since in the last couple of decades the number of couples with the same professional interests has been growing, we focus on labor markets in which both members of couples seek positions. A famous example of such an entrance level labor market with couples in the US is the market where each year thousands of medical school graduates seek their first employment as residents or interns. Since the 1950s this labor market has been administered by the National Resident Matching Program (NRMP).¹ One of the reasons for centralizing this market, and later for reorganizing it,² was that market outcomes did not seem to be ‘stable’ as indicated by unraveling (pre-NRMP) or a significant reduction of voluntary participation in the NRMP (pre-reorganized NRMP). For simplicity, from now on we refer to one side of the market as students and to the other side of the market as hospitals, even though we do not exclusively restrict our attention to centralized markets such as the NRMP

Loosely speaking, an outcome or matching is stable if there are no students/couples and hospitals that are not matched with each other, but in fact would prefer to be. Thus, it is easy to see that whenever contracts are not enforceable, stability is a minimal requirement for the persistence of an outcome for any market (centralized or decentralized). Gale and Shapley (1962), the ‘pioneers’ of two-sided matching theory, provided an algorithm, the famous deferred acceptance (DA) algorithm, that always finds a stable matching in markets with only single students. The algorithm first used by the NRMP in fact was a slight modification of the deferred acceptance algorithm, which was independently developed. Hence, a stable matching can be implemented for any centralized ‘singles market.’ Moreover, Roth and Vande Vate (1990) showed that a stable matching can also be reached by means of decentralized decision making. They demonstrated that, starting from an arbitrary matching, the process of allowing randomly chosen blocking coalitions to match will converge to a stable matching with probability one. Thus, as long as we restrict attention to singles markets, in these markets—centralized or not—stability is possible. Unfortunately, once couples emerge on one side of the market, stability is in danger: stable matchings may not exist (Roth, 1984) and it may be very difficult (i.e., NP-hard) to decide if stable matchings exist for a given ‘couples market’ (Ronn, 1990).

Klaus and Klijn (2005a, 2005b) showed that for a special preference domain for couples, namely the domain of ‘(weakly) responsive’ preferences, stable matchings exist. A couple’s preferences are responsive if the unilateral improvement of one partner’s job is considered beneficial for the couple as well. If responsiveness only applies to acceptable positions, then preferences are weakly responsive. Hence, (weakly) responsive preferences may reflect situations where couples search for jobs

in the same metropolitan area (if one partner switches to a job he/she prefers and the couple can still live together, then the couple is better off). In fact, when couples have weakly responsive preferences, one can derive associated individual preferences for all students that are members of a couple. Klaus and Klijn (2005a, 2005b) demonstrated that any stable matching of an associated singles market is also stable in the couples market.

Klaus and Klijn (2005a, 2005b) also showed that under a restricted unemployment aversion condition, the domain of weakly responsive preferences is maximal for the existence of stable matchings. In view of this result, the class of weakly responsive preferences is a natural starting point for the study of decentralized decision making in couples markets. One may wonder whether there is a path of ‘satisfying’ blocking coalitions that leads to a stable matching. We answer this question in the affirmative. This implies in particular that starting from an arbitrary matching, the process of allowing randomly chosen blocking coalitions to match will converge to a stable matching with probability one. It also shows that to obtain a stable matching by decentralized means there is no need to first determine associated individual preferences of the members of the couples.

Obviously, weakly responsive preferences are quite restrictive because they (a) almost take us back to the model with only singles and (b) do not capture important aspects such as a couple’s desire to live together (see also Remark 4.2). Thus, the next step in our analysis is to try to extend our result to more general preferences. Since then stable matchings may not always exist, particularly if couples preferences are based on distance considerations (e.g., Klaus and Klijn, 2005a, Example 3.8), we check for ‘conditional path convergence,’ that is, we only require the process of starting with an arbitrary matching and allowing randomly chosen blocking pairs to match to converge to a stable matching with probability one if a stable matching exists. For one-sided matching problems, more precisely for roommate problems, Diamantoudi et al. (2004) obtained such a conditional path convergence result.³ Unfortunately, in couples markets conditional convergence cannot be obtained. It turns out that for couples markets with non-responsive preferences, even in the presence of stable matchings, there may not exist a path induced by satisfying blocking coalitions that yields a stable matching. Hence, this ‘cycling’ is inherent in the underlying complementarities in the couples’ preferences, and not so much a problem related to our path construction.

The paper is organized as follows. In Section 2, we introduce a couples market where the labor market modeled consists of a supply side of hospitals and a demand side of couples composed of medical students. In Section 3 we state and prove our result that, for weakly responsive preferences and starting from an arbitrary matching, a stable matching can always be obtained through at least one path of blocking coalitions, for instance, by using our deterministic path algorithm for weakly responsive couples markets (DPC-algorithm). Furthermore, we remark on some features of the DPC-algorithm and discuss differences between our and Roth and Vande Vate’s (1990) path construction. In Section 4 we provide a couples market with non-responsive preferences and a unique stable matching. We show that there are initial matchings starting from which the stable matching cannot be obtained by satisfying blocking coalitions.

2. Matching with couples and stability

For convenience and without loss of generality, we describe a couples market where the labor market modeled consists of a supply side of hospitals and a demand side of couples composed of medical students; $H = \{h_1, h_2, \dots, h_m\}$ and $C = \{c_1, c_2, \dots, c_n\} = \{(s_1, s_2), (s_3, s_4), \dots, (s_{2n-1}, s_{2n})\}$ are the sets of hospitals and couples. We denote the set of students by $S = \{s_1, s_2, \dots, s_{2n}\}$. Each hospital has exactly one position to be filled. Our results can easily be adapted to more general situations that include single students (Remark 3.6) and hospitals with multiple positions.⁴ Furthermore, our results (as well as the results derived in Klaus and Klijn, 2005a, 2005b) do not depend on the students being specifically paired in couples. All results would remain the same if the setting would be extended to triples or any other constellation of ordered tuples of students. Such an extension of the model could include mutual externalities between more than two individuals, for instance between members of a family or among friends or colleagues.

Next, we describe preferences of hospitals and couples.

Hospitals' preferences: Each hospital $h \in H$ has a strict, transitive, and complete preference relation \succsim_h over the set of students S and the prospect of having its position unfilled, denoted by \emptyset . Let $s \in S$. If $s \succsim_h \emptyset$, then student s is acceptable to hospital h ; if $\emptyset \succ_h s$, then student s is unacceptable to hospital h . Hospital h 's preferences can be represented by a strict ordering of the elements in $S \cup \{\emptyset\}$; for instance, $P(h) = (s_4, s_2, \emptyset, \dots)$ indicates that hospital h prefers student s_4 to s_2 , and considers the other students to be unacceptable. Let $P_H = \{P(h)\}_{h \in H}$.

Couples' preferences: Each couple $c \in C$ has a strict, transitive, and complete preference relation \succsim_c over all possible combination of ordered pairs of (different) hospitals and the prospect of being unemployed. Couple c 's preferences can be represented by a strict ordering of the elements in $H := [(H \cup \{u\}) \times (H \cup \{u\})] \setminus \{(h, h) : h \in H\}$. To simplify notation, we denote a generic element of H by (h_p, h_q) , where h_p and h_q indicate a hospital or being unemployed. For instance, $P(c) = (h_4, h_2), (h_3, h_4), (h_4, u)$, etc., indicates that couple $c = (s_1, s_2)$ prefers s_1 and s_2 being matched to h_4 and h_2 , respectively, to being matched to h_3 and h_4 , respectively, and so on. Let $P_C = \{P(c)\}_{c \in C}$.

Whenever we use the strict part \succ of a preference relation, we assume that we compare different elements in $S \cup \{u\}$ or H . Note that as for many classical matching markets with singles the assumption that agents' preferences are strict is also crucial for matching markets with couples (consider for instance Roth and Sotomayor, 1990, Section 2.5).

Couples markets: A one-to-one matching market with couples, or a couples market for short, is denoted by (P_H, P_C) .

Matchings: A matching μ for a couples market (P_H, P_C) is an assignment of students and hospitals such that each student is assigned to at most one hospital in H or to u (which can be assigned to multiple students), each hospital in H is assigned to at most one student or to \emptyset (which can be assigned to multiple hospitals), and a student is assigned to a hospital if and only if the hospital is assigned to the student. If $\mu(s) = u$ ($\mu(h) = \emptyset$), then student s (hospital h) is called unmatched. If $\mu(s) = h$, or equivalently $\mu(h) = s$, then student s and hospital h are called mates.

A matching μ is completely described by the list $\mu(S) = (\mu(s_1), \mu(s_2), \dots, \mu(s_{2n}))$ of hospitals or u assigned to students s_1, s_2, \dots, s_{2n} . Thus, s_k is matched to $\mu(s_k)$.

Stability: Finally, we define stability for couples markets (see Roth and Sotomayor, 1990). First, if a hospital whose position is filled can be better off by unmatching, i.e., leaving its position unfilled, we would expect the hospital to act accordingly. Similarly, if one or both partners in a couple can improve the given matching for the couple by unmatching, i.e., voluntarily choosing unemployment, we would expect the couple to act accordingly. Formally, a hospital $h \in H$ is a one-sided blocking coalition for matching μ if

(o1) $\mu_h(h) \prec h \succ \emptyset$.

A couple $c = (s, sl) \in C$ is a one-sided blocking coalition for matching μ if

(o2) $(\mu(s), u) \prec_c (\mu(s), \mu(sl))$, $(u, \mu(sl)) \prec_c (\mu(s), \mu(sl))$, or $(u, u) \prec_c (\mu(s), \mu(sl))$.

We say that $(c, (\mu(s), u))$, $(c, (u, \mu(sl)))$, or $(c, (u, u))$ is a one-sided blocking coalition (depending on the condition in (o2) that is satisfied).

Second, if one partner in a couple can improve the given matching for the couple by switching to another hospital such that this hospital is better off as well, then we would expect this mutually beneficial trade to be carried out, rendering the given matching unstable. A similar statement holds if both students in the couple can improve. For a given matching μ , $[c = (s, sl), (hp, hq)]$ is a two-sided blocking coalition if $(hp, hq) \succ \{(\mu(s), u), (u, \mu(sl)), (u, u)\}$ and

(t1) $(hp, hq) \prec_c (\mu(s), \mu(sl))$;

(t2) $[hp \in H \text{ implies } \mu_{hp}(hp) \prec hp]$ and $[hq \in H \text{ implies } \mu_{hq}(hq) \prec hq]$.

A matching is stable if there is no blocking coalition.

(In)stability in couples markets: Roth (1984, Theorem 10) showed that stable matchings may not exist in the presence of couples. To make things worse, it can be very 'hard' to decide if a couples market has a stable matching: Ronn (1990) showed that determining if a couples market has a stable matching is an NP complete problem. Klaus and Klijn (2005a, 2005b) established the existence of a stable matching when couples have so-called weakly responsive preferences. In the next section, we introduce this restricted preference domain and show that on this domain and starting from an arbitrary matching, a stable matching can always be obtained through at least one path of blocking coalitions. Clearly, one would be happy to establish similar results for other preference domains, e.g., if preferences capture that couples do want to live together. However, Klaus and Klijn (2005a, 2005b, Theorem 3.5 and Corollary 3.6) essentially demonstrate that none of these preference domains can even guarantee the existence of stable matchings. Furthermore, in Section 4, we show that on the general domain of couples' preferences the existence of stable matchings does not guarantee the existence of any 'path to stability' through blocking coalitions.

3. Weak responsiveness and paths to stability

First, we introduce the domain of weakly responsive preferences (see Klaus and Klijn, 2005a, 2005b for more details). A couple's preferences are responsive if the unilateral improvement of one partner's job is considered beneficial for the couple as well. If responsiveness only applies to acceptable positions, then preferences are weakly responsive.

Weakly responsive preferences: Couple $c = (s_k, s_l)$ has weakly responsive preferences if there exist strict, transitive, and complete preference relations $_sk$ and $_sl$ over the set of hospitals and the prospect of being unemployed u such that (i) for all $h \in H$,

$(u, h) _c (u, u)$ if and only if $h _sl u$,

$(h, u) _c (u, u)$ if and only if $h _sk u$,

(ii) for all $h_p, h_q, h_r \in H \setminus \{u\}$,

$[h_p _sk u, h_q _sl u, \text{ and } h_p _sk h_r \text{ imply } (h_p, h_q) _c (h_r, h_q)]$ and

$[h_p _sl u, h_q _sk u, \text{ and } h_p _sl h_r \text{ imply } (h_q, h_p) _c (h_q, h_r)]$, and

(iii) for all $h_ , h_ \in H, h_ = h_ , u _sk h_ \text{ and } u _sl h_ \text{ imply } (u, u) _c (h_ , h_)$.

A couples' market (P, H, PC) is weakly responsive if all couples have weakly responsive preferences.

Remark 3.1 (Weak responsiveness and associated individual preferences). Let (P, H, PC) be a weakly responsive couples market. Then, for all $c = (s_k, s_l)$, the associated individual preferences $_sk$ and $_sl$ are only uniquely determined for acceptable positions. In other words, if both $[_sk \text{ and } _sl]$ and $[_sk \text{ and } _sl]$ satisfy the three weak responsiveness conditions, then for all $h_p, h_q \in H \setminus \{u\}$, $h_p _sk h_q _sk u$ implies $h_p _sk h_q _sk u$, and $h_p _sl h_q _sl u$ implies $h_p _sl h_q _sl u$.

Let (P, H, PC) be a weakly responsive couples market. Then, from the couples' weakly responsive preferences we can determine associated individual preferences for all students that are members of a couple. By $(P, H, PS(P, C))$ we denote an associated singles market⁵ obtained by replacing couples and their preferences in (P, H, PC) by individual students and their (possibly not uniquely determined) associated individual preferences $PS(P, C)$. It is easy to see that all associated singles markets have the same set of stable matchings (see Remark 3.1). The following existence theorem is due to Klaus and Klijn (2005a, 2005b, Theorem 3.3).

Theorem 3.2 (Stability for weakly responsive preferences (Klaus and Klijn, 2005a, 2005b)).

Let (P, H, PC) be a weakly responsive couples market. Then, any matching that is stable for an associated singles market $(P, H, PS(P, C))$ is also stable for (P, H, PC) . In particular, there exists a stable matching for (P, H, PC) .

However, not all stable matchings for (P, H, PC) may be stable for $(P, H, PS(P, C))$. The intuition is that some matching that would be unstable in a singles market is now stable because a student may not want to block it by taking the position of his/her partner.

Finally, Klaus and Klijn (2005a, 2005b, Theorem 3.5) showed that under a restricted unemployment aversion condition, the domain of weakly responsive preferences is maximal for the existence of stable matchings. In view of this result, the class of weakly responsive preferences is a natural starting point for the study of decentralized decision making in couples markets.

Next, we show that for a weakly responsive couples market, starting from any matching, there always exists a path of blocking coalitions that ends in a stable

matching. To describe this process of satisfying blocking coalitions, we use the following terminology.

Satisfying blocking coalitions: If a hospital $h \in H$ is a one-sided blocking coalition for a matching μ , then we say that a new matching μ' is obtained from μ by satisfying the blocking coalition if h and $\mu(h)$ are unmatched, and all other agents are matched to the same mates at μ' as they were at μ . Formally, matching μ' is obtained from matching μ by satisfying blocking coalition h

(for μ) if

- $\mu'(h) = \emptyset$ and $\mu'(\mu(h)) = h$;
- $\mu'(s) = \mu(s)$ for all $s \in S \setminus \{\mu(h)\}$;
- $\mu'(\cdot \setminus h) = \mu(\cdot \setminus h)$ for all $\cdot \setminus h \in H \setminus \{h\}$.

Similarly, if $[(sk, sl), (hp, hq)]$ is a blocking coalition for a matching μ , then we say that a new matching μ' is obtained from μ by satisfying the blocking coalition if (sk, sl) and (hp, hq) are matched to one another at μ' , their mates at μ (if any, and if not involved in the blocking coalition) are unmatched at μ' , and all other agents are matched to the same mates at μ' as they were at μ . Formally, matching μ' is obtained from matching μ by satisfying blocking coalition

$[(sk, sl), (hp, hq)]$ (for μ) if

- $[\mu(sk) = h \in H \setminus \{hp, hq\} \text{ implies } \mu(h) = \emptyset]$ and $[\mu(sl) = h \in H \setminus \{hp, hq\} \text{ implies } \mu(h) = \emptyset]$;
- $[\mu(hp) = s \in S \setminus \{sk, sl\} \text{ implies } \mu(s) = \emptyset]$ and $[\mu(hq) = s \in S \setminus \{sk, sl\} \text{ implies } \mu(s) = \emptyset]$;
- $\mu(sk) = hp, \mu(sl) = hq, [hp \in H \text{ implies } \mu(hp) = sk]$, and $[hq \in H \text{ implies } \mu(hq) = sl]$;
- $\mu'(s) = \mu(s)$ for all $s \in S \setminus \{\mu(hp), \mu(hq), sk, sl\}$;
- $\mu'(h) = \mu(h)$ for all $h \in H \setminus \{\mu(sk), \mu(sl), hp, hq\}$.

We are now ready to state our main result.

Theorem 3.3 (Paths to stability for weakly responsive preferences). Let (P, H, PC) be a weakly responsive couples market. Let μ be an arbitrary matching for (P, H, PC) . Then there exists a finite sequence of matchings μ_1, \dots, μ_k such that $\mu_1 = \mu$, μ_k is stable, and for all $i = 1, \dots, k-1$, there is a blocking coalition for μ_i such that μ_{i+1} is obtained from μ_i by satisfying this blocking coalition.

Consider a random process that begins by selecting an arbitrary matching μ and then proceeds to generate a sequence of matchings $\mu = \mu_1, \mu_2, \dots$ where each μ_{i+1} is obtained from μ_i by satisfying a blocking coalition, chosen at random from the blocking coalitions for μ_i . Assume that the probability that any particular blocking coalition for μ_i will be chosen to generate μ_{i+1} is positive, and only depends on the matching μ_i (but not on the number i). Let $R(\mu)$ be the random sequence generated in this way from an initial matching μ .

Corollary 3.4 (Random paths to stability for weakly responsive preferences). Let (P, H, PC) be a weakly responsive couples market. For any initial matching μ for (P, H, PC) , the random sequence $R(\mu)$ converges with probability one to a stable matching.

To prove Theorem 3.3, we construct an algorithm that, in a finite number of steps, yields a stable matching for any weakly responsive couples markets.

Before describing our 'deterministic path algorithm' for weakly responsive couples markets, we present how it would look like in singles markets. The advantage of this approach is that (a) the basic structure of the algorithm is easier to understand for singles markets and (b) it is easier to see which additional steps have to be taken to accommodate the presence of couples in the market. In both algorithms that we present now, we will use a virtual room that agents enter and exit throughout the algorithm. This didactic visualization aid was first introduced by Ma (1996).

A modified deterministic Roth and Vande Vate path algorithm to stability for singles markets Let μ be an arbitrary matching for singles market (P, H, PS) . After satisfying one-sided blocking coalitions for μ (first stage) we start putting the students one by one in an initially empty room (second stage). Each student enters with his/her mate under μ , i.e., a student enters with a matched hospital or alone. Whenever a student enters the room with his/her mate, blocking coalitions within the room are satisfied and the hospitals that are 'dumped' (i.e., unmatched while blocking coalitions are satisfied) are put outside of the room. Thus, after this second stage we obtain a matching where all students are matched to hospitals in the room, and for which there are no blocking coalitions within the room. Finally, in the third stage, we let the hospitals outside of the room enter one by one. In each step possibly a blocking coalition within the room has to be satisfied before turning to the next step. A hospital that is dumped is put outside of the room. The blocking coalitions that are satisfied in this stage are 'hospital-optimal,' in the sense that for the hospital involved there is no other blocking coalition available within the room that would give it a better student. In each step the students are not worse off and the number of hospitals in the room does not decrease. Also, in each step there is a student that is strictly better off or the number of hospitals in the room increases strictly. Hence, after a finite number of steps all hospitals have joined the students in the room. Starting from μ we have obtained a stable matching for singles market (P, H, PS) through the satisfaction of blocking coalitions. Note that in difference to Roth and Vande Vate's (1990) original algorithm, in our modified version we use 'dumping' of hospitals instead of choosing optimal blocking coalitions to avoid cycles of blocking coalitions in the room in the second stage. Also, Roth and Vande Vate (1990) do not differentiate between a stage in which the focus lies on the entering students (second stage) and a stage in which the focus lies on the entering hospitals (third stage). For couples markets both steps will play an important role. Henceforth we refer to the modified Roth and Vande Vate algorithm presented above as MRV-algorithm.

A deterministic path algorithm to stability for weakly responsive couples markets Let μ be an arbitrary matching for a weakly responsive couples market (P, H, PC) . After satisfying one-sided blocking coalitions for μ (first stage) we start putting the couples one by one in an initially empty room (second stage). Each couple enters with their mates under μ . Whenever a couple enters the room with their mates, blocking coalitions within the room are satisfied and the hospitals that are 'dumped' are put outside of the room. Thus, after this second stage we obtain a matching where all couples are matched to hospitals in the room, and for which there are no blocking coalitions within the room. Finally, in the third stage, we let the hospitals outside of

the room enter one by one. In each step possibly a blocking coalition within the room has to be satisfied before turning to the next step. The blocking coalitions that are satisfied in this stage are 'hospital-optimal,' in the sense that for the hospital involved there is no other blocking coalition available within the room that would give it a better student. We call the student that is in all the hospital optimal blocking coalitions associated with the entering hospital the best student.

There may be several blocking coalitions that match the entering hospital with the best student. In order to assure the convergence of the algorithm we have to choose the blocking coalition carefully. First, if matching the hospital with the best student while keeping the partner's match constant is improving the matching for the couple, then we satisfy the associated blocking coalition. However, it may be that the couple together with the hospital and the old match of the best student's partner does not form a blocking coalition. In this case, we prove (see the CLAIM in the third stage of the DPC-algorithm and its proof in Appendix A) that the couple together with the hospital and the best student's previous match forms a blocking coalition. We then match the hospital with its best student and the best student's partner with the best student's previous match.

In each step the couples are not worse off and the number of hospitals in the room does not decrease. Also, in each step there is a couple that is strictly better off or the number of hospitals in the room increases strictly. Hence, after a finite number of steps all hospitals have joined the couples in the room. Starting from μ we have obtained a stable matching for couples market (P, H, PC) .

Henceforth we refer to this deterministic path algorithm for weakly responsive couples markets as DPC-algorithm. Note that an additional complication in the DPC-algorithm compared to the MRV-algorithm is the specific choice of hospital optimal blocking coalitions in Stage 3. We now formalize the DPC-algorithm.

Deterministic path algorithm to stability for weakly responsive couples markets or DPC-algorithm.

Input: A weakly responsive couples market (P, H, PC) and a matching μ for (P, H, PC) .

Initialization: Set $A := \emptyset$. We call A the room.

First Stage:

Satisfy all one-sided blocking coalitions (if any) and denote the matching obtained by μ_1 .

After Stage 1 we obtain a matching $\mu_1 := \mu$ without one-sided blocking coalitions.

Second Stage:

If there exists $c = (s_1, s_2) \in C \setminus A$, then put the couple and its mates in the room, i.e., • Set $A := (A \cup \{c, (s_1), (s_2)\}) \setminus \{u\}$.

As long as there is a blocking coalition $[c = (s_1, s_2), (h_1, h_2)]$ with $\{c, h_1, h_2\} \cap A \neq \emptyset$ do:

Begin Loop: Satisfy $[c, (h_1, h_2)]$, and put hospitals that are dumped outside of the room:

- For $i = 1, 2$, [if $(s_i) = h \in H \setminus \{h_1, h_2\}$ then define $\mu(h) := \mu$ and set $A := A \setminus \{h\}$];

- For $i = 1, 2$ with $h_{i-1} \in H$, [if $h_{i-1} = s \in S \setminus \{s_1, s_2\}$ then $h(s) := u$];
- Define $h(c_i) := (h_{i-1}, h_i)$ and for $i = 1, 2$, [if $h_{i-1} \in H$ then $h(h_{i-1}) := s_{i-1}$].
End Loop.

After Stage 2 we obtain a matching $\mu_2 := \mu$ where all couples are in the room and no blocking coalitions exist in the room.

Third Stage:

As long as there exists $h \in H \setminus A$ do:

Begin Loop: Set $A := A \cup \{h\}$. If there is no blocking coalition $[c_{i-1}, (h_{i-1}, h_i)]$ with $h_{i-1} \in \{h_1, h_2\} \cap A \cup \{u\}$, then GO BACK to the beginning of the Third Stage.

If there are blocking coalitions $[c_{i-1}, (h_{i-1}, h_i)]$ with $h_{i-1} \in \{h_1, h_2\} \cap A \cup \{u\}$, then let s_{i-1} be h_{i-1} 's most preferred mate among the ones it could get at these blocking coalitions. Let s_i be the partner of s_{i-1} .

Without loss of generality, $c_{i-1} = (s_{i-1}, s_i) \in C$.

CLAIM: $[c_{i-1}, (h_{i-1}, s_i)]$ or $[c_{i-1}, (h_{i-1}, s_{i-1})]$ is a blocking coalition.

Case (a): If $[c_{i-1}, (h_{i-1}, s_i)]$ is a blocking coalition, then satisfy it and if a hospital is dumped (at most one!), then put it outside of the room:

- Let $h := h_{i-1}$. If $h \in H$, then define $h(h) := s_i$ and set $A := A \setminus \{h\}$;
- Define $h(c_{i-1}) := (h_{i-1}, s_i)$.

Case (b): Satisfy blocking coalition $[c_{i-1}, (h_{i-1}, s_{i-1})]$ (CLAIM) and if a hospital is dumped (at

most one!), then put it outside of the room:

- Let $h := s_{i-1}$. If $h \in H$, then define $h(h) := s_{i-1}$ and set $A := A \setminus \{h\}$;
- Define $h(c_{i-1}) := (h_{i-1}, s_{i-1})$.

End Loop.

After Stage 3 we obtain a matching $\mu_3 := \mu$ where all couples and all hospitals are in the room and no blocking coalitions exist in the room.

Output: A stable matching μ for (P, H, PC) .

Note that throughout the DPC-algorithm every new matching is obtained by satisfying a blocking coalition. The proof that the DPC-algorithm is well-defined and produces a stable matching for any weakly responsive couples market in a finite number of steps is given in Appendix A. In the following example we visualize the DPC-algorithm, showing the process of satisfying blocking coalitions.

Example 3.5 (An application of the DPC-algorithm). Let (P, H, PC) be the couples market given by Table 1. The sets of hospitals and couples are given by $H = \{h_1, \dots, h_9\}$ and $C = \{(s_1, s_2), (s_3, s_4), (s_5, s_6), (s_7, s_8)\}$, respectively. Agents that are not depicted are unacceptable.

For instance, h_1 's preference list is given by $P(h_1) = s_4, s_1, u, \dots$. It is easy to check that all couples' preferences are weakly responsive: associated individual preferences are $P(s_1) = h_3, h_1, h_6, u, \dots$, $P(s_2) = h_2, u, \dots$, $P(s_3) = h_2, h_8, h_6, u, \dots$, $P(s_4) =$

$h_4, h_2, h_3, h_1, u, \dots$, $P(s_5) = h_3, h_5, u, \dots$, $P(s_6) = h_6, h_3, u, \dots$, $P(s_7) = h_3, h_7, u, \dots$, and $P(s_8) = h_8, h_4, h_3, u, \dots$. In Table 2 we apply the algorithm to the initial matching $\mu(S) = (h_1, h_2, h_7, h_8, h_5, h_6, h_3, h_9)$. On the left hand side we give short explanatory comments that guide through the algorithm (we abbreviate the term ‘blocking coalition’ by ‘b.c.’). On the right hand side we depict the matching at each point of the algorithm (whenever a hospital stands below a student it is matched to this student). The vertical bar represents the door of the room: the agents on the left are inside the room and the agents on the right are outside the room. We obtain a path of matchings, each of them being the result of satisfying a blocking coalition for the previous matching. The output is the stable matching $\mu^*(S) = (h_1, u, h_2, h_4, h_5, h_6, h_3, h_8)$. It is tedious, but not difficult, to check that μ^* is the unique stable matching for (P, H, PC) .

Remark 3.6 (The DPC-algorithm and single students). One easily accommodates single students in the DPC-algorithm by letting each single student have a ‘fictitious partner’ that finds all hospitals unacceptable. For example, if single student s has (strict) preferences given by $P(s) = h_1, h_2, u, h_3, \dots$, then replace s by couple $c = (s, s_-)$ (where s_- is a fictitious partner) with (strict) preferences $P(c) = (h_1, s_-), (h_2, u), (u, u), (h_3, u), \dots$. Since the fictitious couples have weakly responsive preferences, one can directly apply the DPC-algorithm to obtain a matching that is stable in the original matching market.

Remark 3.7 (Which matchings can be reached by the DPC-algorithm?). Clearly, if starting from any matching is allowed, all stable matchings can be reached (just start with a stable matching). In particular, the DPC-algorithm can converge to stable matchings that are not stable in the associated singles markets. If in the second stage we allow couples to enter the room with a chosen blocking coalition, then even starting from the empty matching where all students are unemployed and all hospitals’ positions are unfilled we can reach any stable matching for the weakly responsive couples market in question (let all couples c choose their match $\mu(c)$ in a given stable matching μ will trivially induce that stable matching).

Remark 3.8 (The DPC-algorithm and Roth and Vande Vate’s (1990) algorithm). Note that our deterministic path algorithm for couples markets (the DPC-algorithm) is not a straightforward adaptation of the Roth and Vande Vate’s (1990) deterministic path algorithm for singles markets (the RV-algorithm, for short). The second stage of the DPC-algorithm where hospitals that are dumped in the process of satisfying blocking coalitions are put outside the room until the third stage has no corresponding part in the RV-algorithm. Furthermore, even though the loop construction in the third stage uses the same idea of ‘optimally’ satisfying blocking coalitions as in the RV-algorithm, our construction is more complex in order to accommodate special ‘blocking structures’ for couples.

The main problem with any attempt to apply the RV-algorithm directly to weakly responsive couples markets is that blocking coalitions for the associated singles market may in fact not induce permitted blocking coalitions in the couples market and that vice versa a blocking coalition for the couples market may not induce a permitted blocking coalition in the associated singles market. We illustrate this by means of a simple example. Let (P, H, PC) be a couples market with $H = \{h_1, h_2\}$, $C = \{c = (s_1, s_2)\}$, $P(h_1) = s_1, s_2, ?$, $P(h_2) = s_2, s_1, ?$, and $P(c) = (h_1, h_2), (h_2, h_1), (h_1, u), (u, h_2), (u, h_1), (h_2, u), (u, u)$. Clearly $\mu^*(S) = (h_1, h_2)$ is the only stable matching. In the associated singles

market $(P, H, PS(P, C))$, couple c and PC is replaced by 'single' students s_1 and s_2 with $P(s_1) = h_1, h_2, u$ and $P(s_2) = h_2, h_1, u$. If we apply the RV-algorithm to matching $\mu(S) = (h_2, h_1)$, it will converge to $\mu(S)$, but only by satisfying blocking coalitions that do not induce permitted blocking coalitions in the couples market: Starting from $\mu(S)$ as a matching for $(P, H, PS(P, C))$, the only blocking coalitions are $[s_1, h_1]$ and $[s_2, h_2]$. However, satisfying blocking coalition $[s_1, h_1]$ ($[s_2, h_2]$) implies matching $\mu^*(S) = (h_1, u)$ ($\mu^*(S) = (u, h_2)$), which is worse for couple c than the initial matching $\mu(S)$. On the other hand, starting from $\mu(S)$ as a matching for (P, H, PC) , blocking coalition $[(s_1, s_2), (h_1, h_2)]$ would induce satisfying two blocking coalitions in the associated singles market $(P, H, PS(P, C))$ simultaneously, which is not permitted in the RV-algorithm.

If our aim was to extend the RV-algorithm to couples markets as straightforwardly as possible, then we could merge our second and third stage by immediately satisfying blocking coalitions of hospitals that are dumped in the second stage, using the loop construction of the third stage. However, by separating the second and third stage, the proof that the DPC-algorithm is welldefined and converges in a finite number of steps to a stable matching is more transparent.

Remark 3.9 (Exchanging the roles of couples and hospitals in the DPC-algorithm). If we reverse the roles of couples and hospitals in the algorithm, the proof of convergence becomes somewhat more complicated. The changes of the first and second stage (and the proof that they are welldefined) are minor. However, recall that in the third stage all couples are and remain in the room and hospitals enter one by one. Each time a hospital enters the room, some other hospital may be sent out. Thus, we may have to satisfy a chain of blocking coalitions (see loop of the third stage). Now consider the third stage after the roles of couples and hospitals are exchanged. After one couple enters, its blocking may unmatched two students from two different couples. So instead of one chain of blocking coalitions, we have to satisfy two chains of blocking coalitions, which is still possible, but it requires extra notation and 'logistics.'

Starting from the empty matching and depending on whether all students or all hospitals enter the room first, the RV-algorithm only converges to the hospital-optimal or the student-optimal stable matching. Exchanging the roles of couples and hospitals would also change the set of stable matchings that could be reached by the DPC-algorithm when starting from the empty matching. However, for a couples market there may be no optimal matching for either side of the market (Klaus and Klijn, 2005a, Theorem 4.2).

4. Cycles of blocking coalitions

When couples' preferences are not weakly responsive there may be no stable matching. Obviously, in this case the DPC-algorithm cannot work. However, if a stable matching does exist for a general couples market, then one may wonder whether the DPC-algorithm still converges to a stable matching. More generally, one can ask if, for general couples markets with stable matchings, there exists an algorithm that starts from any matching and converges to a stable matching. Before giving the answer to this question, we briefly discuss the answer to the corresponding question for roommate problems.

An important similarity between roommate problems and couples markets is that stable matchings need not exist. In a recent paper, Diamantoudi et al. (2004) showed

that for roommate problems with stable matchings one can indeed construct an algorithm that will, starting from any matching, converge to a stable matching. Since an existing stable matching is explicitly used in the algorithm constructed by Diamantoudi et al. (2004), the algorithm is non-constructive and therefore differs considerably from Roth and Vande Vate's (1990) algorithm and the DPC algorithm for two-sided matching markets.

In contrast to roommate problems, for couples markets it turns out that there may not exist a path induced by satisfying blocking coalitions that yields a stable matching. In other words, for some couples markets that have stable matchings, there are initial matchings from which no sequence of blocking coalitions leads to a stable matching. Thus, this 'cycling' is inherent in the underlying complementarities in the couples' preferences, and not so much a problem related to the DPC-algorithm presented in the previous section.

Example 4.1 (Non convergence).⁶ Let (P, H, PC) be the couples market with $H = \{h1, h2, h3\}$, $C = \{(s1, s2), (s3, s4)\}$, and preferences given by Table 3. (Again, agents that are not depicted are unacceptable.) It is easy to check that $\mu(S) = (h2, h3, u, u)$ is the unique stable matching.

In Fig. 1 we have depicted all possible paths obtained from satisfying blocking coalitions if we start with matching $\mu_1(S) = (h1, h2, u, u)$. Clearly, the path is unique and turns out to be a cycle. Moreover, the cycle does not contain μ . Hence, $R(\mu)$ as defined in Section 3 'cycles,' i.e., does not converge to a stable matching.

Remark 4.2 (The role of weak responsiveness). As mentioned before, the assumption that couples' preferences are weakly responsive is quite restrictive. More realistic domains for couples' preferences would of course include non-responsive preferences that also account for distance considerations between jobs and geographic preferences. We would love to establish more positive results for these domains. However, one conclusion from our analysis is that these results unfortunately cannot be obtained (Example 4.1). Thus, we do not intend to proclaim that weakly responsive preferences are realistic, but that they form the benchmark between 'positive' and 'negative' results in couples markets. We think that even though the results are not as 'nice' as one would wish for, it is still important to describe what happens and where the dividing line between 'positive' and 'negative' results lies.

B.2 P_Net

Haggerty, John & Sherryllene Haggerty

The Lifecycle of a metropolitan business network: Liverpool 1750-1810

Introduction

Historians have become increasingly interested in networks as a tool for historical analysis. Many of these interpretations present a benign view of business networks based on the perceived ability of familial, ethnic and religious networks to reduce moral hazard. More recently, however, historians have complicated their conceptions of networks and have started to look at more formal and civic arenas especially within an urban environment. This research into metropolitan business networks challenges the positivist view of networks. For example, research on guilds and trade associations, using quantitative data similar to that used in this study, has often found them to be reactionary and inward looking to an extent that retarded economic progress. At the same time, many commercial towns remained successful and were able to adapt to new economic opportunities. This suggests a complicated picture of metropolitan business networks.

This paper pushes forward this more nuanced and sophisticated analysis of networks and represents the first serious attempt to measure them to assess change over time.¹ It visualises Liverpool's metropolitan business networks comprising political, trade, social and cultural institutions during the period 1750–1810 and then measures the relationships between these individuals and groups. In doing so, it avoids the issue of source (or ego) centrality often found in network analysis based on qualitative data to provide a wider picture of relationship interplay and its role in economic activity. This is achieved by applying social network analysis and visual analytics to historical data to facilitate analysis of the network relationships. The relationships presented here number over 210,000 amongst 1700 actors, which could not be presented using traditional representational methods. Using different metrics of efficiency based on an actor's or a group's relative position within the network, we demonstrate how the networks changed over time. Moreover, visual analytics, especially when analysing large data sets, is exploratory in nature and therefore may raise questions, rather than answer them per se.

This paper posits that Liverpool's institutional networks during the period 1750–1810 analysed here, coalesced towards bonding rather than bridging networks. This occurred as Liverpool's economy matured, especially with regard to the Atlantic trade, and the slave trade in particular. However, this study also argues that networks are dynamic, fluctuate over time and have a life cycle. We therefore find that various networks were in operation at the same time. This means that, as with regional clusters, some will eventually wither, but this facilitates the emergence of new and more dynamic networks in turn.

This paper first outlines the major themes in the related literature on networks and places our contribution in context. It then outlines the methodology and the eighteenth-century Liverpool case study. This is followed by a discussion on the rise and fall of a section of Liverpool's Metropolitan Business Networks. Finally, we draw our conclusions.

Networks in history

Historians have been using networks as an analytical tool for some time. This is a worthwhile exercise because contemporaries also realised the value of networks, even if they called their contacts 'friends' or 'correspondents' (Hancock, 2005). In particular, business networks are intended to produce various outcomes such as the rapid transfer and synthesis of knowledge or understanding, the conferment of status and legitimacy, lower information and transaction costs, and the ability to adapt to environmental changes (Podolny and Page, 1998). Historians often accept that such networks always performed these positive functions and were easily constructed and maintained ([Davidoff and Hall, 1987], [Prior and Kirby, 1993], [Walvin, 1997] and [Hamilton, 2005]). This may be because much stress has been placed on the 'implicit contract of family' (Ben-Porath, 1980) and ethnic and religious ties, assuming that they reduced moral hazard (Mathias, 2000). It is true that such ties provided reputational mechanisms and specific cultural market information, but ethnic ties could also be inward looking and provide extra obligations ([Greif, 1989], [Mentz, 2004], [Rauch, 2001] and [Haggerty, 2006]). Furthermore, whilst Quakers were perceived to have a highly moralistic business culture, they, like Jews, also faced ostracism ([Doerflinger, 1987], Tolles, 1963 and [Trivellato, 2009]). Indeed, Rauch (2001) has argued that the importance of ethnic networks has been exaggerated due to their observability rather than their importance.

If family, ethnic and religious ties are insufficient to explain business networks, we need to look to more civic ways of networking. Hancock (1996) has shown how merchants came together in order to share skills and capital and has demonstrated that many networks included people who had never met, especially useful considering the distances covered by trade in this period. Clearly then, strong and weak ties worked together in a dynamic way in order for businessmen to conduct their trade. Furthermore, merchants usually belonged to a number of networks at any one time, depending on their fluctuating business interests, and their power or importance within each varied. For example, a merchant might use the tavern or drinking clubs to foster trust relationships for credit, acquire the latest gossip at the coffee house and 'change', but might further career and political interests via a trade association and the town council. Such 'Metropolitan Business Networks' are the 'intangible infrastructure of a region'; consisting wholly or partly of business people, they are significant because their members meet face-to-face, regularly, and are usually efficient because these members are busy men (Casson, 2003). Often articulated as institutions such as guilds

or trade associations in the eighteenth century, these more formalised networks frequently provided a social context for a rational business activity. Interestingly these more institutionalised networks existed as a separate entity, with members coming and going and power relationships shifting over time. This should have allowed them to be flexible and to adapt quickly to a changing environment. Working together on the same issues to defend certain aspects of trade also promoted trust and solidarity and many people became attached to the micro social order of the institutions themselves, as well as to their fellow members ([Baldassarri and Diani, 2007] and [Lawler et al., 2008]). This is not to say that such institutions were without their problems. They often became victims of their own success and protected vested interests. For example, Ogilvie (2004) argues that guilds in German proto-industry inhibited progress and produced friction rather than promoted economic development, and Rosenband (1999) found much the same in his analysis of the early industrial revolution. Crumplin (2007) has also shown that dense metropolitan networks can easily become opaque, which allows cliques to get out of control. Business networks, therefore, are not always efficient.

However, the fact that Liverpool, and indeed many other British metropolitan areas were so successful in this period, suggests that business networks mostly fulfilled their purpose. For example, Pearson and Richardson (2001) argue that Liverpool's permeable networks grew through in-migration in this period, meaning that the trading community was receptive and open, and Greefs (2008) found the same for Antwerp. In this way, international trade was promoted by alleviating problems of contract enforcement and asymmetrical information, and by providing information about new trading opportunities. However, it is also clear that individuals were members of various institutions which linked with various networks, and that all of these were active at the same time. Therefore, whilst one network might have been successful, another one may have been in decline, and still others on the rise at the same time. Therefore, institutionalised networks, like those of individuals, were constantly changing and reforming, as power relationships within them adjusted accordingly. This suggests that, as Swann (1998) found for regional clusters, networks have a life cycle. It is with this dynamism, power relationships and life cycle of such problematic metropolitan business networks that this article is concerned.

Whilst historians and, indeed, social scientists have recently complicated their understandings of networks, they have yet to measure these networks to quantify their dynamics as an aid to analysing actors' or groups' influence in, and reactions to, the wider economic environment. This paper therefore uses visual analytics and graph theory to analyse Liverpool mercantile networks in 1750–1810. Importantly this methodology allows us to visualise and measure the relationships for vast data sets (such as the c. 210,000 analysed here), not possible under 'traditional' representative methods such as graphs or tables.

Case study and methodology

Liverpool 1750–1810 provides an excellent case study for many reasons. Liverpool benefitted greatly from Britain's Atlantic trade and was hailed as the second city of the realm during this period (Montefiore, 1804). Liverpool was the leader in the slave trade after 1750, traded with Nova Scotia, throughout the thirteen colonies/states, and had excellent connections with the British West Indies as well as a vibrant coastal trade. It was the largest outpost in this period with tonnage through the port increasing ([Hyde, 1971] and [Marriner, 1982]). Port development also played an important part in helping the city economically, the first wet dock being opened in 1715 (Power, 1997). This success encouraged in-migration and Liverpool's population grew in consequence. In 1708, the population of Liverpool was around 6500 (Lawton, 1953). By the time of the first census in 1801 it was 77,653 (British Parliamentary Papers, 1801). This rapid rise in population was mirrored by an increase in trade. Imports entering the port amounted to only 14,600 tons in 1709, but had grown to 450,000 tons by 1800 (Marriner, 1982). It also had a variety of formal and informal institutions as befitted its status (Stobart, 2000) providing networking opportunities.

It is their records of membership that will allow us to analyse and measure its metropolitan business networks.

The problem with local and confidential networking is that much of it is not recorded in official records. We do not know what was said after the official meetings, in the Exchange, customs house or tavern for example. We therefore do not always understand the interplay between formal and informal structures and friendship cliques. Institutional records can, however, provide us with the names of individuals involved and therefore the basis for network analysis. We use four of these institutions. Our choice was shaped by the desire to use a range of institutions which would cover merchants in different spheres, various stages of their careers, and also by the longevity of extant data in order to be able to measure change over time.

Town Council (Liverpool Town Books)

Arguably the most formal and civic of the institutions, it was a 'close' (self-electing) corporation and was dominated by elite mercantile men (Power, 1997). It was an Anglican institution very active in protecting and promoting the trade of the town. The records cover the whole period under study.

Committee of the African Merchants Trading from Liverpool (hereafter African Committee) (Committee Book of the African Company of Merchants trading from Liverpool)

This was a formal trade association with links to similar societies in London and Bristol. Run by merchants, this represented one of the important trades of the town. Whilst it was dominated by slave traders, membership was not exclusive to those in that trade.

For example, Edgar Corrie was on the Committee but was against the slave trade. Its records cover the whole period of this study.

Library/Lyceum (Liverpool Lyceum)

First founded in 1758 as the Library, it was later renamed the Lyceum (and is referred to as such throughout this paper). It was an informal cultural institution representative of many of its type in the period. Set up as an elite library, the membership fee was five guineas (£5 5 s) (Wallace, 1795) and among its members were many of the leading liberals and abolitionists of the town, such as William Roscoe and William Rathbone. Its records run from 1769 and continue throughout our period.

Drinking clubs

These were informal institutions with a wide membership of men at varying stages of their careers, mercantile and otherwise. We use two. The Ugly Face Club records (Ugly Face Club) run from 1743 to 1757. Membership of the club included many merchants and ships' captains as well as doctors, a draper and an architect. The Mock Corporation of Sephton (Sephton Mock Corporation) (hereafter Mock Corporation) records run from 1753 to 1796. However, the records are sporadic between 1786 and 1796, with membership recorded only for 1789, 1790 and 1796. Membership was much more diverse than the Ugly Face Club and included merchants and mariners as well as those from many walks of Liverpool life such as potters, corn factors, coopers, grocers, a slater and plasterer, glaziers, and pedlars.

A hundred percent sample was taken of all four institutions for the years for which their records are extant. Fig. 1 illustrates the total number of individual actors recorded across all institutions (removing double counts for multiple institutional membership). As can be seen, the number of actors involved in the clubs in this study rose until the 1770s and then declined despite the introduction of the Lyceum data. However, data available for the Mock Corporation of Sephton is sporadic during the 1780s and 1790s and not available for the 1800s, which affects the total individual actors recorded. This is not to say that actors were not involved in this club, only that data of active participation described above is not extant.

We have also cross referenced institutional data with other sources to provide further information about slave trade investment groups, such as the Trans Atlantic Slave Trade Database (Eltis et al., 2010), trade directories and other qualitative sources such as merchants' letterbooks to illustrate our arguments where appropriate.

Because total actors recorded here number between 200 and 375 at any one time, in a port where the population was approaching 80,000 at the turn of the nineteenth century, this case study is far from being representative of the whole population. Indeed, a study of this period which used a very wide definition of the trading community (including factors, shopkeepers, grocers, etc.) found that community to account for only 2.5% to 3.5% of the whole Liverpool population (Haggerty, 2006).

Moreover, differing assessments by historians and contemporaries as to the precise definition of a merchant (or indeed Hyde et al.'s (1951) merchant of 'substance') makes calculating the representativeness of this sample problematic. However, comparing the numbers of the Town Council and the African Committee which comprised only merchants (see Fig. 2) against numbers of merchants listed in the trade directories (Haggerty, 2006) demonstrates that for the 1760s, 1770s and 1780s, between 42% and 63% of all merchants are represented here. If we include those belonging to the Lyceum and drinking clubs the number is even higher. The high numbers of merchants listed in the trade directories in the 1790s and 1800s mean that this percentage is reduced to between 12.5% and 22.5%. However, given that we are interested in the elite metropolitan business networks, it is likely that we have included most of the significant actors.

All membership records of the Town Council, the African Committee and the Ugly Face Club were recorded. The Mock Corporation had a wide membership but not all were active members. Therefore, only those actors who were recorded as in attendance at meetings or were elected to the council were noted. The Lyceum also had an extremely wide membership, but only those attending committee meetings were recorded. Therefore, active rather than passive participants in these institutions are analysed. In order to make this data manageable we have had to assume that all members listed within one decade met each other and were part of an institutional peer group. This provides the basis for the relationships presented in the analysis below which is therefore grouped by decade. Cross-decade relationships are not analysed here, although we recognise that peer group relationships would not be restricted by our temporal groupings.

Fig. 2 also demonstrates that institutional membership and participation changed throughout the period on an institution-by-institution basis. The number of individual actors belonging to the Town Council (TC), after a brief rise in the 1760s, declined during the period. The number of actors belonging to the African Committee (AC) rose to a high of 107 during the 1780s before declining over the following twenty years. This decline took place during a period characterised by war, credit crises and the abolition of the slave trade in 1807. The Mock Corporation (MC) enjoyed a healthy participation by members until a decline in the 1780s and 1790s (though much of this is due to the lack of extant data). The number of actors actively participating in the Lyceum's (L) committee remained constant throughout the period due to its policy of replacing 50% of the committee's members each year. However, actors who served on the committee could leave and then rejoin the committee at the next available opportunity. Interestingly the rise and decline in institutional membership demonstrated in Fig. 1 and Fig. 2 does not fit neatly with measures of group density, as shown in Table 1. Network density is a widely used measure of group cohesion, measured between 0 and 1, where 0 is an inefficient network and 1 a highly efficient

network (Wasserman and Faust, 1994). Usually, as group sizes increase, the density decreases if the numbers of actor degrees remain unchanged, i.e. the group gets too big for everyone to know each other. However, the network in this study does not follow this expected trend precisely. Between 1750 and 1769, group cohesion as identified by network density rises along with the number of actors. However, this trend is reversed during the next two decades where the number of actors and group cohesion declines. A comparison of the 1790s with the 1800s shows where the number of actors and group cohesion decline. As will be demonstrated below, this supports our argument that Liverpool's metropolitan business networks, as represented here, were more permeable with a bridging quality (Burt, 2004) at the beginning of our period, but coalesced towards more bonding or closed networks towards the end.

In order to further our analysis of these fluctuating networks, we need first to understand exactly how networks are measured. A network is an interconnected group or system and the relations, both logical and physical, between the actors. There is a tendency to assume that just because actors are linked they must form a cohesive social network. However, this is not necessarily the case and the relationships between network members must be explored to fully understand how these networks functioned. In order to analyse complex changes in relationships and behaviour, we analyse an actor's and group's position within the network (or networks) in order to understand and quantify their role, behaviour or potential power through relational information within the wider network(s). We achieve this by using mathematical modelling through graph theory.

It is worthwhile providing a brief overview of graph theory terminology relevant to the analysis. It is not the intention of this paper to present a detailed overview of the field of graph theory. However, it is necessary to explain the methodology used in this study (for detailed explanations see ([Wasserman and Faust, 1994], [Haggerty and Haggerty, 2010] and [Freeman, 1978/79])). A graph comprises a set of points (or vertices) which are connected to other points through edges (or arcs). Two points that are directly connected by an edge are said to be adjacent. The number of other points to which any given point is adjacent is called the degree of that point. The distance between points is calculated by the number of edges in a path. The shortest paths linking pairs of points are called geodesics. Points falling on the geodesics between a given pair of points stand between these points. Fig. 3 provides an example of a small graph.

In Fig. 3, points in the graph are numbered P1 to P4. P1 is adjacent to P2 and P3, but is not adjacent to P4. Points P1 and P2 have degrees of two and P3 has a degree of three. However, due to P4's position on the periphery of the graph, it has a degree of only one. There are two paths from P1 to P4; the first through two other points, P2 and P3, the second directly via P3. The distance is equal to the number of edges in a path. Thus, using the first path P1 has a distance of three to P4 and via the second path, a distance of two. As the second path is shorter than the first, this is the geodesic. As

Fig. 3 illustrates, not all actors are equal. P3 has access to P1, P2 and P4. In contrast both P2 and P1 have access to two other actors (each other and P3), whilst P4 has access to only one other actor, P3. P3 therefore has the best access or 'centrality' having access to all three of the other actors.

Within this paper, we use the following terminology. Network refers to all institutions and actors within the study as well as the relationships therein. An institution (and therefore institutional membership) refers to each of the four associations or clubs. A subnet is a cluster of actors that belonged to two or more institutions and therefore formed their own small groupings through their cross-institutional membership. An actor is an individual member.

We use three measures of centrality to provide an analysis of the network dynamics within our case study. It is not the focus of this analysis to suggest which is the most "correct" measure of a graph. However, a combination of various measures provides an overview of the dynamics of the network from different viewpoints and an appreciation of the actors' or groups' roles therein. All measures used in this paper are those detailed in Wasserman and Faust (1994). These measures are built into the application, SocNetV (Kalamaras, 2009), which has been used to provide the results and visualisations in this paper. The measures of the network are:

Out-degree centrality

This measure analyses the expansiveness, or number of actors that a particular actor possesses or accesses. It therefore measures which institutions or subnets provided the most potential connections within the network. In Fig. 3, this would be P3. Out-degree centrality in this study uses the Actor Out-Degree Centralization (AODC) measure. An AODC value of zero indicates that all actors within the group are equal, as would occur in a regular lattice network. An AODC value of one indicates that a single actor dominates all other actors within the network, as in a star network.

Betweenness centrality

This network measure identifies potential points of control within the network. Those that act as chokepoints (are the most between others) may be centres of power, control and influence because they can choose to retain or share that information. Again, this would be P3 in Fig. 3. Betweenness recognises that communication flow within a given network often does not rely on adjacent actors but moves along geodesics. Because they are focal points of communications within the group and subgroups, certain actors facilitate contact and communications within the network and can therefore be seen as major channels of influence. An Actor Betweenness Centrality (ABC) value falls between zero, indicating that all actors have exactly the same betweenness within the network, and one, when a single actor falls on all geodesics between all the other actors (indicative of a star network).

Closeness centrality

This measure analyses how close an actor is to other actors within the network. Closeness centrality identifies the relative access to information that each actor possessed. Those closer to other actors in the network may be less reliant on a chain of others for their decision-making process and may be more independent. In Fig. 3, P4 is the most dependent. Actors that belong to groupings with a high closeness centrality are likely to have access to the latest information and to be able to make well-judged, timely decisions. The Actor Closeness Centrality (ACC) is measured between zero, as would be indicated by a circular network visualisation, and one, which would be indicated by a star network.

Once the networks are measured, the results are represented using visual analytic tools and techniques. Visual analytics has become important as a field in applied computing and is defined as the science of analytical reasoning facilitated by interactive visual interfaces (Chabot, 2009). Visual analytics facilitates the representation and understanding of large and complex data sets. Visual analytic applications incorporate both visualisation and statistical elements that, combined, provide the researcher with a more nuanced and sophisticated view of their data. Moreover, this area of human–computer interaction is designed to be exploratory (Perer and Schneiderman, 2009). This is particularly important in the study of the networks identified within this study which analyses and measures approximately 210,000 relationships involving 1700 actors. This large and complex information cannot be represented meaningfully in more traditional ways, such as tables. Indeed, specific computer programs have been developed by the authors to analyse and prepare the relationship data used in this study for input into SocNetV as this could not be achieved manually. Thus, visual analytics is ideally suited to this study as it not only enables the representation of a network (and its subnets), but also measures their complicated and interrelated nature over time.

In order to demonstrate, measure and quantify relational changes in the networks over the period under study, the network analysis of institutional membership is broken down into three groups of two decades: 1750–1769, 1770–1789 and 1790–1809. We then use four methods to visualise and measure the relationships between these actors within these two-decade periods using the SocNetV social network application. First, a cluster analysis provides a high-level view of membership by and across institutions and visualises the potential communications flow available to members through relational connections (represented by lines) between actors. In addition, it highlights key bridges between the institutions through identification of subnets, i.e. those actors that belong to more than one institution. Second, this paper uses the three centrality measurements outlined above to assess relationships afforded by institutional membership. In the resulting visualisations of the networks, the closer an actor is to the centre of the diagram, the more central they are in the network by that measure, i.e. the nearer they are to a measure of 1.

To further demonstrate how the networks facilitated business during this period, we have used the Trans-Atlantic Slave Trade Database (Eltis et al., 2010) to follow investment trends in this trade by actors in our networks. Whilst we recognise that this trade was only one of many interests that these actors will have had during this period, this source provides evidence as to how Liverpool's merchants used their institutional membership networks to engage in economic activity. In addition, it supports our argument that, far from being static, networks and the way that actors used them was dynamic.

Fig. 4 illustrates the number of investors in the slave trade versus the number of voyages embarking from Liverpool 1750–1810. Duplicated names have been removed from the data downloaded from the Trans-Atlantic Slave Trade Database (Eltis et al., 2010) to provide the total individual investors by decade. As can be seen, the number of voyages embarking from Liverpool increases throughout the period, apart from 1780 to 1789. It should be noted that the slave trade was abolished in 1807 yet there were only thirteen fewer voyages between 1800 and 1807 than the whole of the previous decade. Interestingly, the pattern of the number of investors follows our temporal two-decade groupings. During 1750–69 there is a rise in investors, followed by decline in the next two decades. The period 1790–1810 again sees a substantial rise in investors, no doubt trying to take advantage of the last rush to purchase slaves. In addition, in contrast to the 1750s and 1760s, many of the investors came from outside the institutions in this study, suggesting a shift in capital requirements and distribution of wealth to small-scale merchants and investors. Moreover, these investment groups were smaller than at the start of our period (McDade), suggesting that fewer merchants were responsible for this increase in voyages. As will be demonstrated in this study, network analysis elucidates the role that institutional membership played in the changing investment patterns and potentially the wider economic activity of the town.

During the 1750s and 1760s as illustrated in Figs. 5(a) and (b) many actors were involved in both formal and informal institutional membership. Three institutions are identified in the visualisations: the African Committee (AC) in the top left, the Town Council (TC) at the foot and the drinking clubs (Ugly Face Club-Mock Corporation) (MC) in the top right. Cross-institutional membership forming subnets are identified by being placed between the institutions to which they belong. The cluster analysis demonstrates that these two decades were characterised by much cross-institutional subnet membership. The largest of these subnets is the African Committee-Town Council cluster with eighteen actors in the 1750s and nineteen in the 1760s, and is a trend that continued until 1810. However, the importance of the drinking clubs is also apparent with many actors in either the African Committee-Mock Corporation or the Town Council-Mock Corporation subnets.

During these first two decades, some actors were members of all three institutions. These actors had a strong position in the network's communications flow in that they were related to all the other actors. Four of these actors, Robert Armitage, Matthew Stronge, Richard Hughes and John Parr, were members of all three institutions during the 1750s. However, by the 1760s these four actors had left the Mock Corporation to belong to the African Committee-Town Council subnet. Three of these four actors during the 1750s appeared to use networks to gain a place on the Town Council. For example, merchant Matthew Stronge joined the Ugly Face Club in 1743, where he was described as having "a Tawny Complexion A Sharp Nose A Hook Mouth Irregular Bad Set of Teeth Like Those of an old worn out Comb, Thoroughly begrim'd A Ghastly queer Grin & Countenance greatly Set off by[?] a long Carrotty Beard" (Ugly Face Club). Despite these rather unattractive attributes, he later joined the African Committee in 1750 and the Town Council in 1751. He served as the corporation treasurer from 1764 to 1767 until he became Mayor in 1768. Richard Hughes joined the Ugly Face Club in 1749 and the African Committee in 1750, and eventually the Town Council in 1752. He became Mayor in 1756. Finally, John Parr joined the Ugly Face Club in 1743 and the African Committee in 1750 before being elected to the Town Council in 1757. During the 1760s three actors were members of all three institutions; William Boates, William Gregson and Ralph Earle. Although we do not see there the same pattern of incremental club membership to gain political office that was obvious in the previous decade, there is still an element of using these networks for career progression. William Boates was a member of the Ugly Face Club as a ship's captain in 1756 and on 28th April of that year was described thus: "a Rugged [sic] face: a turtle nose Pig Eyes sallow Complexion [sic] much Improvd [sic] by going to Africa; Gastly [sic] grin: freckled face & upon the whole well Qualified" (Ugly Face Club), a description obviously reflecting his involvement in the slave trade. By the 1760s he had risen to merchant (Gore, 1766). He joined both the African Committee and Town Council in 1763. William Gregson, another merchant (Gore, 1766), first joined the Town Council in 1760, then the Mock Corporation and African Committee in 1763. Ralph Earle, a timber merchant (Gore, 1766), joined the African Committee in 1750, the Town Council in 1760 and the Mock Corporation in 1764. He became Mayor in 1769.

Figs. 6(a) and (b) illustrate out centrality, or the number of contacts an actor possessed, in the 1750s and 1760s. As discussed above, these early decades were characterised by cross-institutional membership and the influence of this on out centrality is evident. The actors with the greatest expanse of ties were those seven who belonged to all three institutions discussed above. Also of interest is the ranking of institutions and subnets. Actors who were members of the African Committee-Mock Corporation and the Town Council-Mock Corporation subnets are also ranked highly using this measure. The highest ranked single institution is the Mock Corporation, followed by the African Committee-Town Council subnet. The institutions that gave

access to the fewest connections in the wider network were the African Committee and the Town Council.

The importance of the drinking club, and in particular joint membership of it and another institution, for access to a widening of ties reinforces the arguments that these were used for career progression by many in the Liverpool mercantile community. Key actors within the network at this time, such as merchants Thomas Rumbold, Robert Clay and George Clows (Gore, 1766), continued their membership of the African Committee-Mock Corporation subnet during the 1750s and 1760s. Thomas Johnson also continued his dual membership of the Town Council-Mock Corporation during the two decades. Actors such as Unitarian Benjamin Heywood in the 1750s and Robert Seel, John Backhouse and Thomas Hodgson junior in the 1760s, may have used their dual membership of the African Committee-Mock Corporation to gain contacts during the early part of their careers. This may have also been the case with the Town Council-Mock Corporation subnet as three actors, Francis Gildart, Spencer Lawrence and Joseph Clegg, belonged to both during the 1750s but left the Mock Corporation in the 1760s once they had been voted onto the Town Council. Only John Hughes, a gentleman (Gore, 1766), was solely a member of the Town Council in the 1750s but in the next decade was part of the Town Council-Mock Corporation subnet. This may have been due to the fact that membership of the drinking club was no longer generally acceptable, or indeed, necessary, once actors had moved up the socio-economic ladder.

Figs. 7(a) and (b) illustrate the closeness centrality, or relative access to others within the network, during the 1750s and 1760s. This measure again illustrates the importance of cross-institutional membership and largely concurs with the out-centrality analysis in that the single institution with the greatest centrality is the Mock Corporation during these two decades. It suggests that actors who belonged to this institution during this period were able to use their networks to acquire informal information or 'gossip' which may have provided them with a commercial advantage. As noted above, the Ugly Face Club comprised members from a variety of professional occupations, from ships' captains to doctors to merchants. This club, and the Mock Corporation, also contained many prominent members of the Liverpool trading community at varying stages of their career. William Boates, noted above for his rugged face, was a ships' captain when he joined the Ugly Face Club, but had access to key actors in the Liverpool mercantile community, such as William Armitage, Francis Gildart, Benjamin Heywood, Thomas Hodgson and Matthew Stronge. Reputedly an orphan (Wilson, 2008), he would not have had good family networks and such social networking may have been especially important in developing his career from mariner to merchant. Of course, his experience as a slave trade captain was his own human capital to exchange with others (Behrendt, 2007), and he later invested with many others through these associations. It is interesting to note that the least independent

institution using this measure is the Town Council. Despite being the most important institution politically, it did not provide the best access to other people. As will be demonstrated, this became increasingly significant over the whole period of this study.

Figs. 8(a) and (b) illustrate the betweenness centrality, or those who acted as chokepoints of information during the 1750s and 1760s. Any cross-institutional membership during these two decades gave actors who belonged to these subnets an advantage in the control of information over those in single institutions. As might be expected, those actors who belonged to all three institutions had the greatest control over information flow, and are at the centre of the figure. Of note is the ordering of subnet membership using this measure. During the 1750s, the most important subnet is the African Committee-Mock Corporation membership followed by the Town Council-Mock Corporation. However, by the 1760s, the positions of these two subnets are reversed. During these two decades, the cross-institutional subnet that was least likely to act as a chokepoint of information flow is that of the African Committee-Town Council. Despite, or perhaps due to, the fact that it was a durable subnet, it relied on information passed by others to a greater extent than the other cross-institutional subnets.

During the 1750s and 1760s, network density was high, and these measures show that there were many opportunities for information flow via cross-institutional membership including between formal and informal institutions. In the 1750s, four actors belonged to all three institutions and these actors used their advantageous position for career progression or investment in the slave trade. For example, the one slaving voyage that Robert Armitage organised in 1754 was financed by four investors, all of whom were members of the African Committee, Town Council or both. Matthew Stronge organised two voyages, and of the eight investors, four belonged to the African Committee, Ugly Face Club/Mock Corporation and Town Council. Thomas Rumbold belonged to the African Committee-Mock Corporation subnet and organised three voyages involving seventeen investors, fourteen of whom belonged to the same institutions as Rumbold. Robert Clay organised seven voyages involving eight investors, five of whom belonged the same two institutions as he did. This trend continued in the 1760s, if to a slightly lesser extent. William Boates organised twenty-seven voyages involving sixteen investors, six of whom belonged to the same institutions. William Gregson organised nineteen voyages involving sixteen actors, eight of whom were involved in the institutions. Thomas Hodgson organised eight voyages with eleven investors, five of whom belonged to the same institutions. The analysis suggests that Liverpool's mercantile community was using its strong and weak ties gained through the networks under study for investment, in order to take advantage of trading opportunities during this period. The network measures highlight that one of the most advantageous dual-membership subnets to belong to was that involving the drinking clubs.

From 1770 to 1789, as illustrated in Figs. 9(a) and (b), the network dynamics change with the addition of the Lyceum (L) as a cultural institution. Thus, four institution memberships are visualised: the African Committee in the top left, the Town Council in the bottom left, the Mock Corporation in the top right and the Lyceum in the bottom right. It is clear that cross-institutional membership still occurred, but to a lesser extent than in the previous two decades, and no actors belonged to all four institutions. In the 1770s, three actors belonged to three of the four institutions: John Blackburne, a merchant (Gore, 1774), and Thomas Rumbold were members of the African Committee, Lyceum and Town Council, while Richard Hughes belonged to the Town Council, African Committee and Mock Corporation. It should be noted that only one actor, Richard Hughes, belonged to both the African Committee and Mock Corporation, a subnet that had been favoured in the previous two decades. In the 1780s, three actors belonged to three of the four institutions; Thomas Earle (Mayor in 1787), a merchant in partnership with family member William (Bailey, 1787), and Thomas Staniforth, a merchant who had previously been in partnership with Benjamin Heywood in the 1770s (Gore, 1774 and Bailey, 1787), belonged to the Town Council, African Committee and Lyceum, while George Warren Watts, listed as a merchant and insurance office owner in the 1770s and possibly in partnership as merchants Watts and Walker during the 1780s (Gore, 1774 and Bailey, 1787), belonged to the African Committee, Mock Corporation and Lyceum. Indeed, George Warren Watts enjoyed socialising with his peers as he also attended the Unanimous Club (not included in this study due to limited data), in 1775 and 1776 along with other leading Liverpool merchants (Bailey, 1787). Again the fact that membership of these institutions was highly instrumental is highlighted by the fact that between 1770 and 1777, Watts invested in fifteen out of thirty-three slave trade voyages with the following members of the Unanimous Club: Joseph Brooks junior, William Crosbie, Alexander Nottingham, William Pole, Charles Pole and Clayton Case (Eltis et al., 2010). Excluding the Poles, all of these were members of the African Committee during this decade alongside Watts.

The 1770s marked a period of decline in use of the Mock Corporation by members of the African Committee. This suggests that merchants were no longer using drinking clubs for access to other actors, and this trend may have been encouraged by the fact that the conversation at the Mock Corporation tended to be social rather than commercial, as well as by a declining need for expanding networks amongst the more established merchants (Wilson, 2008). However, six actors joined the African Committee-Lyceum subnet, including Benjamin Heywood who had previously used the informal drinking club and was a member of the Lyceum during both decades. This may have as much to say about cultural aspirations as about mercantile ambitions. During 1780–89, newer members of the African Committee again straddle these two institutions. Of these five actors who described themselves as merchants, three, Andrew Black, Cobham Richardson and William Wilson, were first active in the African Committee and then the Mock Corporation. John Hodgson, a merchant in partnership

with his brother Thomas (Bailey, 1787), and Ralph Fisher first became active in the Mock Corporation and then in the African Committee. In the previous decade, John Hodgson had belonged to the Lyceum-Mock Corporation subnet. In the 1770s, eight actors were in the Mock Corporation-Lyceum subnet, suggesting that this subnet still provided networking opportunities, but perhaps not trade-related ones. Indeed, membership of these institutions may have been associated with reputation and civic status rather than commerce. Once these actors are economically established within the wider community, they may have used their financial status for cultural purposes. This is particularly noted in the African Committee-Lyceum subnet which included leading Liverpool merchants, such as Benjamin Heywood, Thomas Staniforth, John Stronge and Edgar Corrie (Gore, 1774 and Bailey, 1787), which illustrates that slave traders and non-conformists were still socialising together. However, these subnets decline slightly in the 1780s despite still including leading merchants such as Benjamin Arthur Heywood and Thomas Hodgson. The strong relationship between the African Committee and Town Council continued with a large number of actors in this subnet.

These second two decades were therefore characterised by a complexity of networks in flux during a transitional period. Cross-institutional membership continued to provide some opportunities for potential contacts within the network during the 1770s and 1780s as further illustrated in Figs. 10(a) and (b). During the 1770s, cross-institutional membership provided greater opportunities for contacts than belonging to a single institution. The exception was the Mock Corporation which was more advantageous for contacts than belonging to the African Committee-Town Council subnet. This position was reversed in the 1780s with actors in the African Committee again using the drinking club so that the African Committee-Mock Corporation subnet ranks second in the network analysis. This may be due to a new generation of merchants entering the community. During both decades, one of the weaker-ranked subnets is that of the Lyceum-African Committee, although it still provided more opportunities for contacts than the African Committee-Town Council subnet.

There is also a shift in single institution ranking whereby the African Committee fell below the Lyceum in the 1770s, but the position was reversed in the 1780s. This is significant due to the rise of the abolitionist movement at this time. For example, some members of the Lyceum, including William Rathbone, William Roscoe, Daniel Daulby and James Currie, were members of the London based Society for the Abolition of the African Slave Trade in 1788 (Trepp, 1928). These actors may have had influence within the Lyceum itself, but the African Committee had greater potential influence in the wider network and therefore in the Liverpool mercantile community.

Closeness centrality (independence and access to timely information) as illustrated in Figs. 11(a) and (b) follows a similar pattern to out centrality in the 1770s. During this decade, the Mock Corporation provided the fastest opportunity for information flow as a single institution, and in comparison to some cross-institutional groupings. This

position changed during the 1780s when it was more advantageous to belong to any cross-institutional subnet rather than to a single institution. It was also when, as noted above, the African Committee provided the best access to information as a single institution. This may mean that it was becoming more efficient and that its members were well placed to mobilise support within the network; or that it was becoming isolated and inward looking. This decade also witnessed a decline in the importance of the Mock Corporation within the wider network, being ranked behind the Lyceum. This supports the assertion that the use of drinking clubs for trade-related networking declined, but that the Lyceum was important culturally.

Figs. 12(a) and (b) highlight the key actors who were chokepoints of information. In the 1770s, Richard Hughes and Thomas Rumbold as members of the African Committee-Lyceum-Town Council subnet and John Blackburne as a member of the African Committee-Mock Corporation-Town Council subnet enjoyed a strong position in the control of information flow. In the 1780s, this position was taken over by George Warren Watts who was a member of the African Committee-Mock Corporation-Lyceum subnet. In contrast to the out centrality measure, betweenness highlights the continuation of the importance of the drinking club in the control of information. For example, in the 1770s, Johnson Gildart and James Gildart enjoyed a strong position over others in the network by belonging to the Town Council-Mock Corporation subnet. The next highest ranked subnet was that of the Lyceum-Mock Corporation, with members such as merchants John Hodgson and James Hatton (Gore, 1774). In previous decades, the more institutions to which an actor belonged, the greater the advantage they gained in the control of information. However, during the 1780s, after George Warren Watts, members of the Lyceum-Mock Corporation subnet were ranked more influential than merchants Thomas Earle and Thomas Staniforth who belong to the Town Council-African Committee-Lyceum subnet, i.e. three institutions. Following these two actors, those involved in the African Committee-Mock Corporation subnet are the next highest ranked. This suggests that whilst the Mock Corporation lost its importance as a key location for accessing information, it remained important for the flow of information.

During the 1770s and 1780s there was still cross-institutional membership but not to the same degree as in the earlier decades. As a result, network density, and therefore group cohesion, started to decline. Importantly, the drinking clubs were not used to the same extent for trade-related networking or career progression. Informal institutional networks were still important for co-investment opportunities during the 1770s, but this trend declined in the 1780s, reflecting the decline in total investors seen in Fig. 4. This may have been due to a number of factors, such as the influx of 'new blood', the cultural aspirations of the mercantile elite, dislocations caused by the American War of Independence and increasing tension between the West India and American merchants (Checkland, 1958).

There is, however, a clear sense of a move away from using institutional networks as a potential for investment, at least in the slave trade. For example, as noted earlier, Thomas Rumbold was well placed within the network as he belonged to three institutions; the African Committee, Lyceum and Town Council. During the 1770s, he organised fifteen voyages involving eight investors. Possibly only one of these investors belonged to any of the institutions, the African Committee.² This is in contrast to the voyages he organised in the 1750s where most of the investors were related through Ugly Face Club or Mock Corporation memberships. In the 1780s John Hodgson belonged to the African Committee-Mock Corporation subnet and organised twelve voyages involving five investors, only one of whom belonged to the African Committee. Ralph Fisher organised twenty-one voyages with twelve investors, only two of whom were members of the African Committee and only one of whom belonged to the Mock Corporation. William Boates had previously used his institutional network connections for investment in the slave trade, but by the 1780s organised voyages either independently or in partnership with Thomas Seaman (a former ship's captain who had previously worked for Boates) and James Percival, neither of whom belonged to the institutions in this study (or any others as far as we can tell). Other actors in this decade did use their institutional networks for investments, but it was the formal rather than informal clubs that provided actors with economic capital. Thomas Earle organised nineteen voyages involving fifteen investors, nine of whom are identified as belonging to the institutions included here. However, eight of these belonged to the African Committee and only one belonged to the Mock Corporation. Thomas Staniforth organised seventeen voyages involving twenty-one investors, thirteen of whom belonged to the institutions in this study. However, only three of his co-investors belonged to the Mock Corporation and they were not members of the African Committee.

Figs. 13(a) and (b) illustrate the cluster analysis for 1790 to 1809. Unfortunately for this analysis, the Mock Corporation data ends in 1792 and therefore the influence of the drinking club in these two decades cannot be fully assessed. However, it is clear that this period was characterised by very little cross-institutional membership apart from the African Committee-Town Council subnet, which was as strong as it had been in 1750. During the 1790s, there was little cross-institutional membership: only two actors, Thomas Midgely, a gentleman, and Thomas Moss Tate, a merchant and tobacco and snuff manufacturer (Gore, 1796), belonged to the African Committee-Lyceum subnet; one actor, William Neilson, a merchant in his own right and possibly in partnership as Neilson and Heathcote (Gore, 1796), remained in the African Committee-Mock Corporation subnet; and one actor, James Williamson, belonged to the Lyceum-Mock Corporation subnet. By the 1800s, the institutions had become further isolated. The Town Council still had strong links with the African Committee but only one actor, James Gildart, listed in 1805 as Receiver of the Dock Duties (Gore, 1805), was involved in cross-institutional activities by belonging to the African

Committee-Lyceum subnet. Indeed, this role would have given him substantial connections with actors and networks involved in activities outside those of the institutions in this study, such as those trading to America and the Baltic. These two decades were therefore characterised by a concentration, or 'entrenchment', of institutions utilising strong ties rather than the weak ties offered by inter-institutional membership. A strong communications link remained between the African Committee and the Town Council to the exclusion of other institutions. However, the strength of this link may also have been its weakness in that it became self-affirming (Podolny and Page, 1998).

The out-centrality analysis in Figs. 14(a) and (b) highlights further the decline in cross-institutional membership. The 1790s followed the trend of the 1780s with the decline of the Mock Corporation although this may be exaggerated by the unavailability of some data in the 1790s and obviously the complete absence of data in the 1800s. Fig. 14(a) also highlights the trend seen in the previous two decades reaffirming the rise in the relative importance of the African Committee-Town Council subnet. However, Thomas Midgely and Thomas Moss Tate had the greatest number of contacts by belonging to the African Committee-Lyceum subnet. In Fig. 14(b), the contracting networks are further demonstrated, with only James Gildart in the advantageous position for potential contacts as he belonged to the African Committee-Lyceum subnet. In both decades, the Lyceum is the highest ranked single institution followed by the African Committee, but, again, the lack of cross-institutional membership suggests that these institutions were being accessed for different purposes. The members of the Town Council remain the furthest 'out of touch' with the wider network despite it being the seat of political power. This analysis suggests that during a period when the trade of many of its members was under political and economic threat, they had less access to actors within the wider network and were therefore less able to mobilise support or gain information. The Lyceum, which as discussed above had significant abolitionists in its membership, had potential to access a greater number of actors particularly in a period when dissent against a prominent trade of the town is being expressed nationally.

As illustrated in Figs. 15(a) and (b), those actors who were best able to access information in the network quickly (were the closest to others) were Thomas Midgely and Thomas Moss Tate in the 1790s and James Gildart in the 1800s due to their cross-institutional membership. However, the African Committee-Town Council subnet has a higher ranking than in previous decades. In the 1790s, it took up an advantageous position above the two actors (James Williamson and William Neilson) belonging to the other two cross-institutional subnets. This subnet was able to pass information around faster than the two other subnets or sole institutions. However, whilst this subnet had closer ties within the wider network, a comparison with the cluster analysis suggests that although it was able to pass information around faster than other subnets, it was

unable to access new information through cross-institutional membership. The African Committee also had an advantageous position in the 1790s in that it had close ties. However, this position is lost to the Lyceum during the 1800s. It should also be noted that the poorest institution for closeness was the Town Council. Even in the 1790s, it was ranked lower than the Mock Corporation despite the fact that the drinking club had fewer actors. The Town Council therefore was the least independent institution in terms of information and had less timely access to that information compared to those actors involved in other institutions.

As seen in previous decades, cross-institutional membership provided actors with the opportunities to act as chokepoints of information flow. However, as illustrated in Figs. 16(a) and (b), with little cross-institutional membership very few actors had an advantageous betweenness position in the network. In the 1790s, the most influential chokepoints were those actors who belonged to the African Committee-Lyceum subnet, the African Committee-Mock Corporation subnet and the Lyceum-Mock Corporation subnet. During the 1790s and 1800s, the importance of the African Committee-Town Council subnet as a chokepoint of information flow is evident. However, this was not because the subnet increased in importance as a chokepoint of information compared to other decades, but due to fewer actors engaging in cross-institutional activities as a whole in this period. This provides more evidence that the African Committee-Town Council subnet failed to use its potentially advantageous network position.

The networks between 1790 and 1810 were therefore characterised by very little cross-institutional membership with actors further coalescing into interest groups. This could be due to actors leaving the networks in this study to join others that reflected new economic trends within Liverpool. Despite the rise in investors in the slave trade during these two decades, seen in Fig. 4, there was a decline in the use of other members of the institutional networks for potential investment. For example, William Nielson, the only actor belonging to the African Committee-Mock Corporation subnet in the 1790s organised twenty-eight voyages with only four investors, only one of whom was also in the African Committee. Thomas Staniforth, who belonged to the African Committee-Town Council subnet, organised seven voyages but used his relationships in the African Committee more. This may be due to his employing business practices that he had used successfully since the 1770s. Of his thirteen investors, six also belonged to the African Committee. Thomas Earle used his relationships within the institutions far less. He organised twenty-four voyages involving seventeen investors. Yet only two also belonged to the African Committee with the others belonging to none of the institutions in this study. This reliance on investors outside the institutional networks continued into the 1800s. Thomas Earle was in the African Committee-Town Council subnet in this next decade and organised five voyages with four investors, but none of them belonged to any of the institutions.

George Case, also in the African Committee-Town Council subnet, organised thirty-five voyages with only seven investors, only two of whom also belonged to the African Committee.

Conclusion

Using visual analytics has facilitated the analysis of a large data set with over 210,000 relationships involving 1700 actors. Moreover, it has provided a nuanced and sophisticated view of the way in which institutional membership networks changed over time, not only in the long term, but dynamically within the short term as well. This study confirms that using tools such as visual analytics can be useful in raising both specific and wider issues from the quantitative data, especially where there is a lack of extant qualitative data. In addition, visual analytics is exploratory in nature, and therefore may raise questions, rather than answer them per se. The results of this research suggest some conclusions about Liverpool's metropolitan business networks, and also about networks more generally.

Despite being the leading British out port during the period 1750–1810, Liverpool's business networks, as represented by the data here, coalesced towards bonding rather than bridging networks. This occurred as Liverpool's economy matured, especially with regard to the Atlantic trade, and the slave trade in particular. The fact that there was a decline in cross-institutional membership may also be a reflection of the increasing tension between the many non-conformists who traded with the thirteen continental colonies and the Anglicans involved in the West India and slave trade. Even banking was split along these lines with the West India men banking with Gregsons & Co. and then later Leyland & Bullins, and the American merchants with Clarke & Roscoe (Checkland, 1958). Indeed, it has been suggested that there were two identities in Liverpool by the late eighteenth century (Stobart, 2000). This tension is reflected in our network analysis.

Furthermore, as these institutional networks coalesced, so did investment groups – at least in the slave trade. Over time investment groups in the slave trade got smaller and smaller (McDade) and relied less and less on institutional networks. However, the number of individuals investing in the slave trade during the 1790s and 1800s grew dramatically following a decline. Moreover, by the end of the period, a few men such as Thomas Case dominated the Council (Ascott et al., 2006). They had political power and were rich men from well connected families. For example, Thomas Leyland may have been exceptional, having died with an estate worth somewhere under £600,000, but William Pole, mayor in 1778, died leaving around £80,000 and Thomas Molyneux, mayor in 1806, left just under £35,000 (Pope, 2007). Therefore, not only did these men not need the drinking clubs for career progression, but they no longer required the more formal institutions for co-investment either. However, their institutional memberships, and the networks and social capital the memberships afforded, meant that they were well placed to provide investment opportunities to the wider Liverpool

business community. Indeed, the rise in investors and voyages during the 1790s and 1800s following the decline in the previous two decades suggests a change in capital requirements and distribution of wealth in Liverpool reflected in our network analysis with small-scale merchants and investors moving into this trade.

The increasing coalescence and dysfunctionality of these metropolitan networks may help to explain why Liverpool merchants were not better able to defend against the movement to abolish the slave trade. They were slow to react to Dolben's Act in 1788 (Sanderson, 1972) and did not fully participate in the London Society for Abolition (Ryden, 2009). It is certainly difficult to explain William Roscoe's election as MP in 1806. True, Roscoe campaigned on an anti-East India Company ticket (Wilson, 2008), but it must have been well known that he supported abolition. He was also nominated by his business partner in his bank, Thomas Leyland, and seconded by Thomas Earle, both prominent slave traders. This seemingly strange occurrence is, however, a reflection of the fact that most Liverpool slave traders engaged in the slave trade as only part of a wider business portfolio (Haggerty, 2008), William Davenport being a lonely exception (Richardson, 1976). Furthermore, most merchants were involved in a variety of partnerships in order to bring together skills, capital and contacts. For example, Thomas Case was in partnership with Nicholas Southworth in the slave trade and in importing sugar and rum from Jamaica, ran a Manchester warehouse, exported dry goods to Philadelphia, acted as an insurance broker in partnership with William Gregson, ran a colliery, and was a proprietor of a fire office (Haggerty, 2008). Therefore, whilst the Town Council was dominated by slave traders who did not engage in cross-institutional membership, its members were able to assess and cater for their needs through their wider trading interests. This is demonstrated by the impressive dock programme in the port throughout the eighteenth century, planned and supported by the town council (Power, 1997), which catered well for all the trades its wider community was engaged in.

Despite abolition of the slave trade, Liverpool's success continued into the nineteenth century. This is no doubt a reflection of the increasing importance of other sectors of the trading community, not included in the institutional membership analysed here. Networking was clearly occurring in other spaces amongst the other groups. For example, there was a Chamber of Commerce established in 1774 and an American Chamber of Commerce in 1801. The members of the American Chamber of Commerce, despite including a few slave traders such as Thomas and William Earle, were clearly another group from those on the Council and African Committee, including the Quaker Rathbones. Indeed, as this group gained in importance, they gained in influence too. For example, American merchants such as William Rathbone, Thomas Martin, and James Cropper were called to give evidence to the House of Commons in 1808 regarding the increasing tensions with the United States (House of Commons Parliamentary, 1808). The first president of the Shipowners' Association set up in 1810

was also a leading American trader, Samuel Holland (Tolley, 1812). It is clear therefore that whilst the networks of the slave trade-dominated Council and African Committee were coalescing, other groups of merchants were on the rise.

Liverpool's experience in 1750–1810 has much to say about networks more generally. Increasingly bonding networks mean that some metropolitan business networks are less able to respond to wider changes in national and international trade. As Granovetter notes, the shape of the network environment determines a group's ability to mobilise in the face of external pressure (Granovetter, 1973). Moreover, it is clear that as a network matures, its members need institutions less and less for access to information and credit as well. Lastly, this case study of Liverpool suggests that, as is the case with clusters (Swann, 1998), there is a life cycle of networks. This means that some will eventually wither, but that this facilitates the emergence of new and more dynamic networks which eventually take over. Importantly, by looking at networks within institutions we can see that 'far from being divorced from the study of political institutions and culture, networks are the constitutive elements that sustain, rupture, and transform social and economic institutions' in turn (Smith-Doerr and Powell, 2005).

B.3 P_Econ

The Probabilistic Counter-Revolution, or How Stochastic Concepts came to Neoclassical Economic Theory

Philip Mirowski

The earliest historiography of the rise of econometrics, such as that found in Schumpeter (1954, p. 962) served to conflate the rise of econometrics with the development of empirical measurement in economics; and it is true many contemporaries did tend to regard the elaboration of mathematical economic theory and that of statistical estimation as aspects of the same research program, at least until circa 1950. The timing of the introduction of procedures such as least squares estimation was regarded by many as a simple process of diffusion of technique across disciplinary boundaries, with historians such as Schumpeter asserting that, "The majority of theorists, including some of the greatest, were completely unaware of the possibility of a theory that might eventually achieve numerical results." Some familiarity with recent discussions in the history and philosophy of science would suggest that an understanding of the "Econometrics Revolution" demands a broader scope than that which presumes a simple model of diffusion of techniques, or one which tries to explain the rise of econometrics solely from such "internalist" considerations as the "demands of the data" or the "logic of the economic problem". Briefly, in our alternative reading the timing and structure of what subsequently became known as "econometrics" was due in large part of a crisis within neoclassical economic theory, ultimately deriving from an advocacy of a strong determinist model of explanation copied directly from physics, just as physics seemed to be repudiating such a model. The genesis of a probabilistic quantum mechanics in 1925-6 exacerbated the crisis, but also provided neoclassical economists with certain resources to rebuff the criticisms and to forge a rapprochement between stochastic principles and their theory.

I. Determinism vs. indeterminism

To understand the history of econometrics, one must first consider the history of neoclassical economics; and, to understand neoclassical economics, one must first have some understanding of the history of physics. Most economists understand instinctively that the neoclassical research program has striven to attain the status of a science, and not just any science, but that most respected of the modern sciences, physics. Yet few realize the extent to which the progenitors of neoclassicism acted to secure that status. Indeed, I have argued elsewhere that the so-called Marginalist Revolution in the 1870s consisted largely of engineers directly appropriating the newly developed formalisms of nineteenth century energy physics, changing the names of the variables, and renaming the result "mathematical economics" (Mirowski, 1984a; forthcoming a). Precisely because this claim may appear controversial, I shall not

attempt a full elaboration here, but merely take this thesis as a point of departure. When economists such as Jevons, Walras, Edgeworth, Pareto and Fisher transmuted the physics of energy into the social mechanics of utility, it was with some ambivalence that they enlisted in the determinist program. On the one hand, it seems fairly clear that they conflated the notion of classical determinism with scientific explanation as a whole. However, when it came to the actual mathematics, the bulk of the work of these protagonists displayed much more fascination with the variational principles than with the conservation principles, to the detriment of the development of an analogous dynamics (Mirowski, forthcoming, chap. 5). Whatever the motivation, this tergiversation over conservation principles had severe consequences for the goal of a neoclassical dynamics which could claim to parallel the classical determinist doctrine (Northrop, 1941). Recourse to the formalisms of Hamiltonian dynamics would have explicitly required that something be conserved through time. Since neoclassical economists were unwilling or unable to specify what that was, their prognostications were effectively confined to static theory. This conflict between the physics ideal of classical determinism and the retreat to a purely static theory was a blight on the neoclassical research program, a canker that could not be ignored by the first decades of the twentieth century.¹

If this deficiency with respect to a legitimate dynamics were not sufficiently distressing, the neoclassical research program was further buffeted by some bad news from the physicists' camp. Just as the economists had come to pay homage to the physicists' belief in determinism, the physicists themselves were contriving to distance themselves from it, and in some cases, even to express doubts concerning its intellectual validity (Hacking, 1983b; Brush, 1983). These doubts began with the observation that it was humanly impossible to know all of the initial conditions for any moderately complicated Hamiltonian, and therefore it was simpler to treat an aggregate of mass points such as, say, an ideal gas, by statistical averages (Porter, 1985). Statistical gas theory then led to an explanation of the second law of thermodynamics using stochastic constructs (Brush, 1983). By 1919, some physicists were suggesting that, "statistical law must be regarded as the comprehensive genus, a concept of higher order than that of dynamic laws and including them as a special case" (Cassirer, 1956, p. 81). But the most devastating jolt came in 1926 with Max Born's interpretation of the formalisms of the new quantum mechanics as explicitly and irreducibly stochastic. It prompted physicists in the 1920s to openly worry about a "crisis in the foundations of physics" (Brush, 1983, p. 128) and to speculate that classical determinism had become untenable. Whatever one's opinions on the wayward course of classical physics, most agreed that the future of physics lay in statistical concepts (Brush, 1983, p. 102; Brush, 1980).

It is difficult now to evoke the extent to which this was regarded as an epoch-making departure in Western physics and philosophy in the 1930s (similar to the reactions to

"chaos theory" in the 1980s). Not only was it the topic of arcane disquisitions amongst physicists, but was also widely bandied about in more popular venues. Heisenberg's "Uncertainty Principle" was thought to support all sorts of outlandish philosophical positions, from solipsism to the necessity of the existence of free will (Born, 1949). It became briefly fashionable in certain quarters to maintain that the very concept of causality had been undermined (Forman, 1971). By the 1930s, any culturally literate layperson could not turn around without bumping into some denunciation of determinism and praise of stochastic concepts:

"[classical determinism] was the gold standard in the vaults; [statistical laws were] the paper currency actually used. But everyone still adhered to the traditional view that paper currency needs to be backed by gold. As physics progressed the occasions when the gold was actually produced became rarer until they ceased altogether. Then it occurred to some of us to question whether there still was a hoard of gold in the vaults or whether its existence was a mythical tradition. The dramatic ending of the story would be that the vaults were opened and found to be empty. The actual ending is not quite so simple. It turns out that the key has been lost, and no one can say for certain whether there is any gold in the vaults or not. But I think it is clear that, with either termination, present-day physics is off the gold standard." (Eddington, 1935, p. 81).

Thus the neoclassical research program confronted a serious dilemma in the first few decades of the twentieth century. The adoption of the mathematical metaphor of utility as potential energy was intertwined with the classical determinist posture, one which equated scientific causal explanation with mechanical prediction. Yet the allegiance to classical determinism was not paying off, as a plausible dynamics still appeared beyond reach, no closer than it had been in the 1870s (Mirowski, forthcoming, chaps 5-6). Moreover, a quantitative empirical program was stymied by the rigors of classical determinism, since anything less than exact conformity of data to a mechanistic model meant the rejection of that model. Time series of economic data were beyond the pale of neoclassical explanation, since the static model could not legitimately be extended through time, due to the perennial complaint that the very structure of the fundamental determinants of the economy had changed in the interim.

Hence, the neoclassical research program at the turn of the century faced an unpleasant choice: either persist in a determinism bereft of the kind of results one associated with physics, or else throw over the entire program of copying physics. In retrospect, it seems that neoclassicals might have remained satisfied with their flawed science, perhaps papered over by various Marshallian devices like the "long run" and the "short run", had it not been for the further bad tidings from the physicists' camp. It slowly dawned upon some neoclassicals that the physicists were changing the rules of the game when it came to "natural law", and it was this cohort which were identified with the rise of "econometrics": Ragnar Frisch, Harold T. Davis, Tjalling Koopmans,

Henry Schultz, Trygve Haavelmo, Gerhard Tintner, Harold Hotelling, Charles Roos, Jacob Marschak and others loosely affiliated with the Cowles Commission.

The first neoclassicals such as Jevons, Walras, Fisher and Edgeworth had appropriated the earlier vintage of physics metaphor in order to arrogate the legitimacy of a "science"; and now their legacy risked being down-graded to a "non-scientific" status as their deterministic conception of explanation was being rendered obsolete. The ideal resolution from the neoclassical point of view would be to admit just a 'little bit' of indeterminism into neoclassical models; just enough to resonate with the evolution of physics, but not enough to relinquish the original commitment to deterministic explanation and utility maximization. Something very much like this ideal package was put together by a new generation of neoclassicals mentioned above in the 1930s and 1940s, and the label on the package read "Econometrics". It is the purpose of this paper to investigate how one might begin to document this thesis.

The struggle to establish econometrics has frequently been misrepresented as a battle between the misguided partisans of "measurement without theory" versus the level-headed exponents of a judicious and balanced empiricism. The canonical text in this interpretation was the famous debate in the 1940s between Tjalling Koopmans and Rutledge Vining (Koopmans, 1947; Vining, 1949a). The problem with this interpretation is that it willfully disregards the fact that the main point of contention between the disputants was the validity of neoclassical economic theory, and not all theory tout court. The real battle was over what neoclassical economists were willing to "see" in the world: more precisely, what was the extent and meaning of the relative proportions of "order" and "chaos" in the economy? One disputant summed it up like this:

"I simply do not believe that any set of econometric models ... will ever suffice for reliable economic forecasting over any great length of time. The element of novel social conception is always breaking in. The social process, I submit, is peculiarly unfitted for description in terms either of simple positivism or mechanically imposed law. Such approaches result from the naive application of 'scientific' notions already out of date in their own field. Koopmans should realize that things have happened since the publication of Newton's Scholium. There is relativity, evolution and the quantum theory. Economics today is striving to become more mechanical and determinate at the very time physical science is loosening its bonds" (Wright, 1951, p. 147).

II. Stochastic economics as an anti-neoclassical movement

Upon reading the pioneering works of the earliest economic statisticians, one is immediately struck by the frequency and insistence with which they all appeal to physics for legitimacy. The very earliest economists who used least squares in order to empirically estimate what they called demand curves, namely Mackeprang, Benini and

Moore, all felt compelled to defend their curious activities from a multitude of detractors by invoking the name and power of science. The earliest instance of what was to become a litany is reported in (Moore, 1908, p. 24):

"Upon the occasion of the meeting of the Italian economists at Parma in 1907, Professor Benini, stimulated by the fact that economists were holding their sessions as one section of a general association of physical and natural scientists, urged that economists should adopt in their investigations the methods of research employed by the natural sciences. In particular, he advised the use of methods of interpolation for the discovery of empirical laws from statistical data, and instanced the utility of the evaluation of the laws of demand and supply."

The next thing which impresses the modern reader is the extent to which these appeals to science were unsuccessful in the first two decades of the twentieth century.² Strangely enough, none of those appeals struck a sympathetic chord with any prominent neoclassical economist, even though they were prone to make their own appeals to "science" in other contexts. In many instances Walras, Marshall and others made extremely disparaging comments about this sort of work, and used their influence to discourage interest in those directions.³ The fact that both the advocates and the detractors of statistical/probabilistic work could invoke the mantle of Science of indicative of the conflicting and confused images of science in the period. Recourse to explicit stochastic models was regarded in some quarters as a backsliding repudiation of scientific determinism, whereas in other quarters the mere fact a technique was used by astronomers was good enough to earn it the scientific stamp of approval.

One of the most curious aspects of the rise of neoclassical theory is that many of the leading lights of marginalism were also instrumental in the development of probability theory and statistics: Jevons, Edgeworth, Bowley, Keynes, Slutsky and Wald, only to name the most illustrious. And yet in the period roughly 1870-1925 none of these polymath theoreticians saw fit to forge direct explicit links between stochastic theory and neoclassical economic theory. This point is subtle, and deserves some careful attention (Menard, 1987). This is not to claim that none of the above ever prosecuted any empirical research program in economics broadly speaking: Jevons obviously used graphical techniques to discuss the price level, and his sunspot theory had some recourse to notions of correlation (Mirowski, 1984c). Edgeworth did work in the theory of index numbers; Bowley published empirical work on wages and the distribution of income; and so on. Rather, the point is that none of the first two generations of innovators of neoclassical price theory such as Jevons, Walras, Marshall, Edgeworth, Bowley and so forth felt compelled to link that particular theory to explicit empirical evidence or to pollute their value theory with stochastic concepts.

For instance, Bowley thought statistics could merely serve as a source of facts, and not as a means of testing functional forms (Darnell, 1981, p. 148). Edgeworth dashed any

hopes of obtaining demand curves by means of statistical observation as "chimerical" (Edgeworth, 1894, p. 473; see also Edgeworth, 1925, vol. I, p. 8), and wrote that, "There is really only one theorem in the higher part of the calculus [of probabilities], but it is a very difficult one, the theory of errors, or deviations from an average. The direct applications of this theory to human affairs are not very considerable" (Edgeworth, 1925, vol. II, p. 287). Indeed, Edgeworth could write an entire paper entitled "Applications of Probability to Economics" (Edgeworth, 1925, vol. II) which had no statistical theory or applications in it!

The irony was that, prior to 1930, it was mainly (but not entirely) economists who distanced themselves from the neoclassical research program and were openly skeptical about neoclassical price theory who pioneered the explicit connection between a stochastically informed empiricism and some sort of economic theory. (The primary exception to this generalization were the agricultural economists, who deserve, but have not yet received, separate attention from historians.)

In the United States, the center of the heterodox movement to construct a stochastic economics was located at Columbia University, and included in its ranks Wesley Clair Mitchell, Frederick Mills and Henry Moore. Mitchell was a student of Thorstein Veblen and a leading partisan of the American Institutional school of economics. Anyone who reads Mitchell's lecture notes on Types of Economic Theory will attest to his profoundly skeptical posture with regard to neoclassical economic theory (Mitchell, 1969). A formidable academic entrepreneur, Mitchell helped found the NBER in 1920, a research organization whose task in its early years was to implement an empirically grounded economics and encourage the development of a modern 'scientific' economics. In his Presidential Address to the American Economics Association Mitchell decried the similarity of orthodox economics to the older Lagrangean mechanics, and suggested economic theory be reconstructed in the pattern of more modern physics (Mitchell, 1925).

One of the first members of the NBER staff was Frederick Mills, another Columbia economist. In a manifesto proclaiming the arrival of a "New Economics" in 1924, Mills denounced the mechanical ambitions of neoclassical theory, and quoted James Clerk Maxwell to the effect that, "The scientific view of nature is thus... neither purely historical nor purely mechanical, it is statistical" (Mills, 1924a, p. 39). Amongst other ringing phrases, he proclaimed that, "In dealing with social and economic phenomena the existence of variation and the consequent inapplicability of the mechanical method must be realized", that "the concept of law is quite inconsistent with the statistical view of nature", and that, "Of contemporary writers, Henry L. Moore has not only clearly described the essential characteristics of statistical laws, but has demonstrated the fruitfulness of the concept when applied to economic problems" (Mills, 1924a, pp. 41, 42, 45).

The opposition of Mitchell and Mills to neoclassical theory and its predispositions went well beyond empty appeals to statistical mechanics or evolutionary biology: it also involved opposition to the conceptualization of stochastic phenomena as "errors" superimposed upon a deterministic structure. It is a little-known fact that doubts about the dependence upon Gaussian distributions and central limit theorems were broached as early as 1915 by, among others, Wesley Clair Mitchell, and that there have been subsequently a parade of empirically inclined economic researchers who have arrayed themselves against the notion of a normally distributed perturbations impressed upon a lawlike neoclassical model.⁴ In the 1920s people at NBER who looked at distributions of price changes noted that they had "tails" that were much too fat to qualify as Gaussian; (Mills, 1927, p. 336) was perceptive enough to notice that this phenomenon might imply "infinite probable error".⁵ These observations were persistently ignored by neoclassical critics, who instead excoriated Mills and the NBER for abjuring the guidance of neoclassical theory in their inquiries (Bye, 1940; Marschak, 1941; Koopmans, 1947).

But the person most responsible for prodding the neoclassicals and forcing them to sit up and take notice in the first few decades of the century was Henry Ludwell Moore. It was he whom Henry Schultz declared was single-handedly responsible for the creation of the new field of the statistical study of demand (Schultz, 1938, p. 63). This statement has irked later generations of econometricians, and indeed, is false from a number of points of view: first, others had fitted things which looked like demand curves before Moore; but second, a careful reading reveals that Moore was not engaged in simply trying to implement the work of Marshall in the empirical sphere; rather, he considered himself to have embarked upon the creation of an alternative and highly novel "synthetic economics". Moore has been ill-served by historians such as (Stigler, 1962) who try to portray him as a fledgling neoclassical econometrician, but then find they must suppress or explain away his numerous eccentricities. What is most essential for our present thesis is to observe that as his empirical concerns ripened, Moore evolved into a sharp critic of neoclassical theory.

In his earliest writings, such as (Moore, 1908), he assumed the mantle of a proselytizer within the neoclassical camp for an explicitly statistical economics, praising Cournot, Edgeworth and Pareto, and hinting that economic laws might be initially discovered by empirical statistical methods, and then later rationalized using the principles of "pure economics".⁶ Yet in 1912 Moore's hopes had been rudely discounted by Marshall; and further, his request of 1908 to Walras concerning his attitude towards statistics was never answered (Jaff6, 1965, vol. III, letter 1685). It seems that by 1914, Moore's view of neoclassicals, and perhaps of neoclassical theory, had grown a little jaundiced. After noting, quite correctly, that the statistical requirement of a large quantity of observations would dictate the collection of data over time, and deducing that it was inevitable that this practice would violate the orthodox *ceteris paribus* conditions and

the static framework of neoclassical price theory, he gave vent to a barrage calculated to provoke even the most thick-skinned of his colleagues:

"In the closing quarter of the last century great hopes were entertained by economists with regard to the capacity of economics to be made an 'exact science'. According to the view of the foremost theorists, the development of the doctrines of utility and value had laid the foundations of scientific economics in exact concepts, and it would be possible to erect upon the new foundation a firm structure of interrelated parts which, in definiteness and cogency, would be suggestive of the severe beauty of the mathematico-physical sciences. But this expectation has not been realized... The explanation is found in the prejudiced point of view from which economists regarded the possibilities of science and in the radically wrong method which they pursued. It was assumed gratuitously that economics was to be modeled on the simpler mathematical, physical sciences, and this assumption created a prejudice at the outset both in selecting the data to be investigated and in conceiving the types of laws that were to be the object of research. Economics was to be a 'calculus of pleasure and pain', a 'mechanics of utility', a 'social mechanics', a 'physique sociale'... The biased point of view implied in these descriptions led to an undue stressing of those aspects of the science which seemed to bear out the pretentious metaphors. One would naturally suppose from this manner of conceiving the science that the economic theorists would have at once entered upon their task with the methods that had proved themselves useful in the physical sciences. But this they did not do. They seemed to identify the method of the physical sciences with experimentation, and since, as they held, scientific experimentation is impossible in social life, a special method has to be devised. This invention was a disguised form of the classical *ceteris paribus*, the method of the static state" (Moore, 1914, pp. 84-86).

Moore coupled this attack with his own preferred method, the ordinary least squares estimation of a brace of curves relating agricultural crop harvests to their prices, as well as pig iron production to its price. In the latter case, he claimed to have discovered an upward sloping "demand curve". Subsequent commentators went to great lengths discounting this finding by elaborating Moore's "mistakes" (e.g., Ezekiel, 1928; Wright, 1930; Stigler, 1962; Epstein, 1987), especially with regard to what was later dubbed "the identification problem" and the conflict with Marshallian theory. Yet all these commentators start from the mistaken premise that Moore intended to empirically implement Marshallian concepts, a practice that he later explicitly renounced⁷ (Moore, 1929, p. 8). If Moore had a precursor, it was more nearly Cournot: both searched for phenomenological regularities cast in the form of a mathematical function; for Moore, these were to be linked in a chain to develop an explicitly macroeconomic theory of business cycles. In Moore's opinion, this was a 'scientific' theory because it linked business fluctuations to exogenous 'physical'

determinants, or as he put it, "The law of the cycles of rainfall is the law of the cycles of the crops and the law of Economic Cycles" (Moore, 1914, p. 135).

In short, by the 1930s some new lines were being drawn. On the one side were many of the neoclassical, holding tight to a incorrigibly deterministic physical metaphor from the mid-nineteenth century and the mathematics of constrained optimization, wary of any hint of indeterminist or historicist arguments and skeptical of the possibility of any theoretically informed precise empiricism; and on the other side, there were those who had started down the road of co-opting the metaphors of the newer sciences, be it statistical mechanics or some version of Darwinian evolutionary theory, stressing the mathematics of probability and dispersion, flirting with a mixture of historicism and institutionalism. Both were intent on occupying the higher ground of 'science' in the debate; but the latter group had a leg up, since they could make reference to a more recent vintage of physics or biology, as well as trumpeting their greater respect for the practice of "scientific empiricism". To once again quote Henry Moore (for he is so eminently quotable):

"But what is the source of the sense of unreality so many experience after having heroically struggled through the writings of Walras and Pareto? ... Foremost among the causes of the sense of unreality are these: the method of proceeding by successive approximations in the approach to a theory of general equilibrium, which give a feeling of indefinitely postponed real solution; the use of the hypothesis of perfect competition with a meaning that does not accord with reality; the limitation of all conclusions to the static state, when, as a matter of fact, all economic phenomena are in perpetual flux; the assumption of an immediate adjustment of changes, when in reality there are always leads and lags; the complexity of the functions which must be derived from reality and the absence of any known method of making the derivation; the assumption that the simultaneous equations may never be solved, first, because their empirical forms can never be known. . ." (Moore, 1929, p. 29).

Perhaps this situation would never have evolved beyond such guerrilla warfare had there not been an incursion of a third destabilizing force into the battlefield in the mid-1920s. The new quantum mechanics of Heisenberg, Born and Jordan irreversibly changed the ontological position of stochastic concepts in science. As (Eddington, 1935, pp. 77-78) put it, "The formulae given in modern textbooks on quantum theory... are explicitly concerned with probabilities and averages... But further it is now recognized that the classical laws of mechanics and electromagnetism. . . are simply the limiting form assumed by the formulae of quantum theory when the number of individual quanta or particles is very large. This connection is known as Bohr's Correspondence Principle. The classical laws are not a fresh set of laws, but a particular adaptation of the quantum laws."

It is simply not feasible to provide a summary in this venue of quantum mechanics for the uninitiated, although Eddington is as good as a guide as any to what was widely

believed in the mid-1930s. For a more modern and accessible primer one might consult (Davies, 1980; Jordan 1986; Cropper, 1970; Landshoff & Metherell, 1979). Nevertheless, even a superficial acquaintance with quantum mechanics would reveal that it clearly tipped the balance in favor of the challengers to neoclassical price theory because it suggested that all deterministic laws were merely limiting cases of a more fundamental stochastic substratum. In such an altered environment, neo-classical economists had two choices. Either they could withdraw deep within the obscurity of the Walrasian ramparts, or else they could try and strike some sort of accommodation with the new stochastic worldview. The Ostrich Option was advocated by many such as (Robbins, 1932, p. 101) who asserted without any justification that statistical laws discovered by the partisans of a stochastic economics could not possess the same status as statistical laws generated by the natural sciences. The Munich Option was the essential strategy of the partisans of the Econometrics Revolution.

III. How quantum mechanics influenced early econometrics

In this reading of the events of the 1930s-1960s, the rise of econometrics as a distinct subfield of economics was not a unified evolutionary process, but rather the multifaceted reaction to a logical crisis in the evolution of the conception of scientific explanation in the research program. Not everyone saw the potential set of problems being linked in the same manner, nor did everyone see the potential set of solutions (or indeed, "econometrics" itself) as comprising an identical set of practices; but it is a fact that the entire discourse was conducted by continuous reference to images of "science". Only in retrospect can be observed that the problems of determinism vs. indeterminism, abstract theory vs. empiricism, statics vs. dynamics, and subjectivity vs. objectivity, and individualism vs. macroscopic laws were all bound up together as problems created by the original physics metaphor which prompted neoclassical theory in the 1870s, and therefore it was neoclassical theorists who were most concerned to find a potential solution under the rubric of "econometrics".

While the advent of quantum mechanics precipitated "econometrics" out of the turbulent cross-currents of early 20th century economics, it did not determine its final structure in anything near the same way that 19th century energy physics determined the structure and formalisms of neoclassical price theory. The early econometricians did not copy, say, the wave equation term for term; for among other reasons, their prime objective was to preserve their 19th century price theory, not to supersede it. Hence the dowry bequeathed to economics by quantum mechanics is rather subtle, and historical evidence will not generally take the format of some key figure blithely asserting they were nothing more than an imitator of Heisenberg or Schrodinger or Born. Instead, historical inquiry will uncover two classes of evidence: suggestions as to the ways in which quantum mechanics focused attention on certain formalisms or structures of explanation (because there was as yet no such thing as a single "correct" stochastic formulation of an economic problem); and evidence as to the personal

influences of quantum mechanics on some of the primary actors. Let us first make a brief list of possible paths of influence, starting from the broadly philosophical and ending with the narrowly technical; and then indicate some relevant biographical particulars.

[1] First and foremost, the success of quantum mechanics fostered the impression that stochastic explanation was eminently scientific, giving it sanction as being grounded in Nature, as indicated in the first section of this paper. Hence the neoclassical school was now willing to co-opt much of the language of the appeals to science of their opponents, saying that, ". . . each of the so-called laws of nature is essentially statistical" (Roos, 1934, p. 9); or "The class of scientific statements that can be expressed in probability theory is enormous. In fact, this class contains all the 'laws' that have, so far, been formulated" (Haavelmo, 1944, p. iv); or, "The analysis of time series has also revealed the present status of economics as a science" (Davis, 1941a, p. 579). But beyond that, it also changed the rules as to what level of analysis was appropriate for the introduction of stochastic considerations. Prior to quantum mechanics, statistical considerations in thermodynamics entered in at the macro level, purportedly due to the inability to track every single individual particle. Now atoms themselves had stochastic aspects, and this encouraged neoclassical economists to likewise allow their appearance at the level of individual behavior (Morgan, 1987a, p. 186). As Roos (1934, p. 8) observed:

"There is necessarily little use made of the static theory of utility and its relation to demand. This does not mean that utility has nothing to do with the determination of demand... As long as physicists dealt with composites of molecules and atoms they were able to discover useful laws. When they attempted the problem of analyzing the atom, they soon came upon the problem of indeterminacy. In economics the individual occupies a role closely analogous to that of the atom in physics."

[2] However, the vintage of quantum mechanics most familiar to the partisans of the econometrics revolution was the sort that favored classical interpretations of quantum relationships. It took quite a while after the genesis of the quantum formalism for physicists to come round to a consensus that it fundamentally ruled out determinism; and by and large the early econometricians were not students of those who advocated such a drastic renunciation. For instance, Tinberger's familiarity with Ehrenfest (see below) may have included familiarity with the "Ehrenfest theorem" of 1927 (Ehrenfest, 1959, pp. 556-558), which stated that the mean of electron states follows a classical trajectory. Likewise, Koopmans' tutelage under Hendrik Kramers may have caused him to share Kramers' "unwillingness, or inability, to detach himself from classical visualizable physics or to make a decisive, irreversible break with classical concepts" (Dresden, 1987, p. 429). This would have translated into a belief that one could ultimately reconcile the deterministic neoclassical model with stochastic considerations in the economic context as well.

[3] Quantum mechanics also suggested a way out of the impasse of importing Fisherian-style maximum likelihood techniques from a context of sampling designs in controlled experiments into a context where such samples were impossible. As Haavelmo (1944, p. 15) insisted in his famous manifesto, "If every theory should be accompanied by a carefully described design of experiments, much confusion on the subject of constant versus changing economic 'laws' would be cleared up"; a dictum seconded by Koopmans (1937, p. 7): "it is clear that the expression 'repeated sampling' required an interpretation somewhat different from that prevailing in the agricultural and biological field". Haavelmo's proposal was to redefine the meaning of "population" and "sample" in economics: the population could be conceptualized as all possible economic decisions (or, what is equivalent, all virtual "economies"), whereas the sample would be the actual observed realizations (Haavelmo, 1944, pp. 51-52). This is, of course, a metaphor: there is no distribution of ghostly 1988 GNP growth rates from which the actual experience is drawn. Nevertheless, the metaphor was rendered more plausible by its resemblance to the new doctrines in quantum mechanics, in particular the concept of superposition of states. It is no accident that many supposed descriptions of underlying stochastic processes in later econometrics resemble the "many worlds" interpretation in later quantum mechanics.

[4] It is well known that quantum mechanics began with explicit consideration of some simple models of oscillators, such as Planck's work on black-body radiation. Heisenberg's first matrix exercise was written for a simple oscillator (Jordan, 1986, chap. 18). Familiarity with such models may have seemed to suggest connections with simple oscillators in economic models, and thence to possible explanations of time series movements and the "business cycle". This process began with the work of Davis (1941a, pp. 35-37) Tinbergen, and Roos, but was best represented by Frisch's conception of "propagation and impulse problems". As he wrote, "I believe that this idea will give an interesting synthesis between the stochastic point of view and the point of view of rigidly determined dynamical laws" (Frisch, 1933b, p. 198 in 1965 edition). Frisch compared to the economy to a pendulum, claiming that the configuration of its swings could be accounted for by an abstract theory of dynamics, but the actual energy of motion would be traceable to "shocks" impinging upon the pendulum from outside the system. These shocks were then identified with stochastic phenomena. The use of linear operators and Fourier decompositions by Frisch (1933b, p. 181 in 1965 edition) and Davis (1941a, pp. 61 et seq.) were unusual for that time, and probably indicated a familiarity with quantum mechanics, or at least the physics of vibrations of elastic solids (Epstein, 1987, p. 62).

Neoclassical economists rapidly became enamoured of this metaphor, irrespective of the fact that they still were bereft of any legitimate dynamic theory, because they could place all the technical emphasis upon stochastic shocks as the ultimate cause of

the movement of the economy (Koopmans, 1947, p. 171). Hence, econometrics bequeathed to neoclassicism an ersatz dynamics, or as Samuelson (1965, p. 147) put it:

"Why should a person interested in economics... spend time considering conservative oscillations of mechanics? Experience suggests that our dynamic problems in economics have something in common with those of the physical and biological sciences. . . Just as Ehrenfest and other physicists had to add probability to the causal systems of physics in order to get around the time-irreversibility feature of classical mechanics that was so inconsistent with the second law of thermodynamics, so we must, in the interests of realism, add stochastic probability distributions to our economic and biological causal systems."

This problem was misconstrued by early econometricians such as Haavelmo (1944, p. 20), who actually claimed that irreversible phenomena could be simply reduced to a mechanical (i.e., reversible) format by the mere addition of more relevant independent variables.

[5] Quantum mechanics may also have inadvertently focused attention upon a particular conception of "structure" which then became the rallying cry of the Cowles Commission's program of "structural estimation". Tjalling Koopmans' mentor, Hendrik Kramers, developed a particular program of explanation in quantum electrodynamics in the early 1930s (Dresden, 1987, p. 341) which consisted of the following steps: (1) start from the classical model of the extended electron; (2) construct a Hamiltonian for the system consisting of electrons and radiation; (3) separate the analytical description of the system into structure-dependent and structure-independent parts; and (4) eliminate structure-dependent features by canonical transformations. As Dresden (1987, p. 343) put it, "To Kramers, 'the elimination of the structure' almost became a slogan". There would appear to be strong resemblances to the Cowles program for econometrics in the 1940s, with its stress on adherence to the neoclassical model, deriving behavioral structural equations from constrained optimization, and isolating the "structure" from stochastic components by means of the process of "identification". Indeed, the whole purpose of canonical transformations is the production of ignorable variables by means of transformations of the original set of coordinates (Lanczos, 1949, chap. 7), which is very similar to the problem of identification.

[5] With respect to the actual formalisms of econometrics, the progenitors did not actually copy quantum mechanics so much as they imitated some of the mathematical formalisms which were characteristic of quantum mechanics. Although matrix algebra was formulated by Cayley and others in the nineteenth century, it was first extensively introduced into physics with the Born-Jordan-Heisenberg formalization of matrix mechanics in the mid-1920s (Mehra & Rechenberg, vol. III, 1982, pp. 65-71). In matrix mechanics, which was a precursor to the full-blown quantum mechanics, the concern was to manipulate matrices of position and momentum coordinates of electrons. In

the elaboration of matrix mechanics, one primarily has recourse to Hermetian matrices: that is, matrix transposes result in a symmetric matrix of complex conjugates. One advantage of a familiarity with Hermetian matrices is that if one arbitrarily restricts the elements of the matrix to be real numbers, then the Hermetian matrix is simply a symmetric matrix, and all analytical results automatically carry over for such matrices. One significant aspect of Hermetian matrices for Born and Jordan was their natural link to the mathematics of quadratic forms, and the reduction of finitely many complex variables to a sum of squares (Mehra & Rechenberg, 1982, vol. III, pp. 120-123).

Potential applications in the formalization of multivariate least squares estimation of this matrix framework by someone schooled in quantum mechanics must have seemed plentiful and obvious. Indeed, it was Frisch and Koopmans who introduced matrix algebra, quadratic forms, the eigenvalue problem and symmetric matrices into econometrics (Koopmans, 1937, pp. 10-13; Frisch, 1934, ch. 1). Koopmans was referring to just this phenomenon in a 1977 unpublished lecture to the Operations Research Society:

"There is the precedent in physics of the 30's when matrices and wave equations came in with quantum mechanics, causing a great deal of distress to physicists who did not have previous exposure to these tools. Similar things happened in statistics in the 30's and 40's, and in econometrics and mathematical economics in the 40's and 50's; in both cases an upward jump in what you might call the intensity of the mathematics used."⁸

[7] The concern of early quantum mechanics with quasi-classical Gaussian states (Holevo, 1982, Ch. 5) may help account for the focus of attention of the early econometricians upon Gaussian distributions to the exclusion of those considerations raised by Mitchell, Mills and others that economic variates generally did not conform to normal distributions.

And now for some brief biographical evidence. The connections of the early Cowles/Econometric Society group with contemporary physics has not received the attention which it deserves. Here we shall just restrict ourselves to a subset of the relevant cast of characters.

Jan Tinbergen, one of the earliest advocates of econometric model building, actually received his Doctoraat in 1929 from the University of Leiden in physics. According to (Klein, 1970, pp. 305-306). He was encouraged to venture into economics by his advisor, the renowned physicist Paul Ehrenfest, one of the pre-eminent authorities on statistical mechanics of the early twentieth century. His thesis (Tinbergen, 1929) was concerned with the connection between variational principles and conservation principles-the critical issue in Hamiltonian dynamics-and their implications for Lorentz's classical model of the electron. In a retrospective interview he suggested he

undertook his physics studies because he "hoped to get sufficient capability to handle things with mathematics and perhaps take physics as an example of a more developed science than economics" (in Magnus & Morgan, 1987, p. 118), but it seems clear that his respect for the physics paradigm waned over the course of his career. His earliest work, true to the format of our hypothesis, began with the application of complex functions in order to render neoclassical theory "dynamic" (Tinbergen, 1933), and to estimate oscillatory solutions to aggregative difference equations using least squares. Yet he revealed an early disinclination to defend the neoclassical theory of pricing and allocation, as well as a lack of interest in segregating deterministic "laws" from stochastic "errors" (Tinbergen and Polak, 1950, p. 78-Dutch edition 1942). He was too early to be a student of Kramers; and throughout his life showed little interest in problems of "identification" and simultaneous equations estimation. Perhaps this reproduces the classical stance of his advisor Ehrenfest towards quantum mechanics, and goes some distance in explaining his later withdrawal from "econometrics" and his gravitation towards work of a more conventional political character (Epstein, 1987, pp. 127-8).

An econometrician closer to our model was Ragnar Frisch. Frisch took a Ph.D. in mathematical statistics from the University of Oslo in 1926; biographical material on his life is somewhat scarce because of the language barrier (ours, not his) and the inaccessibility of his papers at the University of Oslo. Nevertheless, numerous comments in his published writings suggest a familiarity with the physics of his time. Frisch's earliest work explicitly discussed the analogies between neoclassical price theory and rational mechanics, and proposed the use of least squares models to explain errors or divergences from that model (Frisch, 1926, p. 19). By 1929 he was troubled by the possibility of "fictitious determinateness created by random errors" (Frisch, 1933a, p. 9; 1934, p. 6), and proposed his method of "confluence analysis" to uncover the "true determinate" system from the data, using the aforementioned matrix analysis to initiate the discussion of what was later called the identification problem. In 1937, Frisch was lecturing the Cowles researchers on physical analogies with oscillatory equations in economics, and in 1933 he produced the "solution" to the problem of neoclassical dynamics mentioned above. Although Frisch served as the earliest editor of *Econometrica*, it seems he also grew disenchanted with the program in later life, effectively renouncing econometrics as defined by the Cowles Commission by the early 1960s (Epstein, 1987, pp. 127-8).

Another important Leiden alumnus was Tjalling Koopmans, who earned a Ph.D. in mathematical statistics in 1936, following an M.A. in physics at Utrecht in 1933. Koopmans' advisor was the pre-eminent quantum physicist outside of Copenhagen, Hendrik Kramers; and indeed, Koopmans' first academic publication was in quantum physics. The title of (Koopmans, 1934) could be roughly translated as "On the

Relationship of Wavefunctions and Characteristic Values to Individual Electrons".⁹ His subsequent move to economics is best described in his own words:

"Why did I leave physics at the end of 1933? In the depth of the worldwide economic depression I felt that the physical sciences were far ahead of the social and economic sciences. What held me back was the completely different, mostly verbal, and to me almost indigestible style of writing in the social sciences. Then I learned from a friend that there was a field called mathematical economics, and that Jan Tinbergen, a former student of Paul Ehrenfest, had left physics to devote himself to economics. Tinbergen received me cordially and guided me into the field in his own inimitable way. I moved to Amsterdam which had a faculty of economics. The transition was not easy... Also, because of my reading block, I chose problems that, by their nature, or because of the mathematical tools required, have similarity with physics."¹⁰

The areas of similarity with quantum mechanics, such as Kramers' program, the use of matrix methods, and the general stance towards stochastic phenomena have been mentioned above. Koopmans became Research Associate at Cowles from 1944-48, and Director of Research from 1948-54, the periods in which the Cowles approach to econometrics was promulgated.

Some of the economists affiliated with the early Cowles Commission also came to economics from backgrounds in mathematical physics. Charles Roos was a 1926 mathematics Ph.D. from Rice University with a minor in physics. In 1931, Roos was one of the founding members of the Econometrics Society. One of the most prolific and sophisticated, Harold T. Davis, was a 1926 mathematics Ph.D. from Wisconsin who wrote extensively on physics (Davis, 1931). Davis made repeated references to physical analogies in his published works, especially (Davis, 1941a; 1941b), where among other analogies, he developed a formal comparison of the variance of a timeseries to the physical concept of energy (Davis, 1941a, p. 178). But more importantly, Davis' frequent talks and memorandums to both the Cowles Commission and to the Econometrics Society in the 1930s make even more explicit reference to the issues discussed in this paper. Quoting a paper read to a Cowles conference in 1937: "The physicist in his study of microphysics, that is to say, the phenomena of the electron and the quantum, has encountered difficulties similar to our own. Heisenberg was led to postulate the existence of a rectangle of indeterminacy, within which complete statistical chaos existed. Perhaps the main difference between our problem and theirs, statistically considered, is merely in the size of the price-time or production-time rectangle."¹¹

Hence, many of the new generation of neoclassicals associated with the Cowles Commission and the Econometrics Society were eminently well-equipped to attempt to effect a reconciliation between neoclassical economics and the new physics. This new cohort recognized that neoclassical theory was an imitation of nineteenth century energetics, and they openly said so (Davis, 1941, pp. 49, 577; Koopmans, 1957, p. 176;

Tintner & Sengupta, 1972; p. 9; Henry Schultz to Harold Hotelling, letter, May 11, 1932, Hotelling papers). They were trained in universities in Holland and the U.S. where quantum mechanics was regarded as the vanguard of the new physics, and which were avant garde in their reception of the new doctrines (Cobden, 1971). Even though many of their approaches were ultimately different, in the very early stages of the reconciliation of neoclassical theory and stochastic concepts they recognized one another as comrades in arms, engaged in essentially the same research project.

The point here is not that all the protagonists wanted to copy quantum mechanics: rather, it is that some common vocabulary was needed to accommodate the economic orthodoxy with the novel cultural currents. The real underlying issue was order versus chaos: any old advocacy of stochastic descriptions of economic phenomena would not do, because the conception of the market as a natural organizing force was at risk. The subtle shading came, not with the mathematics per se, but rather with the accompanying assertions that a rational stochastic economics was not irreducibly random: it simply reconciled stochastic disturbance and deterministic law into a tidy package. As (Koopmans, 1937, pp. 5-6) insisted, "Following Frisch, each of the variables may be conceived as the sum of two components, a 'systematic component' or 'true value' and an 'erratic component' or 'disturbance' or 'accidental error'. The systematic components are assumed to satisfy the regression equation exactly. .. the erratic component is taken as error in the literal sense of the word."

Might it be possible to suggest that, just as in the case of the history of physics, once stochastic concepts are admitted into the fold, one progressively discovers that the damage to the deterministic world view is not so easily localized? And that, just perhaps, the more that we find out, the more meaningless the paradigm of constrained optimization becomes?

B.4 P_Pov

Joseph W. Hogan & Rusty Tchernis

Bayesian Factor analysis for spatially correlated data, with application to summarizing area-level material deprivation from census data

1. Factor analysis and spatial data

Factor-analytic models are useful for summarizing variance and covariance patterns in multivariate data. A common formulation of factor analysis assumes that measurable variables, such as scores on a test, are manifestations of an underlying latent construct, such as ability or intelligence. The latent variable formulation can be useful for data reduction, that is, summarizing multivariate observations using a lower dimensional variable. A thorough review has been given by Bartholomew and Knott (1999, chaps. 1–3). Recent work from a Bayesian perspective has been done by Geweke and Zhou (1996), Press and Shigamesu (1997), Aguilar and West (2000), and Rowe (2002).

Multivariate spatial data can arise in a number of applied contexts. Wang and Wall (2001, 2003) studied multivariate indicators of cancer risk across counties in Minnesota. Samet, Dominici, Curriero, Coursac, and Zeger (2000) considered the effect on mortality of multiple measurements of air pollution exposure in 20 U.S. cities. Lee, Murie, and Gordon (1995) summarized a variety of methods for combining multiple area specific census variables into scalar measures of socioeconomic standing. The need to analyze or summarize multivariate data that are spatially aligned suggests the utility of factor-analytic models that can incorporate spatial covariation. Wang and Wall (2001, 2003) recently introduced a factor-analytic model for spatially correlated multivariate cause-specific mortality, which provides area-specific scalar summaries of mortality via factor scores. Their first article studied large-sample properties of the model from a frequentist perspective, and the second developed a Bayesian approach that allows discrete manifest variables. Christensen and Amemyia (2002, 2003) developed semiparametric latent variable models for rectangular grids and provided key references to related work in geography and geology.

In this article we use a Bayesian factor-analytic model of spatially correlated data to summarize area-specific material deprivation from multiple census variables. Spatial correlation is modeled on the latent variable scale and can be specified either marginally or conditionally (e.g., conditional autoregressive structures). The Bayesian approach is natural, because the factor-analytic model is hierarchical in nature; furthermore, it confers the distinct advantage that uncertainty about factor scores, which in many cases are of direct interest, can be accurately summarized as a natural byproduct of the posterior parameter distribution. Moreover, constraints on the variance structure that must be imposed for certain CAR specifications pose no special problems in the estimation routines.

The rest of the article is organized as follows. Section 2 is devoted to providing background on the uses of material deprivation indices, and motivates their construction via factor analysis. Section 3 describes a general approach to incorporating spatial correlation in a factor-analytic model, gives several examples of parameterizations, and highlights key differences between marginal and conditional specifications. The development in Section 3 is given in context of measuring material deprivation from multivariate census variables. Section 4 applies several models to census data from Rhode Island, and Section 5 concludes.

2. CONSTRUCTION AND USE OF MATERIAL DEPRIVATION INDICES

2.1 Conceptualization and Measurement of Material Deprivation

Deprivation itself is typically viewed as an underlying construct rather than a measurable characteristic, although it manifests in measurable ways. Deprivation is related to but distinct from poverty, which is a measurable quantity defined in terms of per capita income. Material deprivation is meant to encompass a broader spectrum of access to the materials needed for daily living, and thus is a potentially richer construct than poverty per se.

Indices of material deprivation usually are derived from multiple components and computed in uncomplicated ways; a recent comprehensive review of indices used in the United Kingdom has been given by Lee et al. (1995). One widely used measure, the Townsend index (Townsend, Simpson, and Tibbs 1985) can be considered a metric on which to measure material resources and is representative. For a given area (e.g., census tract or zip code), the Townsend index uses information on four area-level measures: percent unemployed, percent of households without a car, percent of households not owner occupied, and percent of households with more than one person per room (i.e., percent crowding) (Townsend et al. 1985; Lee et al. 1995). Natural log transformations are applied to the unemployment and crowding variables, and each variable is standardized in terms of a Z score. The area-specific index is computed as an unweighted sum of the Z scores, with higher values representing greater deprivation; that is, for area i , the index is $T_i = D_i + Z1_i + C_i + Z2_i + C_i + Z3_i + C_i + Z4_i$. Krieger et al. (2001) described a Townsend index derived from U.S. Census variables, which we use for our analyses in Section 4. Other indices of material deprivation are based on similar constructions (Hutchinson, Foy, and Sandhu, 1989; Jarman 1983; Carstairs 1995; Lee et al. 1995).

2.2 Uses of deprivation indices

Area-level indices of material deprivation play an important role in public health and demographic research and in governmental policy. Public health researchers use these indices to study relationships between, for example, health outcomes and social status (see, e.g., Eachus et al. 1996; Wilson, Chen, Taylor, McCracken, and Copeland 1999; Sundquist, Malmström, and Johansson 1999). Governmental agencies, particularly in

the United Kingdom, incorporate information about area-level deprivation together with mortality rates for making decisions about resource allocation in relation to health care, because these are considered reliable proxies for morbidity and health services use (Hutchinson et al. 1989; Carstairs 1995).

The indices are derived from publicly available data and can be calibrated using data on morbidity and/or specific health outcomes, such as cancer, that are directly related to use of services but typically are available only from sources, such as registries, that are not publicly available. Thus one objective in this area of research is to develop indices with high construct validity as it relates to use of services (see, e.g., Carstairs 1995; Malmström, Sundquist, Bajekal, and Johansson 1998; Barnett, Roderick, Martin, Diamond, and Wrigley 2002). Barnett et al. (2002) provided a recent review of the key issues related to index use in resource allocation, and carried out a detailed analysis showing the relationships between mortality, morbidity, and several deprivation indices. A related and recently opened line of research in the United States concerns the use of area-based measures - in particular, those based on census variables - to quantify and monitor socioeconomic inequalities in health (Krieger et al. 2001, 2002).

Although some researchers object to deprivation indices as proxies of morbidity and health (e.g., Sheldon, Davey-Smith, and Bevan 1993), they remain in widespread use as explanatory variables for epidemiological studies and as components of resource allocation algorithms; furthermore, they show considerable promise for monitoring health inequalities. The primary motivation for our proposed methodology, therefore, is to formalize index development in the context of a model so that the resulting index makes better use of its component variables compared with the empirical indices, for example, by using spatial correlations to incorporate data from neighboring areas into the index, and by calculating measures of uncertainty for each area's index. A potentially important and useful output from the model is a database containing area-specific indices derived from the model, which can subsequently be used by researchers for validation studies. The discussion in Section 5 contains some specific recommendations in this regard.

2.3 Limitations of existing measures and motivations for model-based indices

Despite their intuitive appeal, empirical indices such as Townsend's are based on a number of important structural assumptions. Because the index is an unweighted sum of the Z scores, each variable contributes equally to measured deprivation; furthermore, even though areas may have different population sizes, no adjustment is made for differential precision of the component variables across areas. Moreover, it is assumed that for a specific area, information about deprivation depends exclusively on variables from that area, and not (for example) on variables from neighboring areas. Finally, once computed, the index lacks a measure of uncertainty. For policy-making decisions in particular, this last feature may be problematic, for example, if decisions about resource allocation are based on cutoff values or percentiles of the index.

An effective way to address these concerns is to cast the relationship between the component variables (Z 's) and the deprivation index in a modeling framework. We adopt a factor analytic structure, implemented in a fully Bayesian framework, in which the component variables are the "manifest variables" and deprivation is the underlying latent factor. The deprivation index is defined as the posterior expectation of the latent factor given the manifest variables and model parameters. Under certain distributional assumptions, this approach confers several advantages: first, the model-based deprivation index retains its simple structure as a weighted average of census variables, but the weights are functions of model parameters and therefore are informed by data; second, different sample sizes across areas are incorporated naturally and reflected in posterior variability of the indices; and third, deprivation indices are summarized not as single numbers, but rather in terms of posterior distributions that reflect important uncertainties about deprivation status.

A final limitation of empirical indices is that only data from the specific area are used. We address this shortcoming by introducing spatial dependencies between the latent deprivation indices, thereby incorporating information from census data of neighboring areas. If the underlying latent factor has multivariate normal distribution across the spatially aligned areas, then a wide variety of parameterizations incorporating spatial correlation are available, each with particular implications for model identification and posterior sampling. We consider both marginally and conditionally specified correlation structures; conditional specifications in particular imply marginal variance structures that are not compatible with the usual set of constraints in factor analysis, but the difficulty can be resolved by rescaling.

To illustrate this approach, we use data from the 1990 census to characterize deprivation in 228 of Rhode Island's 232 census tracts. (Two tracts are excluded because they are outlying islands and lack natural contiguous neighbors; two others are excluded because the observed value of one or more manifest variables is an extreme outlier.) We compare the Townsend index and associated ranked values with posterior distributions of model-based indices - and the posterior distribution of their ranks - under several different spatial correlation models. We illustrate two ways to summarize area-level deprivation so as to communicate both ordering and uncertainty.

3. Factor analytic models of material deprivation

Factor analysis provides a natural structure within which to generalize the Townsend and related indices because the latent index (or, more precisely, its posterior expectation conditional on manifest variables and model parameters) is a linear combination of the manifest variables (Bartholomew and Knott 1999). After defining relevant notation, we describe a factor analytic characterization of deprivation index under spatial independence and assuming that area-level samples leading to the manifest variables are of equal size.

The second part of this section is devoted to describing various parameterizations of spatial correlation and their implications for both model interpretation and posterior sampling. We describe both marginally and conditionally specified correlation structures, focusing on Gaussian conditional autoregressive (CAR) parameterizations for the latter.

3.1 Notation and general model structure

We describe the model in terms of census variables, but the ideas here apply more generally. Census variables used here represent proportions (e.g., percent unemployed) estimated from a fractional subsample of the full population. For spatial location i (e.g., block group, census tract), where $i \in \{1, \dots, N\}$, let S_{ij} denote the numerator, and m_i denote the denominator, for constructing the realized census variable $Z_{ij} = S_{ij}/m_i$ ($j \in \{1, \dots, J\}$). The denominator is the sampling fraction times the census tract population. For most areas, the sampling fraction is $1/6$ (U.S. Census Bureau 1999); for our example we assume a sampling fraction of $1/6$ for all areas. Further define $Y_{ij} = g(Z_{ij})$ as an appropriate transformation of the census variable, and let $Y_i = (Y_{i1}, \dots, Y_{iJ})^T$.

Following Cressie and Chan (1989), we use a square root transformation, $Y_{ij} = \sqrt{m_i} Z_{ij}$ ($j \in \{1, \dots, J\}$), to stabilize variances. Exploratory analyses of the Rhode Island census data indicate that the variance-covariance matrices computed from data grouped by quartiles of m_i are approximately equal, suggesting that $\text{var}(Y_{ij}) \approx m_i$ and, more generally, $\text{var}(Y_i) \approx T \text{diag}(m_1, \dots, m_N)$, where T is a 4×4 variance-covariance matrix. Define $M = \text{diag}(m_1, \dots, m_N)$ as the $N \times N$ matrix with m_i along the diagonal and 0's elsewhere, we have $\text{var}(Y) \approx M \otimes T$, where \otimes denotes the Kronecker product.

The general factor-analytic model assumes that each area has an L -dimensional ($L < J$) latent variable, $\pm_i = (\pm_{i1}, \dots, \pm_{iL})^T$, that fully characterizes socioeconomic characteristics, which in turn are manifest through Y_i . For our application to the Rhode Island data, we focus on the case of a univariate index ($L = 1$) and represent the model in hierarchical form. At level I, $Y_i | \pm_i \sim N(\mu_i, \sigma^2/m_i)$, where μ_i is a $J \times 1$ mean vector, σ^2 is a $J \times 1$ vector of factor loadings, and $\sigma^2 = \text{diag}(\sigma_1^2, \dots, \sigma_J^2)$ is a diagonal matrix measuring residual variation in Y_i . That σ^2 is diagonal implies independence between elements of Y_i conditionally on \pm_i . At level II, which in our case characterizes between-area variation, we make the customary assumption that $\pm_i \sim N(0, 1/m_i)$. The assumption of known variance ensures identifiability of σ^2 and μ_i (Bartholomew and Knott 1999). Level III specifies prior distributions on the unknown parameters μ_i , σ^2 , and σ^2 .

It is convenient to write the model in compact form. Let $Y = (Y_1^T, \dots, Y_N^T)^T$ represent the $NJ \times 1$ stacked vector of manifest variables (with a corresponding $NJ \times 1$ vector μ similarly defined), and let $\pm = (\pm_1, \dots, \pm_N)^T$ denote the $N \times 1$ vector of area-specific latent variables. Then levels I and II can be rewritten as

$$I. \mu_i \sim N(\lambda_i, \sigma^2); \lambda_i \sim N(\mu, \tau^2); \mu \sim N(\mu_0, \tau_0^2);$$

(1)

$$II. \mu_i \sim N(\lambda_i, \sigma^2); \lambda_i \sim N(\mu, \tau^2);$$

where $\Lambda \in N \times N$ is the $N \times N$ matrix of factor loadings and I_N denotes an identity matrix with dimension N . This is the spatial independence model, model 1.

3.2 Incorporating spatial dependencies and computing a model-based index

For geographic data, there are at least two motivations for incorporating possible spatial dependencies in the model. First, it is highly likely that adjacent areas have similar socioeconomic characteristics, so that ignoring spatial correlation may result in incorrect posteriors, particularly with respect to uncertainty measures. Second, if spatial dependence is present, then estimation of the posterior index distribution can incorporate information

from neighboring areas; specifically, under multivariate normality in levels I and II, the expectation of μ_i given the full vector of manifest variables Y will be a linear combination of both Y_i and Y_j for areas $j \in \mathcal{N}(i)$. A potential byproduct of incorporating spatial correlation is increased precision for parameter information gained by “borrowing of information” from neighboring tracts (see Sec. 4).

Similar to the approaches of Christensen and Amemiya (2002, 2003) and Wang and Wall (2001, 2003), we induce spatial correlation at level II. Generically, we assume that

$$\mu_i \sim N(\lambda_i, \sigma^2); \lambda_i \sim N(\mu, \tau^2); \mu \sim N(\mu_0, \tau_0^2);$$

where Σ is an $N \times N$ correlation matrix with 1’s on the diagonal and $\tilde{\rho}_{ij} \in \text{corr}(\mu_i, \mu_j)$ on the off-diagonal. When $\Sigma = I_N$, (2) reduces to model (1). The motivation for constraining Σ to have 1’s along the diagonal is to maintain the same scaling for μ_i in models with and without spatial correlation, that is, $\text{var}(\mu_i) = \sigma^2 / \tau^2$. The model-based deprivation index for area i is summarized by the conditional distribution of the “factor score” μ_i given Y and Σ . Following standard arguments for multivariate normal distribution, $\mu_i | Y, \Sigma \sim N(d_i, \sigma^2 / \tau^2)$; where

$$d_i = \tau^2 \Sigma^{-1} \Lambda^T Y_i / (\tau^2 + \sigma^2 \Lambda^T \Sigma^{-1} \Lambda);$$

$$d = D^{-1} \Lambda^T Y;$$

This implies that area-specific indices (d) are linear combinations of the (mean-centered) manifest variables. When $\Sigma = I_N$ (spatial independence), D is block diagonal, and the area-specific posterior expectation of μ_i given Y_i and Σ depends linearly on manifest variables from area i exclusively. More generally, when spatial dependence is nonzero, D is not block diagonal, and the index for area i may depend on manifest variables outside area i .

The matrix Σ of spatial correlations can be parameterized either marginally (Cressie 1993; Conlon and Waller 1998) or conditionally (Besag 1974; Sun, Tsutakawa, and

Speckman 1999). We use both approaches in our application. For the marginal specification, we assume that $\tilde{\alpha}_{ij} \propto \exp(-\beta d_{ij})$, where $\beta > 0$, to ensure that $\tilde{\alpha}_{ij} < 1$, and d_{ij} is the Euclidean distance between centroids of areas i and j , with $d_{ii} = 0$ by definition (Conlon and Waller 1998). For model identifiability, we further assume that α^2 is known ($\alpha = 1$) and refer to this specification as model 2.

Conditional autoregressive specifications of spatial dependency are useful for normally distributed data, because when the conditional distributions are assumed normal, the corresponding joint marginal distributions are multivariate normal and can be derived directly (Besag 1974). Sun et al. (1999) described a rather general structure for Gaussian CAR models.

Let R_i denote the set of indices for areas that are neighbors of area i ; typically, but not necessarily, neighbors are defined by adjacency. If

$$\pm_j \propto \sum_{j \in R_i} \alpha_{ij} (\pm_j - \mu_j) \quad (4)$$

then the joint marginal distribution of $\pm \propto (\pm_1, \dots, \pm_N)'$ follows $N(\mu, B^{-1})$, where B is an $N \times N$ matrix with α_{ii} along the diagonal and $-\alpha_{ij}$ on the off-diagonal, provided that B is symmetric and positive definite (Besag 1974; Sun et al. 1999). A simple and common parameterization of the CAR model gives some insight about parameter interpretation.

Let $R_{ij} = I_{j \in R_i}$ be the indicator that area j is a neighbour of area i , set $\alpha_{ij} = \alpha R_{ij}$, and hold α_i constant (e.g., $\alpha_i = 1$). Then α measures degree of spatial correlation and σ^2 measures residual variation. Elaborations allow overall residual variation σ^2 to depend, through α_i , on area-level characteristics such as number of neighbors (Bernardinelli, Clayton, and Montomoli 1995).

The factor-analytic model requires that the marginal variance of \pm_i be known. In our CAR models, we fix conditional variance, setting $\sigma^2 = 1$. We then rescale the implied marginal model such that $\text{var}(\pm_i) = 1$, so posterior inferences about factor loadings from CAR models can be directly compared with models where spatial correlation is specified marginally. Moreover, one or more parameters must be constrained to ensure that B is positive definite (Sun et al. 1999); the constraints are model specific. We use several CAR formulations to analyze the Rhode Island census data:

Model 3A defines R_i as the set of adjacent census tracts and sets $\alpha_{ij} = R_{ij}$, $\alpha_i = 1$, and $\sigma^2 = 1$, leading to $B = I - R$, where R is an adjacency (weight) matrix with $R_{ii} = 0$ and indicators $R_{ij} = I_{j \in R_i}$ on the off-diagonal, that is, $R_{ij} = I_{j \in R_i}$. Let $\lambda_1, \dots, \lambda_N$ denote the ordered eigen values of R ; then B will be positive definite only if $\lambda_1 < 1 < \lambda_N$ (Sun et al. 1999). The rationale for setting $\sigma^2 = 1$ is that if $\sigma^2 = 0$, then $B = I$, consistent with the convention of assuming known variance for \pm to ensure identifiability. However, it is important to recognize that the variance constraint is being imposed on the conditional distribution of the \pm_i given its neighbors, which in general is not consistent with assuming that σ^2 has 1's on the diagonal. Setting $\sigma^2 = 0$ implies model 1.

Model 3B defines R_i the same as in 3A, but sets $\tau_{ij} = 1/R_{ij}$, $\sigma_i = 1/n_i$, and $\rho = 1$ (where n_i is number of neighbors for area i). This yields $B = \text{diag}(n_i) - 1/n_i$. To ensure that B is positive definite, ρ has the same constraints as for model 3A. This model was suggested by a referee on the grounds that conditional spatial correlation $\text{corr}(y_i, y_j | y_{-i-j}) = \rho / (1 + \rho(n_i + n_j - 2))$ is not dependent on neighborhood size (Cressie 1993, sec. 7.6; Stern and Cressie 1999).

We also consider three CAR specifications that incorporate distance between tracts. The first, model 3C, is based purely on distance between neighbors (Best, Arnold, Thomas, Waller, and Conlon 1999), wherein R_i includes all areas except area i , $\tau_{ij} = \exp(-\rho_{ij} d_{ij})$, $\sigma_i = 1$, and ρ is unknown. The parameterization of τ_{ij} as exponential decay is informed by empirical correlograms for the manifest variables (Fig. 1). The remaining models, 3D and 3E, combine information on adjacency and distance. Centroids of adjacent Rhode Island census tracts are geographically very close in urban areas but not in rural areas, and spatial dependencies appear more pronounced in areas with closer neighbors (Fig. 2). Neighborhoods are defined by adjacency, and spatial correlation may depend on distance within neighborhood (Cressie and Chan 1989). If spatial correlation is positive, then neighbors whose geographic centers are close (e.g., northeastern urban centers) will exhibit greater spatial dependency than those whose centers are further apart (e.g., rural western part of the state). In model 3D, $\tau_{ij} = \exp(-\rho_{ij} d_{ij}) / R_{ij}$ and $\sigma_i = 1$. It is necessary to choose one of the two unknown parameters; following Cressie and Chan (1989), we fit the model using point mass priors for ρ to moderate the role of distance within neighborhood structure. In model 3E we use normalized weights $\tau_{ij} = \exp(-\rho_{ij} d_{ij}) / R_{ij} = \sum_j \exp(-\rho_{ij} d_{ij}) / R_{ij}$ and $\sigma_i = \sum_j \exp(-\rho_{ij} d_{ij}) / R_{ij}$. As with 3E, we use point mass priors for ρ .

3.3 Compatibility between marginal and CAR models

Posterior factor loadings from CAR models cannot be directly compared with those from marginally specified models because they are measured on different scales, due to the latent factors \pm having different variances in the two types of models. At the second level of marginally specified models (e.g., models 1 and 2), we assume that $\text{var}(\pm) = \Sigma$, where Σ is a correlation matrix having 1's along its diagonal. On the other hand, the second level of a CAR model implies $\text{var}(\pm) = B^{-1}$, which in turn implies that $\text{var}(Y) = 3B^{-1} \Sigma^{-1} \Sigma^{-1} \Sigma^{-1}$. Note, however, that because parameters in CAR models 3A–3E are constrained such that B^{-1} is symmetric and positive definite, we can write $B^{-1} = V^{-1} \Sigma^{-1} V^{-1}$, where V is an $N \times N$ diagonal matrix of variances and Σ is the corresponding correlation matrix. Hence, $\text{var}(Y) = 3V^{-1} \Sigma^{-1} V^{-1} \Sigma^{-1} \Sigma^{-1} = 3V^{-1} \Sigma^{-1} \Sigma^{-1} V^{-1}$, where $e_3 = V^{-1} \Sigma^{-1}$, and factor loadings e_3 from CAR models will have the same scale as those derived from marginally specified models.

4. Analysis of Rhode Island census data

4.1 Objectives

The main objective of our analysis is to use of Bayesian factor analysis for computing and summarizing material deprivation at the census tract level. This involves fitting several models to the manifest variables, selecting a model that best represents the observed data, and then computing the posterior conditional distribution $p(\beta; Y; \Sigma; \gamma; \delta; \theta; \phi)$ using the chosen model. Our summary of material deprivation is in terms of posterior ranks of individual β_i drawn from this distribution, and in terms of posterior probability of deprivation. Our suggestion is to define “deprivation” in terms of percentiles of the latent index distribution, for example, the upper 20% of census tracts. Model-based summaries are compared in detail to Townsend indices.

4.2 Manifest variables from Rhode Island census data

Each census-level variable is the square root of a percentage, and these variables are labeled as follows. For census tract (area) i , Y_{i1} corresponds to the (square root) of percentage of “crowded” households (more than one person per room), Y_{i2} denotes percent unemployed, Y_{i3} is percent of individuals without a car, and Y_{i4} is percentage of households not owned by the occupant(s). Figure 1 shows, for the transformed variables, histograms in the left panels and correlograms in the right panels. For census variable j , the correlograms estimate the function $\frac{1}{2} \int_{-h}^h C_j(h) dh / C_j(0)$, where $C_j(h) = \text{cov}(Y_{ij}, Y_{i+j})$ is the average between area correlation as a function of Euclidean distance between area centroids. The correlogram indicates substantial correlation for areas close together. On the map that we used, maximum Euclidean distance between centroids of two areas is .801, and maximum distance between centroids of two adjacent areas is .180. The spatial distribution of each standardized variable is shown in Figure 2. From Figures 1 and 2, spatial clustering appears evident for all variables.

4.3 Model selection and posterior parameter distributions

Each of the models described in Section 3 was fit to the manifest variables; details on prior selection and posterior sampling are given in the Appendix. Model selection was made using a posterior predictive criterion that balances goodness of fit and predictive variance under a squared error loss function (Gelfand and Ghosh 1998; see also Spiegelhalter, Best, Carlin, and van der Linde 2002 for a criterion based on deviance). Define, for model q , $\hat{Y}_{ij} = E_q(Y_{ij} | \beta_j)$ and $\hat{\sigma}_{ij}^2 = \text{var}_q(Y_{ij} | \beta_j)$, where Y_{ij} is an observation from the posterior predictive distribution (i.e., a future observation that has the same distribution as the observed Y_{ij}). The model choice criterion for model q is $C_q = \frac{1}{n} \sum_{i,j} \frac{(Y_{ij} - \hat{Y}_{ij})^2}{\hat{\sigma}_{ij}^2} + k \log \frac{1}{\hat{P}_q}$; where

$$G_q = \frac{1}{n} \sum_{i,j} (Y_{ij} - \hat{Y}_{ij})^2 \quad \text{and} \quad (5)$$

$$P_q = \frac{1}{n} \sum_{i,j} \hat{\sigma}_{ij}^2$$

This criterion penalizes both for lack of fit and for high posterior predictive variance due to underparameterization or overparameterization. The constant k can be used to calibrate the role

of P ; that is, when $k \rightarrow 0$, model choice is based purely on t , and when $k \rightarrow 1$, t and variance are weighted equally. Table 1 summarizes C_{ij} for the models listed in Section 3. For many models, incorporating spatial variation actually leads to poorer model performance relative to the spatial independence model, but in nearly all cases it reduces predictive variance by borrowing information from neighboring areas. The best model with respect to fit and variability is 3B, a CAR formulation with spatial correlation based in adjacency (this holds for several other choices of k).

The posterior distribution of β suggests strong spatial association on the latent variable scale (median .150, 95% posterior interval .145 to .154; this is very close to its upper bound). The clearest manifestation of incorporating spatial correlation is reduced variability in posterior parameter distributions. Table 2 summarizes the posterior distribution of factor loadings. Variability in the posterior distribution of each β_j is measured using the interquartile range (IQR); in each case the IQR is smaller for the model incorporating spatial variations. Reductions in posterior IQR range from 14% for β_4 (housing not owned) to 23% for β_2 (unemployment). The attenuation of factor loadings in model 3B occurs because each factor loading is a variance Component - β_j $\text{cov.} Y_{ij} ; \pm i /$ - and some of the total variation in each Y_{ij} is absorbed by the spatial parameter.

Because residual variances σ_j^2 differ across j , the factor loadings themselves measure covariance on different scales and cannot be directly compared to assess strength of association between census variables and material deprivation. Instead we can examine squared correlation coefficients $\beta_j^2 = \beta_j^2 C_{jj} / \sigma_j^2$, for $j \in \{1; 2; 3; 4\}$; which represent proportion of variation in the j th manifest variable Y_{ij} that is explained by the latent index β_i . Table 3 gives a comparison of these between models 1 and 3B. Under both spatial independence and CAR dependence, the correlation is greatest between lack of car ownership and material deprivation, followed by nonownership of housing, unemployment, and crowding. The contribution of the last two variables is essentially equal.

4.4 Summarizing distribution of material deprivation

Material deprivation across areas is summarized in two ways. First, for each area, we computed the posterior distribution of the rank of $E_{ij} = Y_{ij} / \sigma_j$, including its mean and 99% probability interval. Using output from the Markov chain Monte Carlo algorithm, we first drew one replicate β_j from its joint posterior distribution, then drew the vector d_j according to (3), and then assigned ranks to the N elements of d_j . The simulated posterior rank distribution was determined from 1,000 replicates of d_j (see also Laird and Louis 1989). Figure 3 summarizes area-specific posterior ranks, indicating wide variability in precision of measuring deprivation, and makes clear that fine distinctions in deprivation level are unlikely to be possible.

Second, to simplify the reporting of deprivation, it is possible to derive ordered categories that represent degree of deprivation on a reduced scale, for instance, based on quintiles of the distribution of \hat{y}_i . The vertical lines on Figure 3 indicate quintile-based categories. The variation exhibited by the rank distribution makes clear the potential difficulties of classifying all areas because many 99% intervals straddle two categories (and in two cases, three categories). An alternative to strict classification is a summary of posterior probability of class membership, that is, the posterior probability that the rank for area i falls between a specific pair of quintiles. In cases where resource allocation (for example) is based on identifying only the 10% or 20% most deprived areas, posterior probability summaries can be used to avoid arbitrary cutoffs and artificial distinctions between areas that may be separated by a small number of ranks. If those above the highest quintile are to be identified as “materially deprived,” then any area with positive posterior probability of being ranked above that quintile would be included. Figure 4 illustrates this method of reporting using maps of the posterior probability of inclusion in quintile-based groupings 1, 3, and 5 (5D greatest deprivation). Darker shading indicates higher probability.

4.5 Model-based versus Townsend indices

Limitations of the Townsend index, spelled out in Section 1, include (a) that each census variable contributes with equal weight to the summary index; (b) census tracts are considered independent, despite similarities between neighboring tracts, possibly resulting in less efficient use of information; and (c) measures of uncertainty, particularly due to varying census tract population, are not part of the index. The model-based method of summarizing information on deprivation addresses each of these shortcomings while retaining the structural format of the Townsend index as a (weighted) linear combination of census variables. When spatial correlation is incorporated, the index for area i may depend on census variables recorded in other areas.

A direct comparison of area-specific indices derived using the Townsend formula and using model 3B is shown in Figure 3(b), which plots ranks of the Townsend index against posterior rank intervals derived from the model. As expected there is generally positive correlation between the ranks, but some notable discordancies are evident. Table 4 presents data from the five most discordant areas. Apparently, discordancies are due in large measure to differential weighting of census variables. In areas 61 and 72, for example, discordancy can be attributed to a higher-than-average percent with no car ($Z_3 > 0$), leading to model-based deprivation indices that are much greater than the Townsend index. (Recall that percent with no car contributes the greatest amount of information to the deprivation index.) The distribution of ranks is much less variable in the upper right tail, where in general there is more agreement between the Townsend and model-based indices.

5. Discussion

We have used a fully Bayesian factor analysis model for spatially correlated data. Spatial correlation is introduced on the latent variable (factor) distribution, and many common parameterizations of spatial covariation can be used, including direct and conditional (CAR) specifications. The model is applied to the problem of calculating an index of material deprivation from census variables, using tract-level data in Rhode Island from the 1990 U.S. Census.

Measuring area-level material deprivation is important in both policy and research. Standard indices, usually constructed as unweighted sums of census variables, are used for health services resource allocation and as proxies of socioeconomic position in epidemiologic research, and more recently have been proposed as public health monitoring tools for small areas (Krieger et al. 2002). The potential use of material deprivation indices as area-level health status surrogates holds great interest because the indices can be constructed from publicly available census data; however, proper validation requires the use of health outcomes data, such as from a non-public cancer registry.

Although any measure of a latent construct like deprivation is bound to be imperfect, our model-based approach provides a method for maximizing available information about a fixed area. The model-based index given here is a weighted average of census variables from the target area and from surrounding areas. Spatial correlation allows manifest variables from surrounding areas to inform the index. An advantage of allowing different weights for each variable is that the similar information from different manifest variables does not contribute more than once to the index. The Bayesian formulation leads naturally to indices that are reported with measures of uncertainty, such as posterior probability intervals associated with deprivation rank. In principle, there is no reason why deprivation cannot be summarized on the scale of the $\pm i$ themselves. Although the use of ranks does not convey scaling, it must be noted that a priori scaling of $\pm i$ in terms of the standard normal distribution is itself somewhat arbitrary.

An important question arises as to how researchers and policy makers should use this technology. A major consideration in both policy and research is whether model-based indices lead to increases in validity, whose definition depends on the intended use of the index. Toward this end, the model can be used to generate a public-use database consisting of census variables and appropriate summaries of the posterior index distribution. In the simplest case, the database could contain a posterior mean (i.e., scalar R value) for each area, $\mu_i = E(\pm i | Y) = \sum_j w_j Y_j / \sum_j w_j$. Alternately (or additionally), the posterior mean rank could be listed. Using standard analytic tools, researchers could then compare the validity of standard indices, such as the Townsend index, to the model-based indices. Another way to make the data available is in terms of replicate datasets, in the style of multiply imputed versions of census data (cf. Little

and Rubin 1987), to convey uncertainty in the deprivation measure. A database of this type would list for each area (say) 10 draws of E_i from its posterior distribution. An immediate application of these data would be to regression analyses where the index is used as a proxy for socioeconomic position; inferences would be drawn by combining parameter estimates derived from each imputed dataset using formulas given by Little and Rubin (1987).

Our work can be extended in several directions. First, factor analysis can be used as an exploratory tool to discover multivariate structure in deprivation, or more generally in socioeconomic position (Kreiger et al. 2001), by including more census variables. Thus one natural extension would permit multivariate factors with a spatial correlation structure. Second, as suggested earlier, the advantage gained by using a model versus an empirical index could be quantified through validation studies, such as the one carried out by Barnett et al. (2002). The analyses by Barnett et al. (2002) also suggest that the components of material deprivation may differ between urban and rural areas, suggesting the need to incorporate covariates, which could enter either at the first or the second level (Sammel and Ryan 1996). To validate measurement of deprivation per se, one can use external indicators measured at the neighborhood level (as in Raudenbush and Sampson 1999); to validate the index as a surrogate of morbidity or health services use, data on the specific health outcome is needed (e.g., Krieger et al. 2002). In principle, our model also could be elaborated to include temporal trends when data from different time points are available (e.g., Waller, Carlin, Xia, and Gelfand 1997), for example, to study the relationship between changes in deprivation and corresponding changes in morbidity or health services use.

B.5 P_CEO

Managerial pay and governance in American nonprofits

Kevin F. Hallock

Although there has been substantial work on executive compensation in the for-profit world (Murphy 1999, very little academic work has focused on the non-profit sector [see Oster (1998) for a recent exception]. This article examines several areas in the compensation of managers of nonprofits, including pay levels, pay dispersion, pay for performance, the effects of government grants, and governance, using detailed panel data from the Internal Revenue Service (IRS) on more than 30,000 manager-years from 1992 to 1996. It is particularly difficult to understand incentives for managers in situations like nonprofits where performance is unusually difficult to measure or quantify (Weisbrod 1989). This study provides an empirical analysis of several of the issues.

For a variety of reasons nonprofits provide an interesting context in which to study the pay of managers during this period. First, nonprofits are an important part of the economy. There are more than 1 million nonprofits in the United States. They employ more than 10 percent of all workers and account for about 6 percent of gross national product (GNP) (Bowen et al. 1994). Second, although there has been increased scrutiny from the IRS over the pay of managers of firms, recent legislation has focused on the pay of top managers of charities in the United States (Taxpayer Bill of Rights 2, 1996). This new law not only requires organizations to carefully document how much they pay top managers, but it also requires the boards of the charities to be able to document how the salaries of their chiefs are determined. If the salaries are found to be higher than expected and higher than those found in similar charities, the heads could be fined and required to return the amount by which they were overpaid. Moreover, the members of the boards of directors of these charities also could be fined if pay levels of the top managers are found to be in question. The problem with this legislation is that there are no strong guidelines for applying these laws. The third reason nonprofits are an interesting context in which to study compensation is that there is a theoretical literature that deals, in part, with how managers in nonprofits may be paid, but there is almost no empirical work in the area.

The data needed to study managerial pay in nonprofits has become available only recently. This article uses panel data on nonprofits and their top managers from a large sample of IRS tax returns of thousands of nonprofits for each year from 1992 to 1996 (inclusive). I also use a sample of top executives of firms collected from Standard & Poor's EXECUCOMP for comparison.

Clearly, nonprofits are different from for-profit firms (see Hansmann 1980, 1996; Mirvis and Hackett 1983; Steinberg 1990a, 1990b). However, for at least two reasons it seems natural to study the compensation of those at the top of nonprofits in the same

way as the compensation of senior executives of for-profit firms. First, both types of top managers are quite visible and clearly focal people within their organizations. Second, recent scrutiny of nonprofits from the IRS and others is similar to that faced by large firms in the United States. We know a great deal about the pay of executives in large firms by very little about compensation within nonprofits.

Researchers have studied compensation for well-educated people in the non-profit sector but have not concentrated primarily on top managers. For example, Weisbrod (1983) suggested that lawyers who work in the “public interest” sector have different preferences from those who work in private practices, and Preston (1989) found that the non-profit/for-profit wage differential was negative for many groups of workers by statistically more negative for managers and professionals. She suggested that the managers and professionals were more “closely tied to social benefit provision” in the organizations and therefore were more likely to “donate” wages to the nonprofits by accepting lower wages for similar work. Both these articles study differences in pay levels across the for-profit and non-profit sectors generally. Neither studies managers or directors, nor do they consider such a volume of data across such a wide variety of organizations.

In this article I use panel data from IRS returns on 32,146 organization-years to examine several issues related to top executives in non-profit organizations. The first section briefly describes the non-profit sector and some reasons we might expect pay to be different in nonprofits than in for-profit firms. The second section introduces the unique data and outlines an industry classification system known as the National Taxonomy of Exempt Entities (NTEE), which is similar to the Standard Industrial Classification (SIC) system used for firms. The second section also includes basic summary statistics on the pay of managers and directors by NTEE. We know from studying large for-profit firms that there is a well-documented link between chief executive officer (CEO) pay and organization size, but evidence for a link between other for-profit performance measures and CEO pay is mixed. The third section, therefore, is a discussion of measures of performance (including organization size) and whether there is a link between managerial pay and performance in nonprofits. The fourth section studies the connection between government grants and managerial pay in nonprofits and examines whether managerial pay increases in years when nonprofits get larger grants from the government. The fifth section examines governance and whether the size of the board of directors affects managerial pay in nonprofits. Finally, the last section offers concluding comments.

I find that managerial pay in nonprofits is substantial and varies within and across organizations. I also document certain possible measures of performance of the managers of nonprofits; however, evidence linking pay to performance is only robust when performance is measured as some function of organization size. Also, although in the cross section it appears that nonprofits with larger government grants pay their

managers more, this is not true within organizations. Finally, I find that the larger the number of paid board members, the lower is the pay of the top officers and the other staff.

Structure and uniqueness of the nonprofit sector

Distinct Institutional Features.

This section places 501c(3) (charitable) nonprofits into the larger context of all nonprofit organizations. These institutional features are discussed in order to provide a background for the discussion of managerial pay in nonprofits that follows.

In order to become officially designated as a nonprofit, an organization must file forms with the IRS. Such organizations do not have to pay taxes, although if they have greater than \$25,000 in annual net revenue, they must file IRS Form 990 (described in more detail in the data section below). Among the 28 possible groupings for nonprofits, by far the most common designation is 501c(3), "Charitable and Religious". 501(3) organizations are considered charitable because, according to the IRS, they serve "broad public purposes include[ing] educational, religious, scientific, and literary activities, among others, as well as the relief of poverty and other public benefit actions" (Stevenson, Pollak, and Lampkin 1997). Table 1 shows that of the roughly 1 million nonprofit organizations in the United States in 1992 to 1996, approximately half were 501(3) organizations. 501(3) nonprofits have the added benefit that contributions to the organization are deductible to the contributor.

There are at least two key features of nonprofits that distinguish them from for-profit organizations. First, nonprofits inherently have a different bottom line in that they are not created to generate returns to their shareholders. Second is Hansmann's "nondistribution constraint." Hansmann (1980, 1996) notes that while nonprofit organizations are free to make profits, those profits cannot be distributed to those with formal control over the organization: "A nonprofit organization is, in essence, an organization that is barred from distributing its net earnings, if any, to individuals who exercise control over it, such as members, officers, directors, or trustees" (Hansmann 1980). However, this does not imply that nonprofit organizations cannot make profits in a technical sense; rather, "It is only the distribution of profits that is prohibited" (Hansmann 1980).

It is also worthwhile considering why some organizations might form as nonprofits. Hansmann outlines several reasons. The first is that the buyer and the recipient of goods and services are not always the same person. For example, people are more likely to donate funds to a relief organization organized as nonprofit because they know that the managers are less likely to abscond with the residual money. Another reason that an organization may form as a nonprofit comes out of the idea of public goods (Hansmann 1980). The fact that one person enjoys the services of a public good does not preclude others from doing the same. For example, donors are more likely to

contribute to nonprofit public radio stations because they are more likely to believe that the funds will be used for programming rather than for the private benefit of the managers. Museums and operas are similar in that they practice “voluntary price discrimination” (Hansmann 1980) by offering low-priced seats, but those who can afford to contribute do so with the knowledge that the residual money cannot be distributed to those in control. Hansmann (1980) also describes “implicit loans” in higher education, where students are more likely to donate money back to their nonprofit colleges when they feel that those in charge cannot keep the money for themselves.

Another interesting and important example of an institution that optimally may form as a nonprofit is a nursing home or day-care center (Weisbrod and Schlesinger 1986). Parents (in the case of a day-care center) and children (in the case of a nursing home) are more likely to place their loved ones in the care of institutions organized as nonprofits because they believe that the additional dollar is more likely spent on quality care, food, etc., and not on the well-being of the manager.

Reasons why pay levels for managers of nonprofits may be different.

The first possible reason for differences in pay between employees of nonprofit organizations and for-profit firms was presented by Preston (1989), and some of her findings may be applied to managers as I do here. The idea has to do with “labor donations” and the fact that workers may trade lower pay for higher social benefits. The limiting case of labor donations is volunteer labor, where workers are paid nothing for the time they donate to an organization. If labor donations are at work, it is expected that wages for similar workers will be lower in nonprofits than in for-profit firms.

Another possible reason for differences in pay comes from the screening hypothesis of Hansmann (1980) and is related to the choice of organizational form described earlier. Hansmann (1980) examined a simple model in which there are two types of managers. He suggested that if consumers are not good judges of the quality of a service provided by an organization, then some managers, by virtue of accepting lower wages, could make productive signals to the uninformed consumers, and “the nonprofit form may both restrain the managers of the organization, whatever their personal desires, from profiteering at the expense of the organization’s patrons and may select as managers precisely that class of individuals whose preferences are most in consonance with the fiduciary role that the organization is destined to serve.” In other words, heads of nonprofits essentially take lower wages to signal that donations will be used effectively.

Workers in nonprofits may accept lower wages in exchange for a host of pleasant amenities on their job, such as flexible hours, more stable job prospects, and a slower pace of work (compensating differentials). An additional reason for lower pay in

nonprofits may be the differences in returns to characteristics of managers and ability bias. Controlling for ability across sectors is beyond the scope of this article, although it is empirically testable if one had data on those individuals who switch management jobs between sectors.

Data

Primary sample

The data I use in this article are from tax returns of 501(c)(3) tax-exempt organizations. The primary sample includes organizations filing Form 990 returns each year from 1992 to 1996 (inclusive). The original sample consists of 35,109 organization-years (10,539 unique organizations). I delete from the sample those organizations which have assets, expenses, revenue, program service expenses, grants, and other expenses or top officer, director, or trustee compensation that is missing or less than \$1000. I also dropped organizations that have negative values for management and general expenses, fund-raising expenses, or payments to affiliates. The financial criteria reduce the sample by 2175 observations (6.9 percent), and the compensation criteria reduce the sample by 788 observations (2.24 percent). These selection criteria reduce the sample to 32,146 organization-year observations over the 5 years (from 9776 unique organizations). Due to the selection criteria, the results are more generalizable to larger nonprofits. These data come from the Statistics of Income (SOI) files of the IRS. See Stevenson, Pollak, and Lampkin (1997) for additional details on the data.

The IRS data contain a host of interesting information on the compensation of all officers and directors and of the top five nonofficer/director/trustee employees. For each officer, director, trustee, or other “key employee,” I have collected three measures of compensation: (1) base compensation, including “salary, fees, bonuses, and severance payments paid” (Internal Revenue Service 1996), (2) contributions to employee benefit plans and deferred compensation, including “medical dental and life insurance” (Internal Revenue Service 1996), and (3) expense account and other allowances, including “expense allowances or reimbursements that recipients must report as income on their separate income tax returns”. Examples include amounts for which the recipient did not account to the organization or allowances that “were more than the payee spent on serving the organization.” Organizations are required to include such payments as “the value of the personal use of housing, automobiles, or other assets owned or leased by the organization” (Internal Revenue Service 1996). Means and standard errors (in parentheses) for the sum of these three variables are reported in Table 2. (Breakdowns by industry and type of pay are reported in Table 3). The first officer or director earned, on average, \$160,098 (all financial data are reported in real 1996 dollars). A significant part of the analysis that follows will focus on the total compensation, which includes the sum of cash, benefits, and expense accounts of the top officer, director, or trustee and of the top nonofficer, nondirector, or nontrustee employee.

The average total compensation packages for the second and third highest paid officers and directors were \$122,381 and \$111,530, which are both substantially less than the average of \$160,098 reported for the top officer. The increase in average pay from position 3 to position 2 is 9.7 percent, and the increase in pay from position 2 to position 1 is 30.8 percent. This result (ever-increasing pay as one moves up the organizational hierarchy) is consistent with tournament theory (see Lazear and Rosen 1981). The top panel of Table 2 also reports mean total compensation for the three highest paid nonofficer, nondirector, or nontrustee employees. These are relatively lower, ranging from \$112,659 for the highest paid nonofficer, nondirector, or nontrustee employee to \$83,127 for the third highest paid nonofficer nondirector or nontrustee employee.

These numbers can be compared with the average compensation of asset of 6677 CEO-years collected from EXECUCOMP firms in the Standard & Poor's (S & P) 5000, S & P Midcap 400, and S & P SmallCap 600 (1836 unique CEOs over 5 years from 1992 to 1996). The average salary and bonus for this sample of CEOs is \$919,942. The average compensation including salary, bonus, other annual compensation; total value of restricted stock granted; total value of stock options granted; long-term incentive payouts; and other is \$2,183,024. Clearly, on average, CEOs of large firms earn more than heads of nonprofits. It is important to note, however, that these firms are much larger than the sample of nonprofits (this will be explored further later in the article).

The data also include financial information such as total assets, total expenses, and revenue. The expenses are further broken down into several categories, including (1) program services expenses, which are "mainly those activities that the reporting organization was created to conduct and which, along with many activities commenced subsequently, form the basis of the organization's current exemption from tax," (2) fund-raising expenses, which are "total expenses incurred in soliciting contributions, gifts, grants, etc.," (3) management and general expenses, which are the "organization's expenses for overall function and management, rather than for its direct conduct of fund-raising activities or program services," including "legal services, accounting, insurance, office management, personnel," and other expenses (Internal Revenue Service 1996), and (4) payments to affiliates. These data are reported in the second panel of Table 2. The assets of these organizations are substantial, averaging \$62 million. The average revenue is \$60 million, and the average expenses are \$55 million. A substantial fraction of the expenses are spent on program services (\$48 million) and a relatively small fraction on fund raising.

I also have collected data on sources of support for the nonprofits under study. In particular, I have included information on three measures of outside support for the organizations as defined by the IRS: (1) public direct support, (2) public indirect support, and (3) government grants. Public direct support is the sum of "contributions, gifts, grants, and bequests that the organization received directly from the public"

(Internal Revenue Service 1996). The average organization in the sample received \$3.9 million in public direct support each year during the sample period. Indirect public support includes “total contributions received indirectly from the public through solicitation campaigns conducted by federated fund-raising agencies and similar fund-raising organizations (such as the United Way organization and certain sectarian federations)” (Internal Revenue Service 1996). The average public direct support each year during the sample period was \$653,911. Under the law, a government grant is “treated as a contribution if its primary purpose is to enable the donee to provide a service to, or maintain a facility for, the direct benefit of the public rather than to serve the direct and immediate needs of the grantor even if the public pays part for the expense of providing the service or facility” (Internal Revenue Service 1996). The average charity received roughly \$3.2 million in government grants in a typical year during the sample. The last panel of Table 2 reports the average number of paid officers and directors (4) and the average number of uncompensated officers, directors, or trustees (20).

Industry classifications

There is evidence in the literature on executive compensation in firms that executive pay levels vary by industry. For example, Joskow, Rose, and Shepard (1993) show that CEOs in regulated industries earn less than those in unregulated industries. Until recently, there has been no system such as SIC codes for firms to classify the substantially different missions of nonprofits even among those classified as 501c(3) organizations. However, the NTEE has been developed to help classify charitable nonprofits into specific groups.

The four-character NTEE codes are similar to SIC codes for firms. There are 26 “major groups” (listed in the left-hand column of Table 3), such as Arts, Culture, and Humanities and Educational Institutions and Related Activities, which comprise the first character of the codes. These 26 major groups can be generalized into 10 “major categories”: (1) Arts, Culture, and Humanities, (2) Education, (3) Environment and Animals, (4) Health, (5) Human Services, (6) International, Foreign Affairs, (7) Public, Societal Benefit, (8) religion related, (9) Mutual/Membership Benefit, and (10) Unknown. Under each of these 26 groups are two digits of subcodes that make finer distinctions in organization type. The last of the four characters, the “common code”, includes classifications such as Alliance Organizations, Management and Technical Assistance Services, and Public Education. There are a total of 645 unique subgroups in the NTEE. See Stevenson, Pollak, and Lampkin (1997): Appendix B), Hodgkinson (1990), Hodgkinson and Toppe (1991), Gronbjerg (1994), and Turner, Nygren, and Bowen (1992) for more details on the NTEE. In several parts of the remainder of this article I will control for industry effects by concentrating on the 26 major groups.

Table 3 reports some simple results of differences in compensation by major nonprofit industry type for the top officer, director, or trustee and for the top nonofficer,

nondirector, or nontrustee employee. This table also separates total compensation into its three components: base pay, benefits package, and expense account. The pattern that the top officer or director earns more than the top nonofficer, nondirector, or nontrustee as reported in Table 2 is still evident in Table 3 (in fact, it is true in each group). It is also clear that there is substantial variation in the levels of compensation within the nonprofit sector (across NTEE classifications). For example, the average base pay for the top officer/director/trustee in the medical research sector (category H) was \$205,025, and the average in health, general and rehabilitative (category E) was \$187,379, whereas the average base pay for the top officer/director/trustee in the category of religion-related (category X) was \$72,315, and the average in Housing, shelter (category L) was \$72,092. There is also substantial variability in the value of the benefits package offered to the top officers/directors/trustees from a high of \$27,383 (Public, Society Benefit: Multipurpose and Other, category W) to a low of \$4,177 (Housing, shelter). In addition, there is also significant variability in the expense account category, which ranged from a high of \$6,581 (Medical Research, category H) to a low of \$946 (Housing, shelter). The ranges for the top non-officer/director/trustees employees are similarly striking. Although there has been only little work on specific industries within nonprofits [e.g., Ehrenberg, Cheslock, and Epifantseva (2000) for universities and Bertrand, Hallock, and Arnould (2000) for hospitals], continued work might help us to understand reasons for differences in pay across nonprofit industries.

Pay for Performance in Nonprofits

The relationship between performance of firms and the compensation of CEOs has been widely studied in for-profit firms, but little attention has been paid to nonprofits. In a few articles, however, Steinberg (1990a, 1990b) and Weisbrod (1989) examine incentive compensation for employees in nonprofits mostly from a theoretical point of view. We know that designing incentive compensation plans in for-profit firms is difficult (Lazear 1995), but it may be even more difficult to measure the performance of managers in nonprofits because these organizations are likely to be striving to create something much different from returns to shareholders. One feature that makes nonprofits distinct from for-profit organizations is the nondistribution constraint (described earlier). This nondistribution constraint does not imply, however, that employees cannot be paid based on incentives (see Abelson 1998; Steinberg 1990b).

A classic example of how difficult it is to measure performance for managers in the nonprofit sector is the case of a manager of a nursing home (Weisbrod and Schlesinger (1986). A manager could be paid based on the profits he or she accrues, but this gives him or her incentives to provide lower-quality care to the residents. The output sought by the board of the nursing home (say trustworthiness) is difficult to observe. We could try to measure trustworthiness by using an easy-to-observe measure such as the

mortality rate. However, this might induce management to admit residents who are mostly healthy, which is, no doubt, at odds with the mission of the board.

The literature on CEO pay and firm performance provides several common measures to test the top manager's performance, among which are stock returns and the change in market value or size of the firm. Since nonprofits are not owned by shareholders, there is no stock price or value of the nonprofits in the formal sense. The rest of this section is devoted to several possible measures of performance for managers in nonprofits.

Charity size and managerial pay

A host of articles, including Murphy (1985), Finkelstein and Hambrick (1989), Kostiuk (1990), Lambert, Larker, and Weigelt (1991), and Rosen (1992), have documented a link between the size or scale of a firm and the pay of the CEO. Firm size has been measured in the executive pay literature by such variables as market value, assets, and number of employees. One reason why we would expect managers of nonprofits with more assets to earn more is that they have responsibility over a much larger number of people and resources. Larger organizations are more difficult to manage and therefore may require top executives with scarce (and expensive) human capital.

Table 4, panel A, reports the results of an empirical examination of this issue using the following very simple empirical specification:

Where C is the log of manager compensation, S is the log net ending assets of the organization, i represents organizations, t represents time in years, and \dots is a composite error term containing possible permanent effects.

The first column of Table 4, panel A, reports results from an ordinary least squares (OLS) regression of $\ln(\text{total compensation})$ for the top officer, director, or trustee on $\ln(\text{assets})$, a set of yearly indicator variables, and a constant for the nonprofits. Log assets is a naturally important measure of organization size of nonprofits (Salamon 1992). The coefficient on $\ln(\text{assets})$ is 0.218 ($t = 71$). This return number is large and significantly positive, as expected, and is somewhat smaller than for CEOs of firms. This large effect, taken at face value, suggests that larger firms pay their top managers more. We know, though, that the compensation of managers varies by industry (recall Table 3), so in column 2 of Table 4, panel A, I also control for the industry in which the nonprofit operates. Even controlling for industry effects, firm size [measured as $\ln(\text{assets})$] is still a substantial and important contributor to top manager pay in this specification.

I performed a brief but similar analysis of the compensation of managers in firms for comparability using data on 6677 firm-years from 1992 to 1996 from the S & P EXECUCOMP data described earlier. The dependent variable is the log of total compensation. The main independent variable $\ln(\text{assets})$ (I also control for yearly time indicators) is also collected from EXECUCOMP. The regressions in Table 5, panel B,

show that the return to assets is much higher in for-profit firms than in nonprofits. My results are consistent with those of other studies on CEO pay in firms. Using other measures of CEO pay for firms (e.g., just considering salary and bonus alone) as dependent variables yields similar results. Just as there is a well-documented link between the size to scale of a firm and CEO pay, there is a strong and positive link between the assets of a charity and the pay of its top manager, although it is somewhat weaker in nonprofits.

Clearly, however, firm size and industry effects do not fully explain top manager pay in this sample of nonprofits (R^2 in column 2 of Table 4, panel A, is only 0.172). We need, therefore, to control for the possibility that other characteristics of managers and organizations are confounding the relationship between firm size and managerial pay in nonprofits. To help remedy this situation, I make use of the 6 years of IRS panel data. If one is willing to assume that the source of endogeneity arises only through the permanent component of the error term ξ and not through the transitory component ϵ , then the standard fixed-effects estimate of the preceding equation will yield consistent estimates of the parameters. These results are displayed in the third column of Table 4, panel A, and show that, even controlling for organization fixed effects, the relationship between organization size and top manager pay remains. In other words, even within organizations, as organizations grow, managers are paid more. This same set of relationships holds for top nonofficer, nondirector, or nontrustee employees in the right three columns of Table 4, pane A, although the coefficient estimates are smaller in each case.

As mentioned earlier, one problem with studying nonprofit organizations is their myriad of different potential missions. With for-profit firms, the goal or true measure of performance is increasing shareholder value, which many authors (e.g. Jensen and Murphy 1990) measure with firm market value (firm size). Table 5 relaxes the idea that the relationship between firm size and managerial pay should be the same across all different nonprofit industries while retaining the idea that net ending assets are an appropriate measure of size for each organization. This table represents the results for a set of regressions where the log of total compensation of the top officer, director, or trustee is regressed on log net ending assets and a set of annual time indicators by NTEE industry. It is clear that within each industry the relationship between organizational size and managerial pay is positive and significant. When individual organization fixed effects are controlled for (not reported in the tables), most industries have a relationship between firm size and managerial pay that is not statistically different from zero. This may reflect the fact that the samples are relatively small within industries.

Clearly, organizational size is strongly linked with top manager pay. The effect is tempered substantially when organization fixed effects are controlled for. Oster (1998) compares the relationship between managerial pay in nonprofits and organization size

in various ways across selected industries such as hospitals, social service organizations, and foundations. In five separate samples of between 31 and 95 observations, she finds coefficients on measures of size such as $\log(\text{assets})$ in a managerial pay regression in the range of 0.67 to 0.376. Oster's (1998) work is very informative but based on limited samples and is only cross-sectional. She therefore cannot consider changes within organizations using organization-specific effects.

Other measures of performance

Although it is clear that firm size is strongly related to managerial pay in nonprofits (even within organizations), it is useful to investigate whether there might be other measures. Another possible choice is the fraction of expenses spent on program services, a measure on which some groups of nonprofits seem to concentrate. Each column of table 6 represents the results from a different OLS regression. Columns 1 through 3 report results when the dependent variable is $\log(\text{total compensation for the top officer, director, or trustee})$. Columns 4 through 6 give the results when the dependent variable is $\log(\text{total compensation of the top nonofficer, nondirector or nontrustee})$. The regressions also control for $\ln(\text{assets})$ and yearly time indicators. In the simple cross section for the top officer, director, or trustee, it is clear that higher levels of expenses going to program services are associated with higher pay (the coefficient estimate is 0.098). This result does not hold when we control for NTEE effects in column 2 or when we control for organization fixed effects in column 3.

Columns 4, 5, and 6 examine the compensation of the top nonofficer, nondirector, or nontrustee employee, and the results differ somewhat. In the simple cross section, the link between performance and pay is much stronger (0.591), and the result is still significant when 26 NTEE (industry) classification (26 possibilities) while still accounting for individual organization fixed effects (I have not reported the results in the tables). Only two of the individual industries show significantly positive returns on the fraction of expenses spent on program services. There is only a very weak link between pay and performance based on this measure. Nonprofits that spend higher fractions of their total expenses on actual program services tend to pay their top officers, directors, and trustees and top nonofficer, nondirector, and nontrustee staff more in the cross section, but further analysis suggests that this result is not particularly robust.

While, clearly, success in fund raising may be an important metric by which to evaluate managers of nonprofits, another possible measure of performance could be the combination of government grants and direct and indirect support from the public. The results of this investigation are reported for the two types of managers in Table 7. In this case, even conditional on assets of the organization and with controls for NTEE industries or organization indicators, the higher the outside fund raising, the higher is the pay of the manager.

There is a host of other potential ways to measure performance in nonprofits. These include nonprofit profits, defined as total revenue minus expenses (median level in this sample \$1,117,222), and return on assets, defined as profits/assets (median level in this sample 7.01 percent). However, from an empirical point of view, neither of these measures is really related to compensation of the managers of the nonprofits. This may be due to the fact that there are wide differences in the missions of most nonprofits. In the end, the measures of nonprofit size (including net ending assets) seem most appropriate as measures of performance.

As an additional examination of the issue of performance in nonprofits, I analyze (but do not report in the tables) the issue of relative performance evaluation [see Antle and Smith (1986) and Gibbons and Murphy (1990) for a discussion of relative performance evaluation in for-profit firms]. In this analysis I estimated empirical specifications much like those in Table 6 with the additional covariate of average performance (program service expense/total expense) in that organization's industry (defined as one of the 26 NTEE industries detailed in the left hand column of table 3). I find no evidence of relative performance evaluation for top managers in nonprofits. While performance by some metric may matter (although the results are not particularly robust) for the pay of managers in nonprofits, there is no evidence that it is measured relative to others.

Government grants and the pay of managers in nonprofits

All 501c(3) nonprofits "are entitled to receive tax-deductible contributions and are defined by the IRS as 'charitable' because they serve broad policy purposes" (Stevenson, Pollak, and Lampkin 1997). Considerable work has been done on issues of government crowd-out with regard to donations. From a conceptual point of view, government grants may mean one of several different things for the pay of managers of nonprofits. First, they could be a sign of organizational quality or managerial ability. That is, in the cross section, we might expect that those organizations with grants or those with larger grants should have managers who are paid more. However, within organizations (where industry type, organization quality, and managerial ability are fixed), more government grants may be a sign of some sort of outside monitoring of the manager or an increase in funding attracted by the manager.

These ideas suggest the following simple modification of the earlier equation:

And this includes the addition of government grants G_{it} . I describe the precise definition of G_{it} below. Regressions of this form are reported in Table 8. The first three columns are for the compensation of the top officer, director, or trustee, and the last three are for the compensation of the top nonofficer, nondirector, or nontrustee. The top panel defines $G_{it} = 1$ if there is any government grant and 0 otherwise, and the bottom panel defines $G_{it} = \log(\text{grant size in dollars})$ for all organizations with positive grants. In column 1 of Table 8, panel A, the log of total compensation of the top manager is regressed on an indicator for whether a government grant exists, a

constant, a measure of organization size (log assets), and a set of yearly time indicators. The coefficient 0.165 suggest that organizations that have government grants pay their managers on average 17 percent more than those which do not have grants. Once individual organization fixed effects are controlled for in column 3 of Table 8, panel A, the effect of grants is slightly negative (but not significant). Roughly the same results come out of similar analyses in the rest of the panel.

This suggests that in the cross section, organizations that receive grants pay their managers more but that within organizations (i.e., once organizational quality or managerial ability are fixed) managers with grants may be paid the same or slightly less. That is, within an organization, adding a grant does not increase and actually may decrease managerial pay.

Managers of nonprofits are paid more on average when they generate grants, but this is not true within organizations. Clearly, certain kinds of organizations that pay managers will generate relatively large amounts of grant money. However, top managers in nonprofits are not additionally compensated when their organizations generate new (or additional) government grants.

Governance in nonprofits: the role of boards in managerial pay

There is some evidence that the makeup of the board of directors may have some influence on managerial pay or turnover in for-profit firms (e.g., Weisbach 1988; main, O'Reilly, and Wade 1995; and Hallock 1997, 1999). Despite some excellent work in the area (e.g. Bowen 1994), little is known about governance in nonprofit organizations, and even less is known about the compensation structure for the managers who lead these organizations. This section contains an investigation of the effects of board structure on the compensation of managers and top employees in nonprofits.

Hallock (1997) presents cross-sectional results that those that, controlling for the levels of assets, stock performance, and CEO characteristics, CEOs who lead for-profit firms with larger boards earn more. However, nonprofit organizations are different in many ways, and their board may be serving a different function (Bowen 1994). Perhaps larger boards suggest that there is more monitoring, and therefore, lower-quality managers are needed or at least managers who are paid less.

Table 9 presents the results from regressions of a similar form as in the preceding tables with the addition of a variable on the number of board directors in the nonprofit organization. The results for the top officer, director, or trustee are reported in the first three columns and for the top nonofficer, nondirector, or nontrustee employee in the last three. Panel A reports results on the link between managerial pay and the number of paid board members, whereas panel B reports results on the link between managerial pay and the number of unpaid directors.

Some researchers including Main, O'Reilly, and Wade (1993) have tested tournaments (Lazear and Rosen 1981) in large for-profit firms by including the number of vice

presidents in a standard CEO pay regression. The idea is that, conditional on the other factors, CEO pay should be higher in the presence of more vice presidents because more individuals are in competitions for the top spot in the firm. In the case of this article, however, board members are not as likely as vice presidents in for-profit firms to compete for the top position, so the tournament theory is not necessarily testable in this way in nonprofits.

Panel A of Table 9 shows that the larger the board, the less the top officer, director, or trustee is paid. This is true even when controlling for 26 NTEE classifications (column 2) and for individual organization fixed effects (column 3). For the top nonofficer, nondirector, or nontrustee employee, in the simple cross section and when controlling for 26 NTEE industry effects, the larger the board, the higher is the pay of the employee. However, when individual organization fixed effects are controlled for, the same negative result (larger board, less pay) appears. The results are less robust in panel B of Table 9 when the effect of unpaid directors is examined.

Again, the results in this section are not entirely clear. Perhaps board members serve as substitutes for management experience. Therefore, as board sizes grow, there is less need for a highly paid top officer, director, or trustee or nonofficer, nondirector, or nontrustee employee because larger paid boards may be able to more easily manage organizations than smaller boards. Another explanation may be that it is easier for larger boards to monitor the operations of organizations, and therefore, organizations with larger boards may not require managers with high salaries. The differing results for the unpaid/paid members suggests that a more detailed examination of the effects of boards on managerial pay in nonprofits would be interesting.

Concluding comments

There is an expansive theoretical and empirical body of work on the compensation of managers of firms in the United States (e.g., Murphy 1999) and some basic theoretical work on compensation in nonprofits that can be applied to managers of nonprofit organizations. However, little is known about the compensation of managers of nonprofits from an empirical point of view. This article presents evidence on the compensation of top managers of nonprofits in the United States using a very large amount of previously unexplored data from tax returns of the organizations.

Among the main empirical findings are that pay levels of top officers and directors of charities average about \$160,000, not close to those reported for heads of large U.S. firms. Just as there is a well-documented link between organization size and managerial pay in for-profit firms, there is a similar relationship in nonprofits. Also, there is substantial variability in the pay of top managers of charities, within the charities and across types of charities, as defined by their NTEE industries. In addition, a reasonably large fraction of the compensation of top officers, directors, and trustees and of top nonofficer, nondirector, or nontrustee employees comes in the form of

benefit plans and expense accounts. It is also the case that although in the simple cross section nonprofits with higher levels of government grants pay their heads more, this is not true once organization fixed effects are accounted for. This may be due to the fact that government agencies monitor managerial actions, and managers can be paid less in the presence of this monitoring. Similarly, there is a weak negative relationship between the number of paid board members and managerial pay in nonprofits, suggesting, perhaps, that paid boards of directors substitute for managerial talent.

Recently, there has been increased emphasis on the compensation and governance of managers in nonprofits (e.g., Taxpayer Bill of Rights 2 1996). However, very little is known about the pay of these managers. 501c(3) nonprofits seem to be a relatively narrow set of organizations. However, one possibility is that they actually perform significantly different tasks and therefore are organized quite differently from one another. This may lead, in turn, to differences in the way they pay their managers. This article is an important first step in studying the pay of managers in these types of organizations. Hopefully, these basic findings can help inform the issues and motivate further research in the area.

B.6 S_Rand

Random matching in the college admissions problem

1 Introduction

The study of two-sided matching has been mainly devoted to centralized markets.

These matching markets work by having each agent of the two sides of the market submit a rank ordered preference list of acceptable matches to a central clearinghouse, which then produces a matching by processing all the preference lists according to some algorithm. Typically, such mechanisms are deterministic in the sense that the outcome depends on the submitted lists in a way that involves no element of chance. As a consequence, the existing results do not generally allow us to address behavior in many labor markets and other two-sided matching situations where lotteries ultimately determine the outcome. In discrete problems where agents have opposite interests randomization is surely one of the most practical tools to achieve procedural fairness.¹ Hence, equity considerations provide an important justification for the introduction of chance in many instances of centralized matching. On the other hand, lotteries are especially attractive as a means of representing the frictions of a decentralized market. Indeed, in the extremely complex environment of a real life market, decentralized decision making will often lead to uncertain outcomes: the question of who will match with whom depends on the realization of random events—random meetings.

This paper studies a class of matching mechanisms that are random: given agents' behavior, chance determines the final outcome. These mechanisms may be used in centralized markets as a means to promote procedural fairness. Or they may arise in the context of decentralized decision making: starting from an arbitrary matching, agents from the two sides of the market meet bilaterally in a random fashion. We assume that each individual has preferences over the other side of the market and the prospect of being unmatched; however, they are not compelled to behave in a straightforward manner, according to these true preferences. Instead, agents are confronted with a game in which they act in what they perceive to be their own best interest. Hence, upon meeting, the paired agents match if this is consistent with their strategies, and separate otherwise. Since one of the clearest lessons from the study of deterministic procedures is that understanding such incentives is crucial to understand the behavior of the market, the paper is devoted to equilibrium analysis.

Our study was largely motivated by Roth and Vande Vate (1990, 1991). In the context of the marriage problem where matching is one-to-one, Roth and Vande Vate (1990) proved that, starting from an arbitrary matching, the decentralized decision making process of allowing randomly chosen blocking pairs to match will converge to a stable matching with probability one. Under a stable matching no individual or pair of agents has incentives to circumvent the matching. It is argued that such process can be

thought of as an approximation to real life dynamics. In the related paper Roth and Vande Vate (1991), strategic considerations are made for the marriage market, focusing on the class of truncation strategies, i.e., strategies that are order-consistent with true preferences, but may regard fewer partners as acceptable. In a one-period game in which every agent states a list of preferences and then a matching stable with respect to those preferences is selected at random, it is shown that all stable matchings can be reached as equilibria in truncations. However, certain unstable matchings can also arise in this way. A multi-period extension is then considered to rule out such undesirable outcomes.

As in Roth and Vande Vate (1991) we assume that random meeting among agents will eventually converge to a stable matching with respect to the chosen strategy profile. Hence, such process induces a lottery exclusively over stable outcomes. However, the present paper extends their contribution in two ways. First, we take equilibrium analysis further, going beyond the analysis of truncations. A concept of equilibrium based on first-order stochastic dominance is used, given that preferences are ordinal in nature and probability distributions over matchings are to be compared. The notion of ordinal Nash equilibrium guarantees that each agent plays his best response to the others' strategies for every utility representation of the preferences.² Second, the analysis is conducted in the context of the college admissions problem. In this setting, agents belonging to two disjoint sets (henceforth firms and workers) have preferences over the other side of the market; in addition, each firm can employ at most some fixed number of workers, while each worker can fill only one position. Strategic issues in this context have been studied for a deterministic stable matching rule. Roth (1985) shows that no stable matching rule exists that makes it a dominant strategy for all players to report their true preferences. Moreover, he proves that there are equilibrium misrepresentations that generate any individually rational matching with respect to the true preferences.³ Ma (2002) shows that in order to obtain stability with respect to true preferences, we have to use a refinement of the Nash equilibrium concept and restrict to truncations at the match point (i.e., strategies that preserve the ordering of the true preferences, but rank as unacceptable all the agents that are less preferred than the current match). More precisely, all strong equilibria in truncations at the match point produce stable outcomes. Further, Ma (2002) establishes that every Nash equilibrium profile admits at most one stable matching with respect to the true preferences; if, indeed, such a matching is admitted, it will always be achieved.

In this paper we characterize equilibria arising in the game induced by a random stable matching mechanism, providing simultaneously some results that extend to deterministic mechanisms. First, we show that when ordinal Nash equilibria are considered, a unique matching is obtained as the outcome of the random process. In addition, this outcome is individually rational with respect to the true preferences. Since every individually rational matching for the true preferences can be achieved as

an equilibrium outcome, we establish that a matching can be reached at an ordinal Nash equilibrium if and only if it is individually rational for the true preferences. We then turn our attention to equilibria where firms behave straightforwardly. In fact, there are reasons to contemplate truth telling as a salient form of behavior in situations involving uncertainty; further, sophisticated strategic play does not even make sense in settings where firms follow an objective criterion to fill their positions. We prove that, even though workers may not play straightforwardly, stability with respect to the true preferences holds for any matching that results from a play of equilibrium strategies in which firms reveal their true preferences.

Conversely, every matching that is stable for the true preferences can be achieved as an equilibrium outcome. In closing, we relate the equilibrium strategy profiles in the games induced by both random and deterministic mechanisms. In particular, for any random stable matching mechanism that always assigns positive probability to two different stable matchings (when they exist), we show that a strategy profile is an ordinal Nash equilibrium if and only if it has a unique stable matching and it is a Nash equilibrium in the game induced by some deterministic stable mechanism.

We proceed as follows. In Sect. 2 we present the college admissions problem and introduce notation. We describe random matching mechanisms and the equilibrium concept used in Sect. 3. In Sect. 4 we turn our attention to individual decision making. The matching process is modeled as a one-period game and its equilibria are then characterized. In Sect. 5 we briefly discuss equilibria in the context of a sequential game. Some concluding remarks follow in Sect. 6.

2 The model

The agents in the college admissions problem are two finite and disjoint sets, the set $W = \{w_1, \dots, w_p\}$ of workers and the set $F = \{f_1, \dots, f_n\}$ of firms. We let $V = W \cup F$ and sometimes refer to a generic agent by v , while w and f represent a generic worker and firm, respectively. Each worker w can work for at most one firm and each firm f has a quota q_f , the maximal number of workers it may employ. Each worker w has a complete, transitive, and strict preference relation P_w over the set $F \cup \{w\}$. For example, the preferences of w on $\{f_1, f_2, f_3, f_4\} \cup \{w\}$ can be represented by $P_w : f_1, f_2, w, f_3, f_4$, indicating that the best firm for w is f_1 , his second choice is f_2 , and he prefers being unemployed than working for either f_3 or f_4 . Each firm f also has a complete, transitive, and strict preference relation P_f over the set $W \cup \{f\}$. For example, the preferences of f on $\{w_1, w_2, w_3, w_4\} \cup \{f\}$ can be represented by $P_f : w_3, w_1, f, w_2, w_4$, indicating that the best worker for f is w_3 , its second choice is w_1 , and it prefers having a position unfilled to hiring any other worker. A worker is acceptable if the firm prefers to employ him rather than having a position unfilled. Formally, the set of acceptable workers for f is $A(P_f) = \{w \in W : w P_f f\}$. Given P_w , we can similarly define an acceptable firm and the set of acceptable firms for w as $A(P_w) = \{f \in F : f P_w w\}$. In the above examples, the set of acceptable workers for f is $A(P_f) = \{w_1, w_3\}$

and the set of acceptable firms for w is $A(P_w) = \{f_1, f_2\}$. We let $P = (P_{f_1}, \dots, P_{f_n}, P_{w_1}, \dots, P_{w_p})$ denote the profile of all agents' preferences; we sometimes write it as $P = (P_v, P_{-v})$ where P_{-v} is the set of preferences of all agents other than v . We let P_v denote the set of all possible preference relations for agent v and let $P = \prod_{v \in V} P_v$ be the set of admissible preference profiles. We write $v \succ_v v'$ when v is preferred to v' under preferences P_v and we say that v prefers v' to v'' . Since agents will have to compare two potential partners v' and v'' that may actually be the same, we write $v \succeq_v v''$ to denote that either $v' = v''$ or else $v \succ_v v''$. In this case, we say that v likes v' at least as well as v'' . The set of agents that v likes at least as well as v'' is $U^{P_v}(v'') = \{v' \in V : v \succeq_v v''\}$.

Each firm with quota greater than one must be able to compare groups of workers. Following Roth (1985), we assume firms' preferences over groups of workers are responsive to the preferences over single agents. A preference P_f for f over sets of workers is responsive to its preference P_f over single workers if, for all $S \subseteq 2W$ such that $|S| < q_f$, 1. For all $w, w' \in W \setminus S, S \cup \{w\} \succ P_f S \cup \{w'\}$ if and only if $P_f w \succ P_f w'$; 2. For all $w \in W \setminus S, S \cup \{w\} \succ P_f S$ and for all $S' \subseteq 2W$ such that $|S'| > q_f, S \cup S' \succ P_f S'$. Responsive preferences are assumed throughout the paper.

Since firms may have to compare two groups of workers S and S' that may actually be the same, we use $\succ R_f$, a responsive extension of R_f . We write $S \succeq R_f S'$ to denote that either $S = S'$ or else $S \succ P_f S'$. We let $U^{P_f}(S) = \{S' \subseteq 2W : S' \succeq R_f S\}$ denote the set of groups of workers f likes at least as well as S . An outcome for the college admissions problem (F, W, P) is a matching μ , a mapping μ from the set V into $2W \cup V$ satisfying the following:

1. For all $w \in W$, either $\mu(w) \in F$ or else $\mu(w) = w$;
2. For all $f \in F, |\mu(f)| \leq q_f$ and $\mu(f) \subseteq 2W$;
3. For all $(w, f) \in W \times F, \mu(w) = f$ if and only if $w \in \mu(f)$.

Observe that, while a worker may be matched to a firm or to himself—the latter meaning being unmatched—a firm is always matched to a subset of workers and being matched to the empty set stands for being unmatched. We denote the set of all matchings by M .

We can extend preferences over partners to preferences over matchings in the following, natural, way: each worker's preferences over matchings correspond precisely to his preferences over his own assignments at the matchings; similarly, firms' preferences over matchings are tantamount to the preferences over its assignments. For instance, w prefers μ to μ' when $\mu(w) P_w \mu'(w)$, while f prefers μ to μ' if $\mu(f) \succ P_f \mu'(f)$.

A matching μ is individually rational if, for every $w \in W, \mu(w) R_{ww}$ and if, for every firm f and w in $\mu(f)$, $w P_f f$.⁵ A firm f and a worker w are a blocking pair for μ if they are not

matched under μ but prefer one another to one of their assignments, i.e., $w \succ_{\mu} \mu(f)$ but $f \succ_{\mu} \mu(w)$, $w \succ_{\mu} \mu(f)$, and either (i) $|\mu(f)| < q_f$ or (ii) if $|\mu(f)| = q_f$ then there exists $w' \succ_{\mu} \mu(f)$ such that $w' \succ_{\mu} \mu(w)$. A matching μ is stable if it is individually rational and if there is no blocking pair for μ . Note that the stability of μ depends on preferences over individuals, irrespective of the responsive extension that is being used. We let $I R(P)$ and $S(P)$ denote the set of all individually rational and the set of all stable matchings respectively with respect to a profile P . A firm f and a worker w are achievable for each other if f and w are matched under some stable matching.

The proof of existence of stable matchings in Gale and Shapley (1962) is constructed by means of the deferred-acceptance algorithm. For a given preference profile P , proposals are issued by one side of the market accordingly, while the other side merely reacts to such offers by rejecting all but the best in P . In the case that firms make job offers, the algorithm arrives at the firm-optimal stable matching $\mu^F [P]$, with the property that all firms are in agreement that it is the best stable matching. The deferred-acceptance algorithm with workers proposing produces the worker-optimal stable matching $\mu^W [P]$ with corresponding properties. Further, the optimal stable matching for one side of the market is the worst stable matching for every agent on the other side of the market, a result presented in Knuth (1976) but attributed to John Conway.

3 Random matching and ordinal Nash equilibria

Many matching markets do not employ centralized procedures. Agents are free to issue offers and make acceptations and rejections as they please and matching is performed over the telephone network, using the mail, or through the Internet. In such environments, randomness determines the order in which agents communicate: it may depend on which telephone call goes through, on the speed of the mail, or on how fast firms react to eventual proposals. When a central clearinghouse does exist, chance is widely used to restore procedural fairness—any deterministic mechanism is bound to favor a subset of the agents involved. In two-sided matching markets, the need for compromise solutions is especially intense given the strong polarization of interests of agents reflected in the structure of the set of stable matchings. Some real life applications of random procedures concern allocation problems as on-campus housing, namely in American universities, or public housing.

6 Student placement mechanisms that assign students to colleges are another example of mechanisms where randomness plays a role, as well as procedures used to match students to optional courses or even children to summer camps.⁷ Finally, randomness is present in any matching mechanism where the position in a queue or the order of arrival may influence assignments.

Formally, a random matching rule $\tilde{\mu}$ is a mapping from preference profiles to lotteries over the set of matchings: $\tilde{\mu} : P \rightarrow \mathcal{M}$. A random matching $\tilde{\mu}[Q]$ is the image of a

preference profile Q under a random matching rule, i.e., a lottery over matchings. Throughout the paper, we consider only random stable matching rules by restricting the range of random matching rules to the set of lotteries whose supports are subsets of the sets of stable matchings, i.e., we consider $\tilde{\mu}$ such that, for every Q in P , the support of $\tilde{\mu}[Q]$, denoted by $\text{supp } \tilde{\mu}[Q]$, is included in $S(Q)$. While $\tilde{\mu}[Q]$ denotes a lottery over matchings, we let $\tilde{\mu}_v[Q]$ represent the probability distribution induced over agent v 's achievable matches. Whenever the probability distribution $\tilde{\mu}[Q]$ is degenerate, we abuse the notation slightly by letting $\tilde{\mu}[Q]$ denote the unique outcome matching; similarly, if the distribution $\tilde{\mu}_v[Q]$ is degenerate for some agent v , $\tilde{\mu}_v[Q]$ denotes v 's unique match in the random stable matching $\tilde{\mu}[Q]$. Observe however that in general $\text{supp } \tilde{\mu}[Q]$ is a subset of the set of stable matchings $S(Q)$. In contrast, a deterministic matching rule μ is a function from preference profiles to matchings: $\mu : P \rightarrow M$. We consider only deterministic stable matching rules that produce a unique stable matching $\mu[Q]$ for every profile of preferences Q . In particular, μ^F and μ^W denote the deterministic stable matching rules that yield the firm-optimal $\mu^F[Q]$ and the worker-optimal $\mu^W[Q]$ stable matchings, respectively, for every Q in P . Finally, we let $\mu_v[Q]$ denote v 's partner under the matching $\mu[Q]$.

In a matching market (F, W, P) , we consider the game induced by a random stable matching rule $\tilde{\mu}$ in which agents are each faced with the decision of what strategies to act on. As a first approach, we examine a one-period game where the strategy space of player v in the game is the set of all possible preference lists P_v .

Given the true preference ordering P_v , each player v may eventually reveal a different order Q_v over the players on the other side of the market, and then a matching μ stable with respect to the stated preferences Q is selected at random among all the potential matchings, i.e., the elements of $\text{supp } \tilde{\mu}[Q]$. To be precise, we consider the mechanism $(P, \tilde{\mu})$, where P is the set of admissible strategy profiles and $\tilde{\mu}$ is a random stable matching rule; we refer to $(P, \tilde{\mu})$ as a random stable matching mechanism. Once the preferences P of the agents are specified, the above mechanism induces the game $(P, \tilde{\mu}, P)$. Analogously, (P, μ) is a deterministic stable matching mechanism that induces the game (P, μ, P) . In Sect. 5, we discuss an extension of the obtained results to a more complex setting where agents' strategy spaces are broader.

In the game $(P, \tilde{\mu}, P)$, agents compare probability distributions over matchings when deciding which strategic course to take. Since preferences are ordinal, there is no natural utility representation of these preferences for expected utility calculations. It follows that to address strategic questions we need to develop ideas about what constitutes a "best decision" to be taken by an agent. With this purpose in mind, let \hat{Q} be a strategy profile and consider $w \in W$. Let $\tilde{\mu}_w[\hat{Q}](S)$ be the probability that w obtains a partner in $S \subseteq F \setminus \{w\}$ when the profile \hat{Q} is used in the game $(P, \tilde{\mu}, P)$; in particular, let $\tilde{\mu}_w[\hat{Q}](UP_w(v))$ be the probability that w is matched to a partner at least as good as v when the profile \hat{Q} is used in $(P, \tilde{\mu}, P)$.

Given a random stable matching rule $\tilde{\mu}$ and given \hat{Q}^w , we say that the strategy Q^w stochastically P^w -dominates Q_{-w} if, for all $v \in F^w$, $\tilde{\mu}[Q^w, \hat{Q}^w](UP^w(v)) \geq \tilde{\mu}[Q_{-w}, \hat{Q}^w](UP^w(v))$. This means that, for all $v \in F^w$, the probability of w being assigned to v or to a strictly preferred agent is higher under $\tilde{\mu}[Q^w, \hat{Q}^w]$ than under $\tilde{\mu}[Q_{-w}, \hat{Q}^w]$. Similarly, given $\tilde{\mu}$ and given \hat{Q}^f , we say that the strategy Q^f stochastically P^f -dominates Q_{-f} if, for all $S \subseteq W$ and for every responsive extension \tilde{P}^f of P^f , we have $\tilde{\mu}^f[Q^f, \hat{Q}^f](U^{\tilde{P}^f}(S)) \geq \tilde{\mu}^f[Q_{-f}, \hat{Q}^f](U^{\tilde{P}^f}(S))$.

This means that f is not able to increase the probability of obtaining any set of workers S_{-f} (with whom it may end up matched) and all sets ranked higher than S_{-f} in its list of preferences \tilde{P}^f , when using Q_{-f} instead of Q^f . Hence, if we consider the problem that agent v faces given the strategy choices \hat{Q}^v of the other players, a particular strategy choice Q^v may be preferred if it stochastically dominates every other alternative strategy. This provides the basis for the solution concept we will adopt throughout the paper.

Definition 1 The profile of strategies Q is an ordinal Nash equilibrium (ON equilibrium) in the game $(P, \tilde{\mu}, P)$ if, given Q^v , Q^v stochastically P^v -dominates every alternative strategy Q_{-v} for every agent v .

It follows from the above definition that Q is an ordinal Nash equilibrium when no agent v can gain in expected utility terms by unilaterally deviating from Q^v , no matter what utility function is used to represent its true preferences. We will then be concerned in finding a profile of strategies Q that is a Nash equilibrium for every utility representation of agents' preferences.

4 Equilibrium analysis

We now turn to characterize ordinal Nash equilibria in the game induced by a random stable mechanism. Proposition 1 asserts that no ordinal equilibrium supports more than one stable matching. Using the decentralized interpretation, we can say that the outcome in equilibrium is immune to the order in which agents meet when players behave strategically, even though truth revealing often leads to a lottery over matchings. Agents manipulate to protect themselves against uncertainty.

Proposition 1 Let Q be an ordinal Nash equilibrium in the game $(P, \tilde{\mu}, P)$. Then, a single matching is obtained with probability one.

Proof By contradiction, assume that Q is an ON equilibrium in $(P, \tilde{\mu}, P)$ and $|\text{supp } \tilde{\mu}[Q]| \geq 2$. Then, there exists a worker $w \in W$ and matchings $\mu, \hat{\mu} \in \text{supp } \tilde{\mu}[Q]$ such that $\mu(w) \neq \hat{\mu}(w)$. Let $\mu_{-w}(w)$ be the best match among all given by the elements of $\text{supp } \tilde{\mu}[Q]$, i.e., $\mu_{-w}(w) R^w \mu(w)$, for all $\mu \in \text{supp } \tilde{\mu}[Q]$. Let Q_{-w} be such that $A(Q_{-w}) = \{\mu_{-w}(w)\}$ and let $Q_{-w} = (Q_{-w}, Q^w)$. Note that Q_{-w} is stable for Q and, once w changes his strategy, it remains stable for Q_{-w} (it remains individually rational and no blocking pairs emerge). Further, since the set of matched agents is the same under every stable matching, w is matched to $\mu_{-w}(w)$ under every matching in $S(Q_{-w})$. Then, $1 =$

$\tilde{w}[Q_{-w}](UPw(\tilde{w})) > \tilde{w}[Q](UPw(\tilde{w}))$ and Q_w does not stochastically P_w -dominate Q_{-w} . It follows that Q is not an ON equilibrium in (P, \tilde{w}, P) . As a consequence, in the particular case that the random matching rule always assigns positive probability to at least two different matchings (if such matchings exist), the set of stable matchings of each ordinal Nash equilibrium is a singleton.

In general, however, the set of stable matchings of an ordinal Nash equilibrium may contain several elements. As proved in Ma (2002) for a deterministic stable matching rule, the random stable rule then chooses the matching that is unanimously preferred among all the stable matchings with respect to the submitted profile.

Lemma 1 Let Q be an ordinal Nash equilibrium in the game (P, \tilde{w}, P) . Then, for any matching $\mu \in S(Q)$, 1. $\tilde{w}[Q]R_{\mu}(w)$ for every $w \in W$ and 2. $\tilde{w}[Q]R_{\mu}(f)$ for every $f \in F$ and every responsive extension \tilde{R}_{μ} of R_{μ} .

Proof By Proposition 1, $\tilde{w}[Q]$ is degenerate. The result then follows from Lemma 6 in Ma (2002). For illustration, consider the following example.

Example 1 Let $F = \{f_1, f_2\}$, $W = \{w_1, w_2\}$, and $q_{f_1} = q_{f_2} = 1$. Suppose that the true preferences are as follows:

$P_{w_1} : f_1, f_2, w_1 P_{f_1} : w_1, w_2, f_1$ $P_{w_2} : f_2, f_1, w_2 P_{f_2} : w_2, w_1, f_2$. Let $Q_{w_1} : f_2, f_1, w_1$ and $Q_{w_2} : f_1, f_2, w_2$ and note that the preference profile $Q = (Q_{w_1}, Q_{w_2}, P_F)$ is an ordinal Nash equilibrium in (P, \tilde{w}, P) , the game induced by the mechanism that yields the firm-optimal stable matching. Now let \tilde{w} be a random matching rule that assigns probability 0.5 to both the worker optimal and firm-optimal stable matchings. Clearly, the support of the probability distribution induced by $\tilde{w}[Q]$ includes both $\mu_F[Q] = \{(f_1, w_1), (f_2, w_2)\}$ and $\mu_W[Q] = \{(f_1, w_2), (f_2, w_1)\}$. By Proposition 1, Q is not an ordinal Nash equilibrium in the game (P, \tilde{w}, P) . In fact, every worker can successfully deviate. For example, by using his true preferences, w_1 obtains his preferred firm f_1 with probability one.

In the context of deterministic mechanisms, Roth (1985) shows that by suitably falsifying their preferences, agents can induce any individually rational matching with respect to the true preferences. Unfortunately, this is not a very illuminating result: the set of individually rational matchings includes all the matchings that are remotely plausible. Moreover, the possibility of sustaining matchings where agents hold non-acceptable partners is not ruled out, although individual rationality appears to be a minimum requirement for an equilibrium outcome.

The results that follow establish that \tilde{w} can be supported as an ordinal equilibrium if and only if it is individually rational. Hence, we provide a complete characterization of ordinal Nash equilibria outcomes in the game induced by random stable mechanisms. Furthermore, it can easily be shown that Proposition 3 can be extended to the deterministic case, providing a necessary condition for Nash equilibria in games induced by deterministic stable matching mechanisms.

Proposition 2 Let μ be any individually rational matching for (F, W, P) and let $\tilde{\mu}$ be a random stable matching rule. Then, there exists an ordinal Nash equilibrium Q that supports μ in the game $(P, \tilde{\mu}, P)$.

Proof Let Q_w be such that $A(Q_w) = \{\mu(w)\}$, for every $w \in W$, and let Q_f be such that $A(Q_f) = \{\mu(f)\}$, for every $f \in F$. Clearly, $S(Q) = \{\mu\}$ and μ is reached with probability one. Moreover, no agent can profitably deviate. To see this, take an arbitrary worker w . If $\mu(w) \in F$, the only agent that accepts w is $\mu(w)$. Hence, w faces the choice of holding $\mu(w)$ or being unmatched. Since $\mu(w) P_w w$ by individual rationality of μ , w has no profitable deviation. If $\mu(w) = w$, no firm is willing to hire w , so that w has no profitable deviation: his only alternative is to remain unmatched. Now consider $f \in F$. If $\mu(f) = \emptyset$, only those workers in $\mu(f)$ are willing to accept filling a position in f . Moreover, by individual rationality of μ , $\mu(f) \succ R_f S$, for every $S \in \mu(f)$. If $\mu(f) = \emptyset$, no worker accepts filling a position in f . In neither case can f improve upon $\mu(f)$ by deviating. Hence, Q is an ON equilibrium in $(P, \tilde{\mu}, P)$.

Proposition 3 Let Q be an ordinal Nash equilibrium in the game $(P, \tilde{\mu}, P)$. Then, the unique equilibrium outcome $\tilde{\mu}[Q]$ is individually rational for the true preferences P .

Proof By Proposition 1, a degenerate probability distribution is achieved in any equilibrium play of $(P, \tilde{\mu}, P)$. Let us say $\tilde{\mu}[Q] = \mu$. We will prove that μ is individually rational. First, by contradiction, assume there exists a worker w such that $w P_w \mu(w)$.

Suppose that, instead of acting according to Q_w , w uses the strategy Q_w such that $A(Q_w) = \emptyset$ and define $Q_w = (Q_w, Q_w)$. By considering every firm unacceptable, w is alone under every matching in $S(Q_w)$. Hence, $1 = \tilde{\mu}[Q_w](U P_w(w)) > \tilde{\mu}[Q](U P_w(w))$ and Q_w does not stochastically P_w -dominate Q_w . It follows that Q is not an ON equilibrium in $(P, \tilde{\mu}, P)$.

Now suppose that there is a firm f and a set of workers $S \subseteq \mu(f)$ such that $S \succ P_f \mu(f)$. Let S be, among all the subsets of $\mu(f)$, the one that is preferred by f . Consider Q_f , an alternative strategy for f , where only the elements of S are considered acceptable. We will show that Q_f does not stochastically P_f -dominate Q_f .

To start, consider the matching μ_S such that $\mu_S(f) = S$ and $\mu_S(f') = \mu(f')$, for every $f' \neq f$. Let $S_B = \mu(f) \setminus S$ (note that $S_B \succ P_f \mu(f)$) and $Q_f = (Q_f, Q_f)$.

Now consider the matching market $(F, W \setminus S_B, Q_R)$, where Q_R is the same profile as Q , but restricted to $W \setminus S_B$. We will prove that μ_S is stable for Q_R in this reduced market. Note that, when Q is considered, $\mu(w)$ is acceptable for every worker w , all elements in $\mu(f)$ are acceptable for every firm $f' \neq f$, and S is the preferred subset of $\mu(f)$ for f . It follows that μ_S is individually rational for Q_R .

Now suppose that (f', w) blocks μ_S , i.e., $w \succ P_{f'} \mu_S(f')$, but $f' Q_R w \succ P_{f'} \mu_S(w)$, $w Q_R \mu_S(f')$, and either (i) $|\mu_S(f')| < q_{f'}$ or (ii) if $|\mu_S(f')| = q_{f'}$ then there exists $w' \succ P_{f'} \mu_S(f')$ such that $w' Q_R \mu_S(f')$. Since only the elements of $\mu_S(f)$ are considered acceptable in

$Q_{-R} \hat{f}$, we must have $f_{-} = f$. Hence, $Q_{-R} \hat{f} = QR \hat{f}$, where $QR \hat{f}$ is the same strategy as $Q \hat{f}$, but restricted to $W \setminus SB$. By definition of Q_{-} , we have $Q_{-}(f) = Q(f)$, for every $f_{-} = f$, and $Q_{-}(w) = Q(w)$, for every $w \in W \setminus SB$. The above expression thus becomes $\hat{f} QR w Q(w)$, $w QR \hat{f} f$, and either (i) $|Q(f)| < q \hat{f}$ or (ii) if $|Q(f)| = q \hat{f}$ then there exists $w_{-} \in Q(f)$ such that $w QR \hat{f} w_{-}$. Hence, in the unrestricted market, $\hat{f} Q w Q(w)$, $w Q \hat{f} f$, and either (i) or (ii) holds with $w Q \hat{f} w_{-}$, for some $w_{-} \in Q(f)$. This means that (f, w) blocks Q under Q , contradicting $Q \in S(Q)$. Thus, Q_{-} is stable in $(F, W \setminus SB, Q_{-R})$. Note that, since f is matched to SG under a stable matching, it must hold exactly SG under the firm-optimal stable matching for $(F, W \setminus SB, Q_{-R})$, by definition of $Q_{-R} \hat{f}$ and of the firm-optimal stable matching.

Suppose SB join in. By Theorem 5.35 in Roth and Sotomayor (1990), every firm must be at least as well off in the new firm-optimal stable matching. Since only SG are considered acceptable by f in the strategy $Q_{-} \hat{f}$, f cannot improve upon SG . Thus, it must be matched to SG under the firm-optimal stable matching of the market (F, W, Q_{-}) .

Finally, notice that since $|Q(f)| \geq q \hat{f}$ and $SB_{-} = \emptyset$, we have $|SG| < q \hat{f}$. Hence, Theorem 5.13 in Roth and Sotomayor (1990) guarantees that f must hold the same workers under every stable matching in (F, W, Q_{-}) . Therefore, by deviating and acting according to $Q_{-} \hat{f}$, f will get SG with probability one instead of $Q(f)$. Concluding, $1 = \tilde{v}_f(Q_{-})(U \hat{f} f(SG)) > \tilde{v}_f(Q)(U \hat{f} f(SG))$ and Q is not an ON equilibrium in (P, \tilde{v}, P) .

The above result is as uninformative as large the set of individually rational matchings may be. Ma (2002) shows that one way to make a sharper prediction of equilibrium outcomes and guarantee stability is to go as far as refining the notion of Nash equilibrium to strong Nash and require the use of a particular kind of strategies: truncations at the match point (i.e., deleting the $(m+1)$ th and less preferred partners when matched to the m th choice). We provide a different sufficient condition for stability in the game induced by a random stable mechanism: every ordinal Nash equilibrium where firms behave straightforwardly is stable for the true preferences.

Truth telling by firms is natural in markets where firms obey some kind of objective criterion to fill their positions (e.g., universities admit students on the basis of examination scores, student placement mechanisms assign students to public schools according to the area of residence, firms hire workers according to scores given by recruiting agencies). Moreover, in situations involving uncertainty agents may have no clue about the form that effective strategies might have and straightforward behavior is always an easy resort.

Proposition 4 Let $Q = (PF, QW)$ be an ordinal Nash equilibrium in the game (P, \tilde{v}, P) . Then, the unique equilibrium outcome $\tilde{v}[Q]$ is stable for the true preferences P .

Proof By Proposition 1, a unique matching is achieved as the outcome of an ON equilibrium in (P, \tilde{v}, P) . Let us say that $\tilde{v}[Q] = \mu$. By Proposition 3, $\mu \in I R(P)$.

We will prove that $\mu \in S(P)$ by contradiction. Suppose that (f, w) blocks μ when the true preferences are considered, i.e., $w \succ f$ but $f P w$, $w P f$, and either (i) $|\mu(f)| < q_f$ or (ii) if $|\mu(f)| = q_f$ then there exists $w' \succ f$ such that $w' P f$.

Consider Q_w , an alternative strategy for w , such that $f Q_w w$ and $v Q_w w$ if and only if $v Q_w w$, for every $v, v' \in F \setminus \{f\} \cup \{w\}$. Let $\hat{Q} = (Q_w, Q_f)$. By stability under Q_w , if w is not matched to f with positive probability under $\mu \in S[\hat{Q}]$ —so that we cannot show Q_w is a deviation for w —then, under every matching in $\mu \in S[\hat{Q}]$, each position of f is filled with a worker f finds better than w . We will prove that, in this case, f has a profitable deviation when the other agents use Q_f , so that \hat{Q} is not an ON equilibrium in the game (P, μ, P) .

Let μ_f be such that $\mu_f(f) = R_f \wedge \mu(f)$, for every $f \in \text{supp } \mu \in S[\hat{Q}]$ and let Q_f be such that $A(Q_f) = \mu_f(f)$. Note that $|A(Q_f)| = q_f$. Let $\hat{Q} = (Q_f, Q_w)$. We will show that μ will get $\mu_f(f)$ under every matching in $\mu \in S[\hat{Q}]$, so that Q_f is a profitable deviation for f .

By contradiction, assume μ_f is not stable for \hat{Q} . Since the definitions of \hat{Q} and of Q_w ensure $\mu_f \in R(\hat{Q})$, this implies that there exists a blocking pair (f', w') for μ_f when \hat{Q} is considered. As, by definition of Q_f , we have $f' = f$, this means $f' \hat{Q} w' \succ f'(\mu_f)$, $w' P f'$, and either (i) $|\mu_f(f')| < q_{f'}$ or (ii) if $|\mu_f(f')| = q_{f'}$ then there exists $w'' \succ f'(\mu_f)$ such that $w'' P f'$. As a consequence, (f', w') blocks μ_f for Q_w unless $w' = w$ and $\mu_f(w') = f$. But this contradicts the assumption that, under μ_f , each position of f is filled with a worker f finds better than w . Hence, $\mu_f \in S(\hat{Q})$.

It follows from $|\mu_f(f)| = q_f$ and from Theorem 5.12 in Roth and Sotomayor (1990) that f has all positions filled under every matching in $S(\hat{Q})$; as the set of agents f finds acceptable in Q_f is exactly $\mu_f(f)$, f is matched to $\mu_f(f)$ under every matching in $S(\hat{Q})$ and, in particular, under every matching in $\mu \in S[\hat{Q}]$. Two remarks are in order. First, this result can easily be applied to games arising from deterministic stable mechanisms: stability for the true preferences is obtained in any Nash equilibrium where firms are truthful. Second, in accordance with the claims in Roth and Sotomayor (1990) concerning deterministic mechanisms, the analogous result with workers telling the truth and firms acting strategically does not hold, although it would hold when all quotas equal one.⁸ The college admissions problem, unlike the marriage problem, is not symmetric between the two sides of the market and there are substantial differences between the two when strategic issues are contemplated. Any firm with a quota greater than one resembles something like a coalition rather than an individual. Hence, allowing for manipulation on the firms' side is similar to giving such powers to sets of agents in a marriage market and, in equilibria where workers tell the truth, stability is lost.

The converse result is given in Proposition 5, asserting that every stable matching for the true preferences can be supported as the outcome of an ordinal Nash equilibrium

where firms act according to the true preferences. In fact, workers can compel any jointly achievable outcome in the game induced by a random stable mechanism, while firms behave straightforwardly.

Proposition 5 Let μ be any stable matching for (F, W, P) and let $\tilde{\mu}$ be a random stable matching rule. Then, there exists an ordinal Nash equilibrium $Q = (P, \tilde{\mu})$ that supports μ in the game $(P, \tilde{\mu}, P)$.

Proof Define Q_w such that $A(Q_w) = \{w\}$ for every $w \in W$. Clearly, $S(Q) = \{\mu\}$ and μ is reached with probability one.

Let us now prove that Q is an ON equilibrium in $(P, \tilde{\mu}, P)$. Take an arbitrary worker w and suppose that there exists a firm f such that $f P_w \mu(w)$. We claim that w cannot deviate to get matched to f . In fact, the stability of μ with respect to P implies that either $f P_f w$ —in which case f declares w unacceptable—or, if $w P_f f$, then $|\mu(f)| = q_f$ and $w _ P_f w$, for every $w _ \mu(f)$. In the latter case, since $\mu(w _) = f$ for every $w _ \mu(f)$, then $Q_{w _}$ satisfies $A(Q_{w _}) = \{f\}$ and f ends up matched to $\mu(f)$. Now consider firm f . The only workers willing to accept f are those in $\mu(f)$. Furthermore, individual rationality of μ implies that $\mu(f) \succ R_f S$, for every $S \subset \mu(f)$. It follows that f cannot improve upon $\mu(f)$ by deviating. In conclusion, Q is an ON equilibrium in $(P, \tilde{\mu}, P)$. Our next results establish a strong link between equilibria in games induced by random and by deterministic stable mechanisms. We start by pointing out that every ordinal Nash equilibrium of the random process must be a simple Nash equilibrium of a game induced by some mechanism where chance plays no role.

Proposition 6 Let Q be an ordinal Nash equilibrium in the game $(P, \tilde{\mu}, P)$. Then, there exists a deterministic stable matching rule μ such that Q is a Nash equilibrium in the game (P, μ, P) .

Proof Assume that Q is an ON equilibrium that yields μ in $(P, \tilde{\mu}, P)$. Proposition 1 guarantees that μ is the only element in $\text{supp } \tilde{\mu}[Q]$ and, by Proposition 3, $\mu \in R(P)$. Now suppose, by contradiction, that there exists no game induced by a deterministic stable matching rule μ where Q is a Nash equilibrium. In particular, consider any μ such that $\mu[Q] = \mu$ —such a rule exists since $\mu \in S(Q)$ —and assume that some agent has a profitable deviation.

Let such agent be a worker, w . Then, there exists a strategy $Q_{w _}$ such that $\mu_w[Q_{w _}] P_w \mu(w)$, with $Q_{w _} = (Q_{w _}, Q_{w _})$. This implies that $\mu_w[Q_{w _}] \in F$ since $\mu \in R(P)$. Let $f = \mu_w[Q_{w _}]$ and define $Q_{w _}$ such that $A(Q_{w _}) = \{f\}$. Observe that under any matching in $S(Q_{w _}, Q_{w _})$, w is matched to f — $\mu_w[Q_{w _}] \in S(Q_{w _}, Q_{w _})$ since it remains individually rational and no blocking pairs emerge once w uses $Q_{w _}$. Therefore, under every matching in $\text{supp } \tilde{\mu}[(Q_{w _}, Q_{w _})]$, w holds f and Q_w does not stochastically P_w -dominate $Q_{w _}$. We get a contradiction: Q is not an ON equilibrium in $(P, \tilde{\mu}, P)$.

Now assume that $f \in F$ can profit by deviating from Q in (P, μ, P) . This means that there exists Q_{-f} such that $\mu_f(Q_{-f}) \succ \mu_f(Q)$, with $Q_{-f} = (Q_{-f}, Q_f)$. Since $\mu \in R(P)$, $\mu_f(Q_{-f}) = \mu$. Define Q_{-f} such that only the workers in $\mu_f(Q_{-f})$ are considered acceptable. Since $\mu_f(Q_{-f}) \succ S(Q_{-f})$, once only the workers in $\mu_f(Q_{-f})$ are considered acceptable by f , we can guarantee that $\mu_f(Q_{-f}) \succ S(Q_{-f})$. The definition of Q_{-f} and the fact that under every stable matching firms have the same number of positions filled (Theorem 5.12 in Roth and Sotomayor 1990) imply that f holds $\mu_f(Q_{-f})$ in every element of $S(Q_{-f})$. Therefore, $1 = \mu_f(Q_{-f})(U \succ \mu_f(Q_{-f})) > \mu_f(Q)(U \succ \mu_f(Q)) = 0$ and Q is not an ON equilibrium in (P, μ, P) .

In Proposition 7, we establish a partially converse statement: the set of ordinal Nash equilibria in the game induced by a random stable mechanism includes all the strategy profiles that are simultaneously equilibria in the games induced by the rules that yield the firm-optimal and the worker-optimal stable matchings.

Proposition 7 Let Q be a Nash equilibrium in both (P, μ^F, P) and (P, μ^W, P) . Then, Q is an ordinal Nash equilibrium in the game (P, μ, P) for any random stable matching rule μ .

The following Lemma is useful in proving Proposition 7.

Lemma 2 Let Q be a Nash equilibrium in both (P, μ^F, P) and (P, μ^W, P) . Then, the set $S(Q)$ is a singleton.

Proof Assume that Q is a Nash equilibrium in both (P, μ^F, P) and (P, μ^W, P) . Suppose, by contradiction, that $|S(Q)| \geq 2$. Clearly, this implies that $\mu^F(Q) = \mu^W(Q)$. Lemma 1 in Ma (2002) implies that, for any matching $\mu \in S(Q)$, we have $\mu^F w(Q) R w^F(w)$, for every $w \in W$. Since Q is an equilibrium in (P, μ^W, P) , the same lemma guarantees that $\mu^W w(Q) R w^W(w)$, for every $w \in W$ and for any $\mu \in S(Q)$. It follows that $\mu^F w(Q) = \mu^W w(Q)$, for every $w \in W$ and we contradict the initial assumption that $\mu^F(Q) \neq \mu^W(Q)$.

Proof of Proposition 7 Suppose that Q is a Nash equilibrium in both (P, μ^F, P) and (P, μ^W, P) . By Lemma 2, $|S(Q)| = 1$. Let us say that $S(Q) = \{\hat{\mu}\}$ and assume, by contradiction, that there exists a random stable matching rule μ such that Q is not an ON equilibrium in (P, μ, P) .

Suppose then that there exists a worker $w \in W$ and an alternative strategy Q_{-w} such that Q_w does not stochastically P_w -dominate Q_{-w} . This implies that there exists $\mu \in \text{supp } \mu(Q_{-w}, Q_w)$ such that $\mu(w) P_w \hat{\mu}(w)$. Note that, since Q is a Nash equilibrium in the game induced by a stable matching rule, $\hat{\mu} \in R(P)$. Hence, $\hat{\mu}(w) R w$ and it must be the case that w is matched to a firm under every matching in $S(Q_{-w}, Q_w)$. Let $\mu_{-w}(w)$ be the best match for w in $\text{supp } \mu(Q_{-w}, Q_w)$ and define Q_{-w} such that $A(Q_{-w}) = \{\mu_{-w}(w)\}$. Since $\mu_{-w} \in S(Q_{-w}, Q_w)$ (it is still individually rational and no blocking pairs emerged), Theorem 5.12 in Roth and Sotomayor (1990) ensures that w is matched to $\mu_{-w}(w)$ under every matching in $S(Q_{-w}, Q_w)$. Then, in no game induced by a stable matching rule is Q a Nash equilibrium, since for every stable

matching rule μ , $\mu_w[Q_{w-}, Q_w] = \mu_w(w)$ and $\mu_w(w)P_w \succ^w \mu_w(w)$. It follows that no worker can profitably deviate in the game induced by μ .

Then, there exists a firm f and a strategy Q_f such that Q_f does not stochastically P_f -dominate Q_{-f} , i.e., there exists $\mu \in \text{supp } \tilde{\mu}[Q_{-f}, Q_f]$ such that $\mu(f) > P_f \mu(f)$. Since $\mu \in \text{IR}(P)$, we have $\mu(f) > R_f \mu$ and, under every matching in $S(Q_{-f}, Q_f)$, f has at least one position filled. Let μ_{-f} be such that $\mu_{-f}(f) = P_f \mu(f)$, for every $\mu \in \text{supp } \tilde{\mu}[Q_{-f}, Q_f]$. Define Q_{-f} such that $A(Q_{-f}) = \mu_{-f}(f)$. Note that $\mu_{-f} \in \text{IR}(Q_{-f}, Q_f)$ and that no pair of agents blocks μ_{-f} under the preference profile (Q_{-f}, Q_f) . Therefore, $\mu_{-f} \in S(Q_{-f}, Q_f)$ and, since firms have the same positions filled under every stable matching (Theorem 5.12 in Roth and Sotomayor (1990)), the definition of Q_{-f} guarantees that f holds $\mu_{-f}(f)$ in every element of $S(Q_{-f}, Q_f)$. Finally, for every stable matching rule μ , $\mu_f[Q_{-f}, Q_f] = \mu_{-f}(f)$ and $\mu_{-f}(f) > P_f \mu_{-f}(f)$. It follows that there exists no stable matching rule μ such that Q is a Nash equilibrium in (P, μ, P) , contradicting the initial assumption.

The proof of the above result reveals that a sufficient condition for an ordinal Nash equilibrium in the game $(P, \tilde{\mu}, P)$ is in fact being a Nash equilibrium in every game (P, μ, P) , i.e., in every game induced by a deterministic stable mechanism. This appears to be an extremely strong condition to fulfill. Nevertheless, we will now describe a class of random matching rules for which such condition becomes necessary for an ordinal Nash equilibrium.

In the particular case that μ is the empty matching, Roth and Vande Vate (1990) have shown that, in the marriage model, every element of the set of stable matchings for the revealed preferences can be achieved with positive probability when the random matching rule they define is applied. In fact, starting from a situation in which all agents are unmatched, by successively satisfying all the pairs of a stable matching, we can guarantee that this matching is reached with positive probability. This random process is an instance of what we will name as really random stable matching rule.

A really random stable matching rule $\tilde{\mu}$ assigns positive probability to at least two different elements of the set of stable matchings, i.e., $|\text{supp } \tilde{\mu}[Q]| \geq 2$ for every Q such that $|S(Q)| \geq 2$. In Example 1, the rule that assigns probability 0.5 to the firm-optimal and to the worker-optimal stable matchings is clearly a really random stable matching rule. The following result is an implication of Propositions 6 and 7 in the particular case that $\tilde{\mu}$ is really random.

Corollary 1 Let $\tilde{\mu}$ be a really random stable matching rule. Then, the profile of strategies Q is an ordinal Nash equilibrium in the game $(P, \tilde{\mu}, P)$ if and only if the set of stable matchings $S(Q)$ is a singleton and there exists a deterministic stable matching rule μ such that Q is a Nash equilibrium in the game (P, μ, P) .

Proof Follows directly from Propositions 6 and 7, and the fact that Proposition 1 implies $\text{supp } \tilde{\mu}[Q] = S(Q)$ for a really random stable matching rule $\tilde{\mu}$. \square For

illustration, consider once more Example 1 and note that the set of stable matchings for truth telling is a singleton; further, it can easily be shown that it is an equilibrium in the game induced by the matching rule that yields, say, the firm optimal stable matching. Corollary 1 thus implies that straightforward behavior is an ordinal Nash equilibrium in the game induced by the random stable matching rule described in the example.

5 Non-preference strategies

We have explored the game induced by a random matching mechanism, claiming that one of the main motivations of this paper is the study of some decentralized markets. This may be objected on the grounds that up to this point we have restricted our analysis to a one-period game where strategies are preference lists, which perfectly mirrors the functioning of a centralized market, but falls short of an illustration of a decentralized market. In particular, in matching processes of the kind described by Roth and Vande Vate (1990), at each moment in time, a pair of randomly chosen agents meets and (temporarily) matches if this is consistent with both agents' strategies. This clearly fits the structure of a sequential game. In this context, restricting each agent to hold the potential partner that is higher on some fixed preference ordering sustains the validity of the results of the preceding section. However, in a sequential game, agents can be expected to use richer strategies, conditioning behavior on the history of the game, and not necessarily acting consistently with a unique preference ordering. The strategy of matching with the first partner one meets and rejecting every other agent is an example of such kind of strategies.

One of the difficulties that arises in attempting to capture such complex forms of behavior concerns the very essence of the matching rule that, following Roth and Vande Vate (1990), we assume to be stable with respect to the revealed preferences. In fact, such definition is compromised when, for some play of the game, no list of preferences is compatible with the strategy of a player. Hence, the set of feasible strategies of the sequential game is simply too large and precludes analysis in the theoretical framework we have been using. One potential course of action is therefore to impose that under any play of the sequential game the choices actually made are consistent with some preference ordering, even though they may correspond to incompatible preference orderings when several plays are considered. We can then speak of preference orderings that are "revealed" in the course of the play. A worker who entertains the described strategy in the example above, would match the first firm to tender an offer to him under any play of the game, and reveal that this firm is preferred to every other firm that he eventually meets in the course of that play. Since meeting is random, this worker would reveal distinct preference lists under different plays of the game.

Hence, consider a sequential game where, starting from an arbitrary matching, at each moment in time, a pair of randomly chosen agents, composed of a firm and a worker, meets. Agents match upon meeting if this is consistent with their strategies. We assume that strategies are restricted to those strategies compatible with a preference ordering for each play of the game, the revealed preference ordering, even though the information gathered in the course of the play might allow for other forms of behavior.⁹ According to Roth and Vande Vate (1990), once the probability that a given pair of agents meets is bounded away from zero, each play of the game yields a matching stable with respect to the revealed orderings in the course of that play. Hence, given a profile of strategies that meets the above requirement, every outcome obtained with positive probability is stable for some revealed profile of preferences. We let $G(P)$ denote this sequential game.

In Proposition 8, we show that ordinal Nash equilibria in preference strategies, which correspond to those obtained for the one-period game, are robust to the enlarged strategy space. In fact, given a profile of preference strategies, if by means of a strategy that is not consistent with a unique preference ordering, an agent may improve his position, he is certainly capable of doing so using a simple preference strategy.

Proposition 8 In the sequential game $G(P)$, for any collection of stated preferences Q^v for agents other than an arbitrary agent v , agent v always has a best response that is consistent with a unique preference ordering.

Proof First, consider an arbitrary worker w and fix Q^w . Let s_w denote an arbitrary strategy for w , revealing a preference ordering (not necessarily the same) under each play of the game. Denote by Q_i^w the preference ordering that is consistent with s_w under some play i . In general, we have $\text{supp } \tilde{?} [s_w, Q^w] = \{?_1, \dots, ?_k\}$, where $?_i ? S(Q_i^w, Q^w)$, for $i = 1, \dots, k$. Now let Q^w be such that $A(Q^w) = \{?_j(w)\}$ where $?_j(w) R^w ?_i(w)$, for all $?_i ? \{?_1, \dots, ?_k\}$. Since $?_j ? S(Q_j^w, Q^w)$, we must have $?_j ? S(Q^w, Q^w)$ (it is still individually rational and there are fewer blocking pairs). Hence, given that the same agents are matched under any two elements of the set of stable matchings and the only firm w finds acceptable is $?_j(w)$, this worker is matched to $?_j(w)$ under every matching in $S(Q^w, Q^w)$. It follows that any lottery over $S(Q^w, Q^w)$ gives w a partner at least as good as any lottery over $S(s_w, Q^w)$. Since s_w and Q^w are arbitrary, this completes the proof for a worker w .

Now take an arbitrary firm f . Let s_f denote a strategy for f with the same properties as the strategy for w above. Define Q_i^f as the preference ordering over individual workers that is consistent with s_f for some play i of the game. Let $\text{supp } \tilde{?} [s_f, Q^f] = \{?_1, \dots, ?_k\}$, where $?_i ? S(Q_i^f, Q^f)$, for $i = 1, \dots, k$.

Consider any alternative strategy Q^f for f such that $A(Q^f) = ?_j(f)$ where $?_j(f) R^f ?_i(f)$, for all $?_i ? \{?_1, \dots, ?_k\}$ and for every responsive extension R^f of R^f . Then, $?_j ? I$

$R(Q_f, Q'_f)$ since $\exists j \exists I R(Q_j, Q'_j)$. Moreover, $\exists j \exists S(Q_f, Q'_f)$ since $\exists j \exists S(Q_j, Q'_j)$ and no blocking pairs emerged.

Given that the same positions of a firm are filled under any element of a set of stable matchings and by definition of Q_f , f is matched to $\exists j(f)$ under every matching in $S(Q_f, Q'_f)$. Since s_f and Q'_f are arbitrary, this completes the proof.

Nevertheless, this is far from being a characterization of equilibria in this new setting. In fact, the set of ordinal Nash equilibria is larger here, as the following example demonstrates.

Example 2 (Example 1 revisited) Consider the matching market in Example 1.

Let the strategy of each agent be defined as follows: $s_{fi} =$ “match only with w_i if f_1 is the first firm to meet a worker; match only with w_j otherwise” and $sw_i =$ “match only with f_i if f_1 is the first firm to meet a worker; match only with f_j otherwise”, for $i = 1, 2$. This strategy profile leads to a non-degenerate probability distribution over matchings. Namely, both $\mu = \{(f_1, w_1), (f_2, w_2)\}$ and $\hat{\mu} = \{(f_1, w_2), (f_2, w_1)\}$ are obtained with a 50% probability. Hence, Proposition 1 rules out the possibility that s can be reproduced by an equilibrium in preference strategies. Still, s is an ordinal Nash equilibrium, since any unilateral deviation of a firm or worker may either leave the probability distribution unchanged or leave the deviator unmatched with positive probability.

6 Concluding remarks

At the expense of using an ordinal equilibrium concept, we have provided a characterization of equilibria that arise in the game induced by a random stable mechanism. The analysis is set in the college admissions problem. First, we have proved that every ordinal Nash equilibrium yields a unique matching, while when agents act straightforwardly according to the true preferences several matchings may be obtained with positive probability. Hence, agents avoid uncertainty when behaving strategically. Furthermore, a matching can be reached at an ordinal Nash equilibrium if and only if it is individually rational for the true preferences. Ordinal equilibria where firms best reply by behaving straightforwardly always produce a matching stable for the true preferences. Conversely, every stable matching can be reached as the outcome of an equilibrium play of the game. In a different direction, we relate ordinal Nash equilibria in games induced by a random matching mechanism with Nash equilibria arising in the games induced by deterministic matching mechanisms. In particular, a preference profile is an ordinal equilibrium of the game induced by a matching rule that always assigns positive probability to two different matchings (if such matchings exists) if and only if the set of stable matchings is a singleton and it is a Nash equilibrium in the game induced by some deterministic stable rule. In the last section of the paper we have tried to extend the above results, derived for a one-period game where the set of available strategies coincides with the set of all possible lists of preferences, to the sequential game that may arise in a decentralized market.

Here we assume agents may use strategies that correspond to different preference orderings when different plays of the game are considered. We have shown that ordinal Nash equilibria in preference strategies are robust to the enlarged strategy space.

In what the above results are concerned, a couple of remarks is in order. The first observation concerns fairness and random matching mechanisms. In opposition to deterministic mechanisms, which are bound to favor one side of the market over the other, we have claimed that random mechanisms promote procedural fairness.¹⁰ Nevertheless, “endstate” justice is a different issue. Indeed, the results that relate equilibria in the games induced by random and deterministic mechanisms imply that every equilibrium outcome in the game induced by a random matching mechanism may be obtained by means of a deterministic mechanism. It follows that, based on these results and in what “endstate” justice is concerned, we should not expect random matching rules to improve upon deterministic ones if equilibrium behavior is to be taken seriously.

Second, the aim of the last section is to shed some light on what happens once we move towards allowing for history-dependent strategies, preserving the stability of the mechanism. The purpose of this paper is to explore strategic behavior induced by random stable matching mechanisms, and not to provide a thorough analysis of the incentives agents face in decentralized markets.¹¹ Therefore, relaxing the restriction we impose over the strategy sets would compromise our main goal.

To conclude, equilibrium behavior in random mechanisms has barely been treated in the matching literature. One of the difficulties that arises in attempting to apply the common game theoretical tools stems from the need to compare the probability distributions over matchings generated by a random rules when preferences are ordinal. By means of the concept of ordinal Nash equilibrium we have taken a step towards filling the gap in the literature, providing a fairly complete characterization of equilibrium behavior.

B.7 S_Net

Social capital and economic performance: did networks really work?

In a great extent, allocation of resources and distributive problems find answers in the institutional framework of an economic system. A widespread opportunistic behaviour, such as free riding, cheating, corruption and rent-seeking reflects the state's inability to enforce property rights and contracts or policies biasing the allocation of resources. "Good institutions" are meant systems of law and governance that grant confidence or a tacit impersonal knowledge which economic approaches has correlated for quite a long time to economic performance given the "reputation effect" supported on game theory. In the last tow decades, economic insights have enlarged the theoretical panel to discuss trust, borrowing from sociology James Coleman's social capital metaphor to label the link between social control mechanisms (at a micro or macro-level) and economic efficiency. The concept rose criticism not for what it intends to apprehend – the connection between institutions, shared believes and economic performance – rather for its promoters' intention to include the issue in the range of "capital" theoretical implications, given the similarities with other sort of capital, namely human capital. Such a criticism, however, still instigates intellectual efforts to submit the institutional level of an economic system to quantification, including networks in indexes that measure social capital at a micro-level.

These quantified approaches inserted in a wide inquiry to causes of poverty, reckon informal mechanisms instilling norms and coping with moral hazard and incentive problems under certain contexts, but not always. Political and social turmoil often entail deviationism of the state administration reinforcing the role of networks as suppliers of trust, which in turn may hinder a market-state-based system of law from developing. In any event, there is a consensus about the effects of social change on trust, causing "depletion" in the overall level of social capital, described as "an inverted, U shape curve between the density of social capital and the level of development".

If this consensus proves to have a fundament, contexts of social upheaval, often coupled with a soaring of market transactions, are prone to develop networks with low levels of cohesion which James Coleman considers as not fitted to a low risk in trust. By the same token, periods of market growth press economic organization to depend upon agency to manage distance, which may be jeopardized by opportunistic behaviour. There is an agency whenever someone (the principal) delegates decision making to a second party (the agent), who acts on behalf of the former. Theoretical insights, some more formalized than others, enrol a large scope of problems derived from the probability of the agent incurring in moral hazard and defrauding the principal's expectations, namely by the use of non-shared information (asymmetric information). Hence, agency problems can be critical in periods of transition from traditional systems of trust to market-based systems of law because they reduce

incentives to invest in reputation mechanisms and raise transaction costs. If the system under analysis does not collapse, some devices must have been promoted to cope with this paradox. One may ask what network analysis can tell about cohesion as metaphor of social capital or propensity to trust and cooperate in such contexts of market growth. Do networks always show a configuration compatible with social capital? If they do not, can we presume individuals refrain from transacting instead of setting up gadgets to protect themselves from the risk of distrust? Do those gadgets pay off, both affording a premium and contributing to foster a system of norms in the long run?

This paper follows on this track and questions the view of economic literature on networks which aprioristically consider this informal institution as systems of horizontal relations that grant information flows and peer monitoring. A full exploration of graph theory, whether to defend or contest the advantages of informal, personal mechanisms of social control, is needed and it has been seldom undertaken.

Regarding the making of the Atlantic empires as a representative context of transition, one may expect social capital depletion. Economic history has focused on family and religious common backgrounds as the main contents of the mercantile classes' ties, taking networks as efficient institutions to counterbalance the risk implied in the agency. The assessment of the levels of risk this merchant strategy is supposed to protect from and how efficient it was in seeking this purpose, are questions seldom at the core of inquiry, albeit the foreknowledge arguments based on social network propositions. However, historical records, when providing data on a global network made up of principal-agents relations, bears information to compute cohesion (or the lack of it) and to appraise the links of social capital to agency and economic performance, which is the aim of this paper.

Evidence on the surge of migration to Brazil during the 18th century-gold rush may lead to assumptions on the social heterogeneity of residents in the colony. Agency problems involving peers and familial members are well documented and attest transactions were carried out in a social environment potentially risky for trustworthy ties. The exploitation of an outstanding natural resource triggered migration before the state's actual ability to control a large territory. The embryonic peripheral administration was unable to instil "order", not just for short number of officers appointed to remote posts, but mainly due to the infiltration of the local administration in merchants' networks. Such an interlocking delayed both the making of a specialized bureaucracy and hindered the creation of a skilled group of traders. One may contend that the central administration in Lisbon and private trade organizations faced agency problems given the shortage of incentives to cooperate and widespread opportunistic behaviour.

In this paper, institutions are analysed on the level of informal mechanisms of enforcing trustworthy relations, as principal-agent relations are supposed to be. It takes a system of principal-agent relations in the South Atlantic for unit of analysis. The

Portuguese-Brazilian gold flows provide the data to compute the structure of networks which is a measure of cohesion and density of the group and an assessment of social capital. This will give the probability of trustworthy ties being breached and points the configurations in which asymmetric information would interfere with agency. After the computation of the system's network structure, the governance of trade organizations can be questioned for its adaptive potentials, amounts of gold manipulated being a proxy of efficiency. Modelling business management strategies discovers those better adapted to agency problems and enables the estimation of a premium afforded. It is argued that travelling became an irreplaceable device, revealing that a peer monitoring system may be insufficient to foster norms when opportunities for non-reiterated relation are large.

The first section makes a survey of the economic history contributions to the discussion of social capital in long distance trade to justify the approach followed in this paper. The second section makes explicit the borrowing of social capital and principles of trust from James Coleman's Social Theory to inform the hypotheses considered. The third section describes data used and presents the models to test hypotheses, which explore graph theory and a generalized linear model. The final section provides the results. Conclusion summarizes this analysis' contribution to appraise social capital and economic performance in periods of change.

1- Issues

Long distance trade is a well studied topic in economic history of the early modern times. A multifaceted system of flows and trade firms connected the peripheries of the European empires without the intermediation of the core regions, giving to the Atlantic an economic coherence that led historians to compare it to the Mediterranean Sea¹. Market integration in this large space, however, was limited by technological and institutional hindrances drawing analysis of transactions costs as far as shipping and trade are concerned, whereas social variables, such as social capital depletion, which also raises transaction costs, has not been instigated any educated estimation².

The issue on the social embeddedness of market transactions has led anthropological approaches on the role played by religion in groups' ethos. Family strategies and systems of group's reproduction identify the resources traditional societies explored to protect themselves from risk of distrust³. Conflicting interests undermining principal-agent relations is a topic seldom taken into account although suspicious climate between partners is historical documented and acknowledged by historians. Nonetheless these very situations have not oriented an inquiry that integrates distrust as a common handicap merchants had to cope with, since mercantile classes turned into idyllic examples of systems of trust in historical writings⁴. The question whether or not the social strategies already identified were enough to mitigate agency problems deserves an in-depth analysis. Considering moral hazard is a more common problem in developing countries, being cheated for misinformation or for

unauthorized applications of capital must have been a problem more widespread in early modern times than historiography has thought of.

A risk assessment in principal-agent relations has not oriented an inquiry yet, although these questions dwell upon a set of items transversal to David Hancock and Avner Greif's works⁵. Both authors looked for institutions that assured the endurance of trustworthy relations, which is another way of discussing what makes principal-agent relations work. Each of these authors, following different theoretical and methodological approaches, underlined the agents' compliance with principals' expectations. Systems based on reciprocity, meaning tasks and rights of control were interchangeable in principal-agent relationships, blurred the hierarchical features of the tie and played efficiently to incentive actors to perform as expected, mitigating agency problems.

In the cases studied by David Hancock, such reciprocity in principal-agent relations eventually prompted the formation of a leading group of North-American merchant-planters in London who left relatives and peers as their agents in the colony. These prominent actors, in turn, performed for American peers and concentrated a considerable number of connections, reflecting their reputation and the positive results of confidence supported by social background identity. Avner Greif's works on agency in the medieval Mediterranean also sheds a light on the role of reciprocity, which was correlated to business organizations based on itinerant merchants among the Maghrib Jews. This form of agency is part of a set of features that define collectivist systems, including the small size of the group and its cultural homogeneity. The network granting the connectivity of this ethnic group performed effectively as an informal enforcement institution, affording information on actors' behaviour and becoming an effective sanctioning mechanism. Agency costs could be lower than in other systems, since wages did not need to internalize a premium to prevent cheating.

This paper dwells upon these issues and brings into discussion another case study based on the Portuguese-Brazilian gold flows exploring network analysis to reappraise the importance of travelling in systems where networks did not perform as they are supposed to among the Maghrib Jews. It assesses the risk in trust which any of these authors recognized, but reconsiders the meaning of certain actors concentrating connections. In the Portuguese-Brazilian system it neither derived from horizontal relations nor from symmetric information.

Mobility is taken as a fundamental device in distance management due to technological constraints and social capital depletion in periods of market growth⁶. The mobility of actors, distinguished from migration, needs a revaluation and recent research on the North Atlantic have looked to it as having a particular function in certain phases of the making of the Atlantic space⁷. There is now enough data to support the statement that travelling in the North Atlantic was more frequent than historical literature would lead us to believe some decades ago.

The historiography of the Iberian empires in America, on the contrary, interpreted the presence of itinerant agents as a singular feature of mercantile organizations. The paramount importance of precious metals in the making of the Iberian colonial spaces drew the attention of the political and institutional circles to business management issues, which looked to itinerant agents as a gadget to break down institutional barriers to the entrance of newcomers⁸. In times when travelling merchant-commissioners (labelled as “peruleros” or as “comissários volantes”) disturbed institutionalized interests, they were suspected of being in league with interlopers by their opponents.

Any scheme that could facilitate smuggling justified the attention of governments, which explains the political caution for agency and the parlance around the subject during the Brazilian mining rush. The 1755 legislation banned itinerant agents from the Portuguese-Brazilian trade after suspicions of their cooperation with the British community residing in Portugal⁹. The legislation was a politicized view of merchant strategies, of course. But the Marquis of Pombal was expecting the interdiction to entail a selection of the class which eventually promoted a mercantile ethos coping with the State’s fiscal aims¹⁰.

Historiography of the Portuguese empire did not discuss the political speech as such and took the prejudices that led the legislator’s intentions as an actual picture of mercantile organizations. Hence, the historiography has thought travelling agency to have a direct connection to petty-trade in the Portuguese Atlantic routes and a consequence of the opening of the group of agents to low-ranked people, supposedly submitted to foreigners’ easy credit¹¹. However, once a considerable share of receivers and senders of gold remittances is known, several questions arise before the evidence on itinerant agents among wealthy merchants or having no special connections to foreigners. Thus travelling must have held a rather broader function in the management of distance than serving contraband or British interests.

The legislator’s concern with the widening of the group of dealers bears information to be explored further on. Social backgrounds heterogeneity went along with the enlargement of the number of agents which is assumed to affect the informal mechanisms of social sanctioning traditional societies or small groups can cope with¹². They are not expected to remain effective in contexts that lose cohesion which is deemed as a pre-condition of social capital density. To trace the picture of how much the market transactions increased during the mining rush, it should be recalled that the alien population in Brazil varied from 237,000 to 2,198,000 in the 18th century, mainly due to the Portuguese immigration and to slave importation¹³. Demographical changes in the colony altered the relative importance of the regions along the shores. The rise of the captaincies and jurisdictions in the interior regions of Mina Gerais and Goiás went together with the gold rush pushing the population further to the interior and implying mining regions to be supplied by local connections woven by dealers who

had to stay in Brazil for quite a long time¹⁴. Hence, the most rewarding transactions were increasingly distant from Portugal, which was a factor for Portuguese merchants' difficulties in accessing to and assessing information. On the other hand, the devices generally addressed as affording positive responses to risk of distrust in studies on merchant classes required group identity to remain unchanged, which was not the case under study.

These remarks contend that the knowledge of the group's structure is needed to discuss cohesion, its implications with social capital and to question whether or not agents had better opportunities to cheat. A new light on the role of travelling in the South-Atlantic waters will drive a different appraisal of actors who concentrated relations.

The hypotheses and theoretical foundations are expanded in the next section.

2- Theoretical insights and hypotheses

Principal-agent theoretical insights are operational whenever exchanges are not simultaneous. Given a time lag, a first party who sends goods or capital to a second party must rely on the second party to keep his word in the bargain. This very situation forces the latter to perform as an agent of the former, even if only in discrete periods of time, and establishes a link bearing problems comparable to those raised in a principal-agent relation, no matter the juridical framework that binds individuals¹⁵. Therefore, commodities and monetary flows in the Portuguese Atlantic are thought to have been supported by principal-agent relationships, since they implied one party (the agent) to be authorized to act on behalf of another party (the principal), forging conditions for non-accomplished expectations.

Agency theory takes actors as utility maximizers and potentially seeking conflicting purposes. It foresees systems of incentives to comply the agent with the principal's interests, predicting the reduction of costs of the relationship¹⁶. Still, asymmetric information remains as a handicap and the probability of opportunistic behaviour inserts ethics and trust as variables in the relationship, which are difficult to quantify. However, research on social exchange theory has been dwelling upon this issue and affords useful insights on foundations of trust and on the process through which norms are internalized in a social system¹⁷. It explores social network analysis to measure cohesion and to assess the probability of trustworthy ties being breached¹⁸. James Coleman's approach on trust considers this social resource as not granted in any relationship. It involves, at least, a pair of actors, (a trustor and a trustee) and a few propositions about this relation can be transposed to principal-agent relations, the former being the trustor and the latter the trustee, to inform hypotheses in this paper. The concept of a risk in trustworthy ties, on the one hand, and the rationale determining the actors' assessment of the cost-benefit of acting trustfully are guidelines borrowed from this social theory¹⁹.

Trust is a resource in a relationship which may be accumulated or depleted, depending on the parties' behaviour in a given social context throughout a period of time. It may be increased, say capitalized, with reiterated relations, contributing to the positive reputation of the trustee and entailing a loss both to trustor and trustee if the tie is broken. But trust may also be undermined if a set of conditions assures the trustee a gain higher than a loss when he chooses to cheat, predicting the establishment of a similar relationship with someone else not informed about his behaviour. Therefore, the gain depends on the structure of the network involving both actors, whether or not it provides future trustors information about the defecting trustee. Transposing this to agency issues, it is expected that a non-trustworthy agent in Brazil is who uses information not shared with the principal to his own benefit and does not act to fully serve the principal's interest. However, the gain from behaving likewise depends on the principal's means of exchanging information with other principals on his opportunistic behaviour. So, ties among trustors (principals) alter the trustee's (agent's) assessment of costs and benefits when choosing to cheat.

Small groups have higher potentials of developing trustworthy relationships because the size of the group increased the average probability of trustors becoming acquainted, which makes gossip an effective informal sanctioning device. Thus, it is not the size of the group, per se, the factor promoting systems of trust; it is rather the effect of the size of the group on the structure of a network of social relations which makes trust an abundant resource in small social environments²⁰. Network analysis arises as an operative method to discuss systems of trust.

This methodology and theory has played an important modelling role in fields as diverse as computer science, biology, ecology and economics. In sociology, as in history, it has found application in organizational behaviour and in appraising social capital issues, based on the intuitive notion that social connections display configurations which condition the performance of individuals or organizations. It has been improved by the graph theory since the pioneering work of Jacob Moreno in 1934 who theorized that social configurations could be mapped to visualize the structures through which information or other means of influencing individuals' behaviour could flow²¹.

Social network analysis is now an acknowledged field of research that enables the identification of prominent actors and asymmetry or reciprocity in social interaction²². One of Moreno's main sociometric concepts that have a paramount utility in this paper is that of a "star", which spots actors in a relation system concentrating numerous relations, holding a prominent position. The other fundamental concept is group cohesion²³. It informs social capital metaphor in James Coleman's social theory differently from other approaches that are focused on innovation as a dependant variable on groups' openness and brokerage to grant information flows²⁴. The structure of a network, displaying the channels for norms being internalized, founds

Coleman's concept of social capital as low risk in trustworthy relationships²⁵. The high average probability of two trustors (principals) having a common trustee (agent), being themselves connected, is a feature of networks with significant levels of closure, which are likely to emerge in small groups. This configuration improves the effectiveness of sanctions on actors that defect expectations of others which turns the network an institution able to provide a public good²⁶. The blacklisting of undesirable members becomes an informal but effective device to dissuade the breaching of trustworthy ties, promoting the enforcement of norms and the decreasing transactions costs. These are operative tools to analyse principal-agent relations in the 18th century Portuguese Atlantic world and to question the role of mobility as correlated with social capital.

It is hypothesized that Portuguese-Brazilian transactions were embedded in a risky social environment. Social and cultural heterogeneity was the factor for low cohesion that raised agency problems. The hypothesis is addressed as follows:

H1- Cohesion and the distribution of links in the network gives the probability of moral hazard and measures social capital in the system.

The standard parameters of social network analysis that display levels of cohesion and the prominence of actors (explained in the next section) will test the hypothesis. If values of clustering coefficient will be lower than density and if the degree distribution shapes a power law, pointing to the existence of hubs, the hypothesis is confirmed, i.e., the environment bears a high risk in trustworthy relations.

The second stage of analysis operates the results of the networks' parameters to study agency in this system. It assumes that actors could develop different responses to solve agency problems, but there might have been one model more efficient than others in such an environment. The organization that better managed distance and better overcame the costs of asymmetric information must have reached the possibly highest level of efficiency. The amounts of gold received in Portugal are a proxy for efficiency. Empirical evidence asserts four distinct forms of agency were adopted in the Portuguese Atlantic world, each having pros and cons summarized as follows:

- 1) The principal travelled to Brazil and did not count on any resident agent. This solution implies principal to incur opportunity costs in a long lasting voyage and to expose himself to the perils of the sea at a time when sailing was not a pleasant sport. However he had the advantage of collecting information on the spot. The price of a voyage adds costs to this solution (to be labelled as 'no agency' in following analysis);
- 2) The principal counted on an agent(s) residing in Brazil, which released this solution of opportunity costs inherent to situation 1) while it enabled principal to reap advantages of contacts woven by agents in the colony and sometimes acquainted with local royal officers. However, this additional advantage of counting on people in the

colony represents the typical situation that raises agency problems derived from asymmetric information (labelled as 'resident agents');

3) The principal dealt with agents who travelled back and forth (itinerant agents). This solves the problems raised by resident agents but it prevented the principal from exploiting the social resources afforded by agents in the colony and adds expenses with voyages to wages or commissions charged (labelled as 'itinerant' or 'travelling agents').

4) The principal adopted a complex solution by counting both on resident agents and on itinerants, which is labelled in this paper as a 'mixed agency'. It bore the benefits of solution 2) and 3) and the same costs as 3). The itinerant agent is assumed to monitor, assist and police resident agents. This solution only assures benefits if monitoring is performed by actors with acknowledged trustful reputation.

Each of these cases involves principals and agents in certain positions in the network. The fact that itinerant agents could be assigned to assisting resident agents recalls the importance of principals establishing contacts with reliable peers, a situation that can be identified at a macro-level observation by spotting the central nodes in the network. Thus the structure of the group provides further information on principals and respective ego-networks to be considered in the model which tests the second hypothesis, which is addressed as follows:

H2- In a risky environment, the organization that better deals with the costs and benefits of each form of agency grants the highest returns and is likely correlated with hubs.

The returns in gold are the proxy of the economic performance of merchant organizations. It is the dependent variable. The independent variables consider forms of agency and the features of principals' ego-networks, taking into account the number of agents, the direct links to hubs and the number of "strong ties" measured by reiterated relations.

The significance of strong ties in market performance has driven relevant work on economic sociology after Granovetter's assertion on the "strength of weak ties"²⁷. There is no intention here to add any contribution whatsoever to this interesting debate. The aim is just to test the endurance of relations based on trust. A reiterated relation throughout a period of time tells the strength of a given principal-agent relationship and enables one to infer the role of trust in economic performance.

3. Data and model

Official records of gold sent to Portugal are the data explored in this study²⁸. The number and value of remittances is the most accurate quantitative information on the Brazilian gold arrived in Europe in the 18th century²⁹. These flows of precious metal,

albeit a mirror of the mining sector, were mainly returns from Portuguese exports and re-exports to Brazil³⁰.

Due to fiscal and protection aims, private investors were only allowed to ship gold aboard the military vessels escorting the merchant fleets. The royal navy provided protection to gold shipping and charged a 1% rate over the value, which was justified as a tax rather than a freight charge in diplomas that regulated gold transportation.

Before being kept in the safes onboard, the value of each remittance was written down in royal account books in order to collect the fee further on at the Royal Mint in Lisbon, whether or not remittances had been sent already in kind. The Royal Mint just served as a customs-house as far as these fiscal purposes were concerned³¹. At the moment the receiver claimed his remittance, records made onboard were checked with the actual amount to be delivered and a register in the margins of the book prove the delivery and the payment of the tax. These are the books of the register of the 1% tax on gold remittances, called *Livros de Manifesto do 1% do Ouro*, a documental series kept in the archive of the Lisbon Royal Mint and the source of data analysed in this paper³².

Each record informs the name of a sender, the name (s) of a receiver (s), amounts of gold at stake, whether in kind, in dust or smelted bars. Senders were either people sailing to Portugal or people who stayed in Brazil and just shipped the gold. Receivers were people who claimed the remittance at the Royal Mint in Lisbon and paid the tax. If sender and receiver was one and the same person, it is assumed the individual was aboard the fleet on which the remittance itself had been shipped (as Russell-Wood did), which discloses the cases referring to people travelling back and forth³³. A few cases refer to people with no connections. The majority of cases refer to people who brought gold, both to deliver to others and to claim for themselves. Thinking of these remittances as mostly returns from business developed in Brazil, senders represent agents and receivers represent principals.

These registers started when the tax was levied in 1721, after a royal order determining that the gold was to be shipped in royal vessels. The three main Brazilian ports (Bahia, Rio de Janeiro and Pernambuco) had their own escorted fleet, organized annually. Hence, the three fleets leaving each year for Portugal issued hundreds of registers. The immense information kept in these books advises a research carried out through group sampling.

The test to the first hypothesis needs the networks' structure. The sample conceived an interval of 10 years starting in 1721 and ending in 1761. The benchmarks refer to the institutional framework of maritime communications and the 1755 legislation on agency, since the collection of the tax was supposed to end when fleets were released from sailing escorted in 1765, while the sample dating from 1761 is expected to picture the impact of the 1755 diploma on the system ³⁴.

The senders and receivers' names and sums of gold comprised in the benchmarks were exhaustively collected, which makes five networks available dating from 1721, 1731 (...) 1761, in which principals and agents represent the nodes and gold remittances represent the edges. The fact that there can be a coincidence between sender and receiver, gave the diagonal of the social-matrixes under study a particular meaning, signalling itinerant actors.

The sample gathers 16,212 remittances and involves 11,464 actors in the total five-year period³⁵. Global parameters have been computed to recognize networks' structure and their possible changes throughout the period³⁶. We concentrate on the computation of four parameters³⁷:

a) average degree (k) of a network which represents the average number of edges leaving each element of the network;

b) density (d) of the network which is the actual number of edges in the network over the maximum to make a complete graph; it varies between 0 (empty graph) and 1, (complete graph);

where L is the number of links and n is the number of nodes;

c) clustering coefficient (CC) which measures the average probability that two agents, having a common neighbour, being themselves connected, which finds the shortest cliques in the network involving 3 nodes (equation 2);

where $E(v_i)$ is the size of the neighbourhood (v_i) of the node i and the neighbourhood of i consisting of all nodes adjacent to i ;

d) degree distribution (the distribution of the number of connections of each actor) leads to the discussion of actors' centrality by the number of ties each actor has. The shape of the degree distribution in a network has important bearings: while random and regular networks tend to show a normal distribution of the degree of the nodes, in the scale-free networks the degree distribution approaches a power-law, indicating hubs. We focused on this parameter and passed over another measure of centrality, such as betweenness, which finds actors who are in the middle of a path connecting other actors, providing indirect links in the network and playing roles as brokers³⁸. In the networks under study this position 'in between' implies that a broker would be an agent indirectly connecting two principals or a principal indirectly connecting two agents, which are roles assumed to have conflicting interests, thus such position would rather refer to buffers than brokers who spread information.

The hypothesis considers that cohesion is a form of social capital. Social capital involves confidence and an effective system of rule enforcement or internalization of norms. Then, risk is inferred from results of the computation of the first three parameters. The fourth – the degree distribution – if showing a power law form, indicates nodes providing a scarce resource, in this case, trust and reputation.

The second sample reports to the block of nodes that interlock the five networks, which comprises 115 principals linked to 350 agents through 320 remittances. This set of principals is thought to test the second hypothesis for two reasons, considering principals' ego-networks involve agents acting in different periods of time: 1) the diachronic perspective will tell adaptive patterns to deal with agency problems; 2) the dynamic perspective reckons reiterated relations, indicating the endurance of trustworthy ties.

Each of the 115 principals used the four solutions previously pointed out in section 2, which can appear combined in 13 different sequences. However, these sequences display a pattern we interpret as principals' divergent sensitivity to the risk impinged into the business by resident agents in Brazil. Categories were created to characterize the three patterns disclosed:

Category A: gathers the set of principals that incurred the risk of counting on resident agents without the assignment of scouting services.

Category B: gathers the set of principals who never incurred the risk implied in Category A, choosing to use the mixed agency. One recalls this is the solution that overlapped resident agents with travelling agents, the latter assisting the former.

Category C: gathers the set of principals who never counted on resident agents, either because they travelled themselves or just used traveller agents. This is the pattern mostly based on voyages.

The model to test the second hypothesis is both informed by the networks' structure and by patterns of risk management.

The hypothesis foresees different levels of efficiency according to solutions adopted, which in turn internalize transaction costs implied in agency problems. The dependent variable is the amount of gold received. The independent variables are:

- 1) Number of years of the principal's activity (continuous variable, from 1 to 5).
- 2) Number of years each form of agency had been used (continuous variable, from 1 to 5).
- 3) Pattern of risk management identified by categories A, B and C.
- 4) Position in networks, considering insertions in large or small components. It is described by a dummy (1 for being inserted in the large component).
- 5) Number of agents included in principal's ego-network throughout the period of principal's activity.
- 6) Distance from central actors, described by a dummy (1 for being adjacent to central actors)
- 7) Number of reiterated relations (from 1 to 4).

The dependent variable is not well adjusted by a normal distribution, thus a linear regression model is not adequate in this case. Since the response variable (amount of gold) has proven to be approximately Gamma distributed, a generalized linear model has been used instead, defined as follows:

where \hat{y}_i = expected amount of gold,

and β_i = the model linear component derived from the design matrix Z_i (to identify the relevant explanatory variables) and β , the vector of parameters.

The choice of the best model has been made through both forward and backwards analysis, both leading to the same results. Since backwards analysis starts with the most complex model (with all possible variables) and eliminates the least important one in each step (until all the ones left are relevant and the model fits the data), results from this strategy will be presented, since the significance of all variables is tested and shown³⁹.

4- Results and discussion

The structure of networks revealed a large component connecting 43% to 63% of the population. A component is the minimum set of nodes that can be connected to each other by at least one path. The remaining population is connected in very small components (the largest one observed embraces 35 nodes in 1761), comprising isolated nodes, dyads and triads (Table 1).

The topography of relations in the largest component is given through the computation of clustering coefficient, density and average degree which provides measures to discuss cohesion. Results indicate a low value for clustering coefficient, meaning a low average probability of two actors, having a common neighbour, being themselves connected. The values varied from 0.01 to 0.04 and are inferior to density (varying from 0.035 to 0.055) which assures the lack of cohesion. (Table 2).

Transposing these results to the subject under analysis, one may infer these networks' structures do not assure information flows between principals who had a common agent. Within the largest components, subgroups are connected according to the sociometric concept of the "star" (figura com a estrela – Meira, subgrupo A), agents being at the core, adjacent to several principals, who are not directly connected. Such a structure spots prominent actors who were performing as agents. This configuration asserts the low potential of the network to foster norms and to make the system work as a controlling mechanism over defrauding agents, since communication among principals is absent and would be the condition for the network to provide this kind of public good. The imbalance of power suggests that principals were the vulnerable party in the relationship. The agents certainly controlled information on "business secrets" of their principals who might not have conditions to share information on their common agent, which is consistent with low values of clustering coefficient.

The average degree, varying from 2.3 to 2.5, is also quite small, displaying a sparse network (Table 2). One should note that a similar average degree was found in a socio-matrix made up of Florentine aristocratic marital ties, studied by John Padgett⁴⁰. If such a low average degree does not surprise in the Florentine case, due to the nature of relations, it is quite significant in a market system.

The values of these parameters demonstrate the low cohesion of the group under study. Regarding Coleman's assertion on social capital, such a configuration would not make it "less risky for people in the network to trust one another"⁴¹. Chances of an agent being blacklisted for opportunistic behaviour were scanty, not only because there were no significant links connecting principals who shared agents, but also because the wide number of actors involved increased the probability of reestablishing relations with other principals. Such structures delay the process of internalization of norms. Courts in the colony could not be effective as a third party enforcement institution (the Supreme Court of Bahia should cover the whole territory), and when formal and informal sanctioning are barely efficient, cheating may become a rewarding choice. Social network parameters assessed a high risk derived from low social capital. The matter is not whether networks are alternative institutions to market-based systems of law but rather how markets transform network structures making them lose cohesion.

One may conclude that dealing in the Brazilian market faced high transaction costs⁴². Times of social change and of market growth promote this kind of structures which are losing prior and effective self-sustaining mechanisms whereas a third party that supplies incentives for obligations and expectations being accomplished is not effective yet⁴³.

The structure of the largest component in each network did not change significantly from 1721 to 1761. Results of standard parameters computation are consistent with systems which evolve without the intervention of any external ordering mechanism. The degree distribution shapes a power law (Figure 2 exemplo de 1721) which is a feature found in self-regulated systems, observed in many scientific fields. This distribution reflects a few nodes with very high degree and many with low degree, spotting hubs. In our case, these actors are seen as bearing a reliable reputation and correspond to centres of the aforementioned star-graphs, structuring subgroups around agents. Principals connected to them represent cases in which asymmetric information would be mostly evident. However, the fact that hubs in the system correspond to itinerant agents deserves a particular remark in this matter, because agent's mobility shortened the spatial distance between actors, mitigating the costs derived from such an imbalance of power.

These agents were targeted by 1755 legislation. The reduced dimension of the largest component in 1761 is the outcome of such policy. All in all, 1761 data still reckon a large component where these "charismatic leaders" had an important role for the

connectivity of the system. If the legislation would have been fully accomplished, the connectivity of the network would be much more affected. Actually, the political intents caused tensions between the central government and mercantile factions, who lobbied to constrict the cases in which interdiction should be applied. Diplomas edited in 1762 allowed ship captains to perform as agents, which was just a way of formalizing a situation that had been kept unchanged.

The significance of these hubs stems from what they could provide to the system which is confirmed by the second stage of analysis focused on the sampling of 115 principals and their business organization. Three main aspects confirm their prominence. Firstly, the highest average values of remittances received were delivered by agents having centrality in the network.

Secondly, as it has been already observed at a macro level, also among the 115 principals travelling agents represent hubs, either connected to principals who just used this form of agency or to principals who used a mixed agency.

Third, the importance accredited to these central actors who undertook travelling agency changed throughout the first half of the 18th century. The ratio between agents with high degree centrality and those with no centrality decreased, which means the former became vital in business organization. We infer from this trend, principals sought connections to reputed agents exhibiting the adaptive potentials of the system to environmental challenges, specific tasks being performed without the intervening of a strongly committed (or controlled) third-party mechanism.

Once the role of voyage to deal with a risky environment is reckoned, it is to be expected that business organizations based on this form of agency, included in category C, affords the best economic performance. However the results of the model to test the second hypothesis did not confirm it. The co-variables determinant for amounts of gold received were category B and number of strong ties. The other co-variables had no relevant effect on the dependant variable when considered together (appendix 1).

Category B defines the strategy based on mixed agency, meaning resident agents were regularly assisted by people coming back and forth⁴⁴. Thus, managing distance could not renounce setting up contacts with individuals rooted in the colony.

The historical fundamentals of these results deserve a more detailed explanation. It tells that in spite of the risk implied in counting on agents placed at months of distance, the most rewarding returns needed people to stay there for long periods of time. They assured that goods flooding the market when fleets arrived from Portugal, were sold throughout the following year, exploiting better opportunities, namely when prices rose long after the fleet had sailed back to Portugal. Moreover, the provisioning of interior regions, passing through a chain of middlemen merchants to supply an enlarging market, depended on credit facilities which could only be afforded by

individuals settled for enough time to wait for the return of transactions in the mining areas⁴⁵.

Besides these advantages, residing agents often overlapped trade activities with administrative tasks, which added political capital to principal-agent relationships. Such promiscuity between trade and royal administration had an indirect effect on business by reducing the cost of agency (actually, reducing commissions). Letters of a Portuguese merchant, Francisco Pinheiro, to and from his agent Francisco da Cruz living in Minas Gerais in the 1730's, describe how the local colonial administration intertwined business. In his correspondence to Francisco Pinheiro, Da Cruz recurrently asked him to speed up the bureaucratic procedures at the Court in Lisbon that would give him the post of local court scribe (*escrivão de ouvidoria*) in Sabará and reminded Pinheiro that both would profit from his influence at the Court. He contended that the envisaged exchange of personal favours would also benefit Pinheiro because after getting the post, Da Cruz would have further sources of income that allowed him to lower his commission rate⁴⁶.

Indeed, personal graces and favours became part of business accounting. This is a well enough discussed subject in historical literature. The debate however has been overwhelmed by the fiscal consequences of such promiscuity. It has been strained on royal officers' compliance, turning contraband into an informal institution⁴⁷. The impact of political and administrative local networks on management is seldom referred besides the obvious compensation in tax costs. As far as merchants were concerned it was also a means of reducing costs of agency.

Finally, infiltrations in the southern border with the Spanish empire, namely in "Colónia do Sacramento" (under Portuguese administration until the Treaty of Madrid (1750) recognizing the Spanish rights over the territory), also asked for local connections. Links to this area opened the way to the Spanish silver and to the arbitrage of both precious metals out of European markets, in places where the variation of prices could afford occasionally a better premium⁴⁸.

Resident agents had advantages but costs of agency analysed in this paper were higher in this situation. They could be mitigated either by raising commissions, to include a premium that dissuaded opportunistic behaviour, or by diversifying connections to catch people albeit with a different social backgrounds, well acquainted to local administration. Otherwise, sending regularly reputed agents to collect information on the spot seems a rational strategy. The fact that category B afforded the best results proves that the first alternatives mentioned above (included in category A) had drawbacks. These drawbacks stand out in the negative correlation of this category with the number of reiterated relations ($-.305 (.001)$)⁴⁹. This just reinforces the argument that there were high probabilities of agents residing in Brazil, freed from scouting devices, incurred often in moral hazardous behaviour, leading to the breach of relations.

Travelling agents could collect information on the spot and assist the resident ones. Such a function was undertaken by merchant commissioners with good reputation (hubs in the network). This is the reason for mixed agency, integrated in category B, becoming the pattern that assured the best economic performance. The additional cost with the voyages of travelling agents was certainly lower than the rate by which commissions had to rise in Brazil to cover the risk of distrust.

The other variable determining economic performance (strong ties measured by reiterated relations) regards the role of social capital in economic performance through a different angle. One should start to recall that the dimension of principal's ego-network inserted in category B is higher than that of principals in the residual categories. But it is not this feature which explains the best performance of their business organization. It is rather the higher probability of establishing strong ties when the range of options is larger.

Reiterated relations, lasting more than 20 years in this study, show the important meaning of strong ties in social systems in transition periods. The number of strong ties prompts better results in any of the three categories that represent patterns of risk management. However, the premium for choosing a mixed agency (category B) increases with the number of strong ties.

The analysis on an extended number of actors involved in long distance trade in the South Atlantic highlights the role of travelling in trade organization and how it coped with agency problems. It was far from being an option confined to petty-trade or smugglers, as it has been constantly stated by the historiography on the Iberian empires. It certainly responded to other purposes in a system with low social capital.

Conclusion

Mobility in long-distance trade management has not been correlated to a negative expectation about others' behaviour. This paper considered that hypothesis. Given the endurance of traveller agency in the Portuguese-Brazilian trade, it became a matter of testing whether it could have been an adaptive response to high risks of moral hazard.

Flows of gold remittances in the Atlantic, taken to an overwhelming extent as returns from transactions (commercial or financial), traced a large network of receivers and senders in a period of market growth. Global parameters applied in network analysis revealed extremely sparse networks, with low clustering. The probability of the network being effective as an informal enforcement institution is rather low and so chances of getting rewards for opportunistic behaviour could be high. The analysis on networks gave theoretical support to the hypothesis that social capital, while a metaphor of systems of trust, does not fit the Portuguese-Brazilian market. Business management had to deal with transaction costs derived from low social capital.

The factor for a low risk in trustworthy ties depends on the structure of relations, and on how they trace cohesive social spaces. Therefore, Greif's propositions on the effectiveness of a network to assure the spread of information on the reputation of individuals are valid if the structure of relations exhibits cohesion. The author infers it from the size of the ethnic group of the Maghrib Jews. Still, one may question the actual function of travelling agency in this system. As for the Portuguese case, it was demonstrated that travelling agents are to be correlated with low levels of social capital.

Degree distribution showed a power law. It enables the identification of individuals who polarized agency relations and to assess how uncommon they were. As far as the system under analysis is concerned, hubs represent actors who could provide scarce social resources, which certainly included confidence. They would build their reputation by supplying the system with an asset essential to solving agency problems: trust.

Hubs in this system proved to be itinerant individuals, at the centre of star graphs, performing as agents of several principals. Social background identity between pairs is not a constant. The group heterogeneity remained throughout the first half of the 18th century, gathering actors with different economic occupations and, if inserted in the royal local administration, expected to bring other sorts of capital into the business organization. Trustful ties in this system demanded other devices beyond any possible means of blurring the hierarchical feature implied in a principal-agent relationship. If the concentration of relations may signalize efficient responses to agency problems, as happened in the North Atlantic with cases studied by David Hancock, in the Portuguese South Atlantic market such a concentration means shortage of individuals with reputation, and those who have it, travelled back and forth often. Whoever established connections to these individuals could reach a comparative advantage.

The value of gold was taken as a proxy of economic performance of organizations within this system. Two main variables explain the amounts received by each principal: patterns of managing risk of agency and strong ties. Principals linked to central agents, who traveled back and forth and monitored sedentary agents, were better off. Besides, the premium for having strong ties was higher for principals who chose mainly this mixed form of agency.

One may predict that, in due time, the increased importance of travelling commissioners and their demonstrated reputation would contribute to the making of an ethos based on values such as honesty and probity, whereas the group of gold dealers remained opened to the entrance of individuals of distinct social backgrounds. The most reputed ones would reinforce the concentration of relations and would finally blacklist those who persisted in incurring in moral hazard.

The legislation of the Marquis of Pombal, which intended to form a merchant elite, selecting the class through the creation of formal institutions that closed the group, sped up a process which this paper proves was being carried out, although the legislator did not perceive it. The Marquis had targeted an irreplaceable form of agency after he arrived in Lisbon coming from a diplomatic mission in London where he learns the concept of interlopers which he will apply to itinerant agents in Portugal. The legislation could hardly be applied and triggered tensions between the state and merchants, which lobbied to change the spirit of the law to allow the role of travelling commissioners to ship captains. In spite of the opposition raised, the policy was on the right track to create an institutional framework that narrowed the group, fostered its cohesion and eventually improved market performance. However, in the Portuguese case, that prerequisite was being achieved through the mobility of actors. The underestimation of the role of traveling merchant commissioners was Pombal's mistake.

B.8 S_Econ

The Fellowship of Econometrics - Selection and Diverging Views in the Province of Mathematical Economics, from the 1930s to the 1950s

1. The foundation of econometrics

When, in 1926, Frisch defined econometrics for the first time, his aim was the transformation of economics into a positivist science: econometrics should transform 'pure economics, as far as possible, in a science in the strict sense of the word' (Frisch, 1926: 1). Without empirical verification, economics could not be a science in the full sense of the word, Frisch thought, and he devoted all his efforts to the making of that science.

For that purpose, in September 1926 Frisch approached François Divisia with a bold proposal: the creation of an Association Internationale d'Economie Pure and a new journal, as a consequence of their previous correspondence having highlighted their convergence of views on the future of economics. Divisia was a highly respected French economist working on monetary theory to whom, earlier that year, Frisch had sent a copy of his dissertation and a letter outlining his views on the future of mathematical economics. In June, Divisia replied, sharing with Frisch his own ideas on these topics: 'First of all, I believe, as you do, that economic studies cannot today be restricted to the vague reasoning that the classical economists have offered, and that the help of mathematics is necessary; I even believe that economic studies must resort to more complicated mathematical notions than those generally used in sciences for which experimentation is possible'.² Although Divisia did not feel at ease with the more advanced mathematical methods, he nonetheless believed that they represented the way forward: 'Mathematical economics has very few supporters in France; myself, I don't know much about it; nevertheless, I am to be counted among those who consider that economic phenomena must be studied by methods as precise as those used in the other more advanced sciences' (ibid.).

This was hardly an encouragement and even less a commitment, but Frisch only wanted not to be opposed. Divisia, like so many economists of that period, was not exactly a neoclassical economist and suspected the methods of Walras and Pareto.³ Yet, like an even larger number, he was ready to accept the epistemological predominance and guidance of 'pure sciences' and consequently he sympathised with Frisch's move towards a thorough mathematization of economics in order to create an empirically based science, although he felt himself to be in some danger in those deep waters. Consequently, Frisch rightly interpreted the letter as an invitation to proceed. The letter sent from Frisch to Divisia in September 1926 took up the challenge and assumed that new steps could follow immediately and, moreover, that he would lead the effort:

I enthusiastically welcome the idea of a list or some other form of communication between mathematical economists of the whole world. Myself, I had thought of creating an association with a journal discussing these questions. (...) I know quite a few mathematical economists in different countries, and I consider writing one of these days a letter to each of them in order to get their opinion about the possibility of

an 'Association Internationale d'Economie Pure' and the possibility of a journal. What do you say to an *Econometrica* (the sister of *Biometrika*)? (Frisch to Divisia, September 4, 1926)

And so he did: on November 1, 1926, Frisch wrote to four colleagues, Ladislau von Bortkiewicz, Charles Jordan, Arthur Bowley and Eugene Slutsky – no one from the US. Slutsky, whom Frisch had already met in Oslo, was the most enthusiastic about the new association (Bjerkholt, 1998: 31-2), although later on he never adhered to it.⁴ The same day, Frisch informed Divisia of the initiative of this letter.

In spite of these early efforts, the decisive steps in the creation of the econometric movement were not taken until it became a European-American enterprise: when Frisch arrived in the US, he immediately found a like-minded thinker in Charles Roos, then at Cornell University, and together they prepared a five-page memorandum, which Frisch recapitulated at his Nobel lecture (Frisch, 1970: 225). The memorandum argued in favour of rigorous quantification and an empirically based science:

Two important features in the modern economic development are the application of mathematics to abstract economic reasoning (...) and the attempt at placing economics on a numerical and experimental basis by an intensive study of economic statistics. Both these developments have a common characteristic: they emphasize the quantitative character of economics. This quantitative movement in our estimation is one of the most promising developments in modern economics. (October 1927, Frisch-Roos memorandum)

This argument was championed by Frisch two months later, in his presentation to a round table at the joint meeting of the American Economic Association and the American Statistical Association:

Quantitative economics is something more than economic statistics. There is a quantitative aspect of economics which is rational and in one sense more fundamental than the empirical manipulation of numerical data on economic phenomena; namely, that part of economic theory which is concerned with the logic of our quantitative notions (...). We speak of one statistical procedure as giving a better result than another. (...) But I cannot get rid of the impression that we engage (...) in target shooting without any target to shoot at. The target has to be furnished by axiomatic economics. Clearing the ground in axiomatic economics is a job which will certainly not be accomplished within the first few years to come. (Frisch, 1927)

The responses to the memorandum and the call for a new movement were similar to Divisia's: curiosity and sympathy, but only tentative support. It was not until a couple of well-respected economists adhered to the movement that it became a force: this was the case with Schumpeter and Fisher. Schumpeter, who was by then preparing to leave Germany for his American exile, was twelve years older than Frisch, and Fisher was his elder by twenty-eight years. Both were established and leading economists and the idea was not new for them: Fisher had already unsuccessfully promoted the project of a new association of mathematical economists in 1912 (Darnell and Evans, 1990: xv,fn). Yet, the next initiative would not fail. In the autumn of 1927, Frisch met Schumpeter for the first time, at Harvard: their friendship and complicity in matters of the Society became a driving force behind the emerging movement. The American connection was to be the core of econometrics.

In February of the next year, Frisch continued with his tour in support of econometrics and visited Irving Fisher at Yale, and then Charles Roos once again at Princeton: both would soon form part of the Society's first managerial board. Later that month, on February 29, Frisch met Schumpeter and Haberler at the Colonial Club in Harvard: the abstract of the conversation, drawn up by Frisch, indicates that they prepared a new 'list of [77] econometric people' and discussed a name for the projected International Circle for the Promotion of the Econometric (sic) Studies, suggesting *Eranos Oekonommetrikos*⁵ – a scientific corpus under the name of a student club.

The answers from the 'econometric people' were quite prudent. As we saw, Slutsky adhered to the idea but not to the Society. Georges Lutfalla wrote back to Frisch, advising him not to expect crowds at the door: he had had the experience of being unable to find four hundred subscriptions to create a journal of mathematical economics in France.⁶ Despite following the movement since its incipient days, Divisia remained very prudent: 'I believe the formula 'Economic Science [in the title of the journal] would be too dangerous: it would mean we want to monopolise economic science. It may well be the essence of our thought, but I believe it is still not the moment to announce it'.⁷ Norbert Wiener, who attended the foundational meeting, was very pessimistic about the whole enterprise, as Frisch recalled later on (Frisch, 1970: 164n). Others thought the same.

But Frisch did not give up: back in Europe, he went to Italy in June and discussed the matter with Corrado Gini, while developing an intense correspondence with many others about the future Society. Returning to the US at the beginning of 1930 as a visiting professor at Yale, Frisch drew up a list of invitees to the foundational meeting of the Econometric Society, which was to be held in December, and sent out a circular letter to twenty-eight people in the name of Roos, Fisher and himself (June 17). The invitees were: Hans Mayer in Austria; Harald Westergaard in Denmark; Umberto Ricci in Egypt; Clément Colson, François Divisia, Jacques Moret and Jacques Rueff in France; Ladislau von Bortkiewicz and Joseph Schumpeter (who would come to Harvard a couple of years later) in Germany; Luigi Amoroso, Corrado Gini, Alfonso de Pietri Tonelli and Gustavo del Vecchio in Italy; Ragnar Frisch in Norway, Gustav Cassel and Bertil Ohlin in Sweden; Wladislaw Zawadzki in Poland; Arthur Bowley, John Maynard Keynes and Arthur Pigou in the UK; Thomas Carver, John Bates Clark, John Maurice Clark, Griffith Evans, Mordekai Ezekiel, Irving Fisher, Henry Moore, Warren Persons, Charles Roos and Henry Schultz in the US; and Eugene Slutsky in Russia. Jordan was no longer on the list, in spite of having been one of the first to be contacted after Divisia. Considering their answers,⁸ the promoters of the Society decided to go ahead with the inaugural conference.

The conference met as scheduled on December 29, 1930 at the Statler Hotel in Cleveland, Ohio. The meeting was held under the presidency of Schumpeter. Sixteen men, including some added to the preliminary list of invitations, decided upon the foundation of the Econometric Society: from the US, the meeting was attended by Harold Hotelling, Frederick Mills, William Ogburn, J. Harvey Rogers, Roos, Malcolm Rorty, Henry Schultz, Walter Shewhart, Carl Snyder, Norbert Wiener, Edwin Wilson, and from Europe by Frisch, Oystein Ore (who was then at Yale), Ingvar Wedervang, Karl Menger and Schumpeter. In spite of their heterogeneity, this small number of economists, sociologists and mathematicians, some of them Neoclassical, others

Institutionalists, reunited to lay the foundation of one of the societies that would reshape economics. Divergence and scepticism amidst convergence and enthusiasm, that was the founding moment of econometrics, the mathematics of the new 'social physics', as Frisch named it.

2. The creation of the Fellowship⁹ (1931-3)

The conference elected ten men to the first council of the Econometric Society: Fisher, Roos and Wilson, from the USA; and Frisch, Schumpeter, Luigi Amoroso, Ladislau von Bortkiewicz, Arthur Bowley, Divisia and Zawadzki from Europe. Fisher, who was not present at the meeting, was elected President of the Society and Divisia vice-president.

The task this small group was setting itself was immense in three different fields. Firstly, they endeavoured to create a new discipline inside economics: quite originally, the Society was created precisely in order to define its own subject. Secondly, they wanted to emulate physics and established a constitutional goal to create a social science 'to promote studies that aim at a unification of the theoretical-quantitative and the empirical-quantitative approach to economic problems and that are penetrated by constructive and rigorous thinking similar to that which has come to dominate in the natural sciences' – as stated by the Constitution of the Econometric Society, drawn up by a committee composed of Frisch, Mills and Roos.¹⁰ Thirdly, they intended to provide new solutions, rigorous and quantified as they should be, to the traditional economic problems. Mathematics was therefore proposed as the language of the new project. Everything seemed to turn around maths – but, as the events would eventually prove, the econometricians did not agree on what type of mathematics and, even worse, it became obvious that some disdained the very centrality of mathematical formalism.

From the very first day that the Society was created, the members of the council understood that its agenda was too demanding, since they just shared some general ideas and not a concrete response to any of these three goals. In fact, not only did each of them pursue their own research agenda with little connection with each other's, but they also had different visions as far as the future of the Society was concerned. This difficulty emerged immediately in 1931 as the Society took its first task: to define the criteria for membership, and most important, of the selection of Fellows, the distinguished econometricians who should settle the example. The differences, as it appears, concentrated on the value of mathematical achievements for the definition of a career in econometrics.

The founders took opposite camps: Roos and Fisher argued for an open society, whereas Schumpeter, Frisch and Bowley preferred a closed centre of excellence with mathematically trained scientists.¹¹ As a consequence of his different view of the nature of the society, Schumpeter opposed a number of names proposed by Fisher (Bjerkholt, 1998: 39-40). Fisher even complained that mathematics was too emphasised: 'I notice a tendency in the society to stress mathematics and forget economics'.¹²

The disagreement over the criteria for membership was first resolved by the statutory definition of two types of members: ordinary members and Fellows. In spite of these discussions, the Society had grown by the next year: after the meeting of the sixteen founders in 1931, 153 new members joined the association, and, fundamentally, some

of the most influential economists of the time were among that number. Irving Fisher had drawn up a list of 261 mathematical economists and many of them were approached by the founders of the Econometric Society (ibid.: 31).

The inevitable result was that the divergence was consequently translated into the choice of Fellows, and the first years of the Society were indeed dominated by the definition of criteria both for that selection and for the election process itself, which finally took place for the first time in February 1933. It took some months for the election to be held, and some members expressed their anxiety about it: the immediate election of Fellows was instrumental in preventing the discontentment of young members, since, as Cowles put it, 'with the policies of the Society guided by a group of Fellows comprising outstanding econometricians, to exclude from ordinary membership those of lesser attainment in this field, who are nevertheless interested in furthering the aims of the Society, could only result in impeding the progress of econometrics. The creation of a group of Fellows should meet all the requirements of those who crave an esoteric atmosphere'.¹³ In any case, the dominant concept was that Fellows should represent the example of a scientific tradition imposing by their achievements among peers, more than being present in the daily workings of the Society.

Finally, in the first days of 1933, an agreement was reached on the criteria for the election and Fisher formulated the requirements for the choice of a Fellow, as recapitulated by Frisch:

1. The candidate must be an economist acquainted with economic theory.
2. He must have a mathematical foundation.
3. He must have some knowledge of statistics.
4. He must have done some original work.
5. Some of this original work must have been in economic theory. (Frisch to Fisher, January 11, 1933)

Using these criteria, the votes were cast and consequently twenty-nine Fellows were elected to the Council and notified in the following August: Amoroso, Anderson, Aupetit, Boninsegni, Bowley, Colson, Gini, Haberler, Hotelling, Keynes, del Vecchio, Divisia, Evans, Fisher, Frisch, Kondratiev, Mitchell, Moore, Ricci, Roos, Rueff, Schneider, Schultz, Schumpeter, Tinbergen, Vinci, Wilson, Zawadzki and Zeuthen.¹⁴

The diverging concepts about the nature of the Society were expressed in the discussion about the appointments of candidates. Frisch wanted to include Tinbergen, 'an absolutely charming personality', but also Vinci, Gini, Weinberger, Kuhne and 'perhaps' Leontief and Marschak.¹⁵ Divisia opposed Aftalion.¹⁶ Fisher favoured E. Cannan ('one of the first to distinguish between a stock and a flow'), Thomas N. Carver ('the only one who has developed certain points in regard to the coordination of distribution'), E. Kemmerer ('has used a little bit of mathematics'), William Ogburn ('familiar with the application of correlation to mathematics'), a description that indicates the fairly unimpressive state of the art of mathematical economics at that time.¹⁷ Schumpeter suggested Volterra and preferred Taussig to Carver (both from Harvard),¹⁸ although Frisch opposed both of these and proposed John Black ('certainly has the econometric attitude, even if he does not master much of the mathematical

technique').¹⁹ The final list resulted from multiple compromises among these opinions and eventually expressed the prestige and influence of the main candidates: Mitchell received the maximum number of votes possible, 57, in spite of not being involved in the endeavours of the Society, not even to consider his growing hostility towards the econometricians' works; Fisher, Frisch, Schumpeter, Divisia and Roos, the founders, received 54, whereas Keynes received 52.

Nevertheless, as the result of the election of Fellows was unsatisfactory for many, a new list was drawn up the same year. Frisch presented just one candidate, Marschak, since he had previously abandoned this proposal in view of a remark made by Divisia: according to Divisia (whom, as previously indicated, was not familiar with high mathematics), Marschak would not know a partial derivative,²⁰ but Frisch rapidly understood that this just was an unfair insinuation. Taking into account other suggestions, a list was put to the vote of the current Fellows, after eliminating some of the possible candidates, among others Hicks, Sraffa, Hayek and Morgenstern. The list included four who were elected (Allen, Bresciani-Turroni, Marschak and Ezekiel), and thirteen who were rejected (Darmois, Pietri-Tonelli, Fanno, Furlan, Hansen, Hawtrey, Leontief, Mills, Giorgio Mortara, Snyder, Otto Weinberger, E.J. Working and Holbrook Working). Consequently, the election as Fellows of two of those who had been present at the inaugural conference of the Society, Mills and Snyder, was rejected for the second time.

The next list of Fellows was only established four years later, in 1937: Cowles, Hicks, Mortara, René Roy and H. Staehle were all elected. In 1938, it was the turn of Lange, Leontief, J.C. Stamp and T.O. Yntema.²¹ By the end of the first decade of the Society's existence, forty-two Fellows represented the Olympus of the 'econometric people'.

This was, in any case, an immense success: a couple of years after its creation, the Society attracted already some of the most prominent economists and mathematicians: apart from those previously quoted, in 1935 Emile Borel, Constantino Bresciani-Turroni, Jacques Hadamard, Friedrich Hayek, William Jaffe, Otto Kuhne, Emil Lederer, Erik Lindhal, Fritz Machlup, James Meade, Ludwig von Mises, Gunnar Myrdal, Lionel Robbins, Arthur Spiethoff, Sven Wicksell and Vito Volterra were also members, among many others. Not many distinguished economists were absent of the econometric gathering.

The success of the enterprise was matched by the uniqueness of the convergence of different approaches and schools in economics, which were involved by this innovative programme: the Econometric Society was born under the project of reuniting all available capabilities in economics, notwithstanding their divergences. Its pluralistic nature is highlighted by the careful choice of the invitations as well as by the composition of the organs: at the foundational conference the Columbia school was conveniently represented and, as the editorship of *Econometrica* was attributed to Frisch, Frederick Mills was also involved as associate editor as a de facto representative of the Institutionalists. He was also asked to contribute to the redaction of the constitution of the society. Mitchell was one of the five members of the Advisory Council of the Cowles Commission, when it was formed. It was because of Mills's other obligations that he resigned to the post of associate editor in 1934, and not because he did not feel comfortable with the editorial choices of the journal for its first year. On the contrary, at least during the first years of the Society, he played an active role

advocating the virtues of affiliation to the new movement: as Edmund Day, who worked for the Rockefeller Foundation, hesitantly approached Mills in order to weight the arguments for membership, he was convinced by a battery of reasons, including the certainty that the Society would not cultivate mathematical 'esoterism' or any 'kind of separatism'.²²

The Society was moving forwards. But it was self-centred in regard to a number of internal quarrels over vague concepts; diplomacy abounded, but no important steps were taken to establish the dominance of mathematical economics, which still remained ill-defined. The fact was that Mitchell was in 1931 the most popular economist among the selected audience of the econometric people, but, despite being engaged in empirical work like very few others, he could not be taken as the promoter of a mathematically based 'pure' economics, and still less so of a science that aimed at achieving the higher grounds of the positivist realm of social physics: he did not have the appropriate 'econometric attitude', to adopt Frisch's insinuation against Black. At the same time, the econometric people were scarcely prepared for a battle for the reconstruction of economics: according to the President-elect Irving Fisher, the ability to distinguish between a stock and a flow, the 'use of a little bit of mathematics' and 'familiarity with the application of correlation' were sufficient recommendations for membership.

Indeed, the development of the Society required the creation of two instruments: an intense network of cooperation, emulation and competition, such as that provided by regular conferences, and the publication of a journal. Both instruments were at the centre of the preparatory discussions among the founders of the Society, although one was easier to establish than the other: conferences required enthusiasm, but a journal required financing, and the large endowment of the former could not compensate for the scarcity of the latter. Consequently, the priority was to set up the assemblies of econometricians. The Society organised an intense schedule of regular meetings both in Europe and in the US: each September-October in Europe, whereas the US meetings were held in December-January and June, frequently in association with other academic meetings.

A second tool, and a most valuable one, would be *Econometrica*. Since the early days of the 1920s, when Frisch first tried to convince Divisia of his plans for the creation of a new international association, the proposal of publishing a journal had been constantly evoked. Indeed, it was intensely discussed, mostly in respect of its title. Indeed, a curious feature of this early correspondence is the discussion about the name of the future journal: Frisch favoured *Oekonometrika* (admitting the influence of *Biometrika*), Divisia suggested *Oeconometrika* or *Oeconommetrika* whereas Slutsky's choice was *Economometrika*. Fortunately, the option turned out to be the first choice, the more pedestrian *Econometrica*.²³

The journal was a central piece in the project: indeed, since its conception, the association of econometricians had been supposed to be defined both by the organised corpus of membership and attendance at the conferences and by *Econometrica* as the expression of their research. But it was much harder to create the journal, since it required more than just intense work, devotion and imagination – it demanded financing in that disturbed period of the 1930s, right in the midst of the general depression and under the pressure of the imminent outbreak of war. Financing

was even harder to attain given the general ignorance of econometrics and the widespread dismissal of its potentiality: even later, when the Society was beginning to impose itself, the treasurer, Roos, noted the difficulties in obtaining funding from public or other sources, since the referees of projects were very sceptical – mathematicians were rather critical and, if asked, Jacob Viner and Carl Snyder could be ‘quite unfavourable’, whereas Mitchell and Taussig’s attitude was supposed unpredictable.²⁴ The paradox of the situation was obvious: Snyder, Mitchell and Taussig were members of the Society and yet were suspected of not favouring the financing of its projects and activities.

A miracle was needed in order to publish *Econometrica*, and it came in the form of a complete surprise: Alfred Cowles III, the son of a millionaire, the president of an investment counselling firm, Cowles and Co., and a competent statistician interested in stock market predictions, offered to pay twelve thousand dollars a year for the journal.²⁵ ‘An angel suddenly fell down from the sky’, announced Fisher to Frisch, asking for his opinion since, as the ‘original founder’, the decision was up to him, and he accepted after some hesitation.²⁶ *Econometrica*’s first issue appeared in January 1933 and it has been published ever since.

Frisch proposed Hotelling as editor, in order to avoid his own appointment, but he could not prevent it from happening: he finally took over editorial duties from 1933 until 1955. Associate editors were also appointed: first Alvin Hansen (replaced, in 1938, by Schumpeter), Frederick Mills (resigned in 1934) and Harold Davis: one economist, one statistician and one mathematician.

At least for the first decade of the journal’s existence, Frisch was the sole driving force behind its publication: he set the agenda, corresponded with the authors,²⁷ asked for articles, was the referee in most cases, discussed the papers and made suggestions, and, finally, decided on publication, changed the notation for coherence and even corrected the galley proofs. He worked immensely hard and it was his efforts that determined the survival and development of the journal; at the same time, this concentration of decision-making generated delays, since the papers and proofs had to cross the Atlantic twice before each issue, and, worse still, created new editorial problems.²⁸ The editor was the journal. *Econometrica* became one of the leading journals in the profession and fully accomplished its role as a pillar for the development of the econometric movement. Looking back, one can only be surprised by the depth and seminal influence of so many papers, by the diverse and far-reaching strategies of publication, by the attempts to motivate young colleagues, and by the strenuous efforts to combine historical memory with the promotion of technical expertise.

Compared to the *Economic Journal*, an older and well-established journal, *Econometrica* exhibits some revealing differences during this period (Appendix, Table 2). The number of general theoretical papers published in both journals is not significantly different, but the editorial strategies diverge notably in regard to the publication of articles about statistical theory and mathematics (irrelevant for EJ²⁹ and quite substantial for *Econometrica*, with approximately twice as many empirical studies). The other relevant differences were the publication of papers on the history of thought (*Econometrica* publishing twice as many as the EJ) and the publication of papers on empirical studies (EJ publishing three times as many as *Econometrica*).

Econometrica was definitely more inclined to publish papers on mathematically based theoretical and applied research and on the history of economic thought, and less on empirical applications.

The numbers are telling, but they do not reveal the discussions taking place on editorial strategies, in particular those that decided the shape of the newborn journal. In fact, in spite of its dedication to both mathematics, the powerful formal logic that was at the epicentre of econometrics, and statistical theory, Econometrica's editorial policy was not insensitive to the difficulties of affording technical treatment to abstract topics, and several of its prominent leaders frequently emphasised the importance of there being an empirical counterpart in the choice of the papers. After a nasty dispute on the publication of a difficult mathematical piece by Frisch himself, he tried to address his critics' concerns inviting Georges Bousquet, a young economist, to prepare a paper for Econometrica in order to build a bridge between the 'more mathematically oriented men in our group and the broader group of general economists who are interested in the econometric approach but who do not have the time or the background to follow the technicalities of our work'.³⁰

On the other hand, the journal was not supposed to be just a repository of difficult mathematical papers. One of its most impressive early features was the editorial concern with the recuperation of history. This highlighted the nature of the movement, which was looking for legitimacy as the heir to the ideas of giants: in its very first issue, Econometrica surveyed Cournot and Wicksell; in 1934, it included papers on Von Thunen, Edgeworth, Jevons and Walras; in 1938, the famous paper on Pareto and another on Cournot, then the next year on Schultz and again on Cournot; in 1946, a memorial on Keynes. In looking at the past, Econometrica was pointing to the future.

3. Mathematics of induction or of deduction? A 1953 reconsideration of the criteria to select the Fellows

By the end of its first decade, the ES had changed and the paradigmatic coexistence among different views of econometrics was terminated. The epicentre of the mutation was the Cowles Commission, a think tank funded by Alfred Cowles and reuniting a small group of full time and invited researchers, which became the brain of the operation, although it was at first very difficult to convince any prestigious econometrician to take the post of director and to move to Colorado Springs, where its headquarters were settled. After much deliberation, Oskar Lange was selected and given the job in 1937. But in 1939 the Commission moved to Chicago, under the influence of Schultz and Yntema, who took the job – both who were at war with the Columbia Institutionalists –, and the days of the happy convergence were over (Mirowski, 1989). As the Commission was now established at a major university and its prestige had grown, the post of director became a highly desirable position: Marschak succeeded Yntema in 1943 and stayed until 1948; then Koopmans became director from 1948 to 1954.

Marschak was responsible for a major change in direction in the development of the Cowles Commission. He recruited Haavelmo (July 1943) and Koopmans (July 1944) as research associates, as well as Arrow, Klein, Domar, Modigliani, Patinkin and Simon (Hildreth, 1986: 8). This new team drove the Commission towards structural

estimation, following the seminar Marschak had previously held at the New School of Social Research in New York, which was attended by Haavelmo, Schumpeter,³¹ Leontief, Modigliani, Wald, Koopmans, Samuelson and Arrow from 1940 until his move to Chicago, and developing Schultz's programme. For the first time, the elite of US econometrics was directly involved in the day-to-day research of the Commission. A witness of that period, Lawrence Klein, marvelled face to the intensity and diversity of the working of the Commission: 'it was the most unusual group of people there. To think of having Marschak, Koopmans, Haavelmo, Hurwicz, Anderson, Patinkin and eventually Arrow, Herman Rubin, Roy Leipnik and Herman Chernoff, with many visitors like Jan Tinbergen and Ragnar Frisch. It was just a tremendous number of people who were unusually talented, and they all congregated in that one place' (Klein, 1987: 413).

Yet, this programme for generalised use of probabilistic theory for structural estimation was soon exhausted. In spite of massive efforts and the elaboration of sophisticated techniques, these did not lead to very different estimation results in relation to standard OLS. As a consequence, although since 1932 the motto of the Commission had been 'Science is measurement' (Christ, 1952: 61), structural estimation waned away over the course of the 1940s. Consequently, empirical research in structural estimation and 'econometrics started to become a secondary interest of the Cowles staff as the 1940s ended' (Epstein, 1987: 110). Simultaneously, Frisch's parallel research programme on business cycles, avoiding the probabilistic approach, was also paralysed: 'By 1939 Frisch's original research program for the Institute was in shambles, as his high-profiled business cycles/time series project had fallen apart' (Bjerkholt, 2005: 520).

Furthermore, the relation with the economics department at Chicago was tense and deteriorated. The department had always been the centre for intense scientific innovation: during its first period, it was dominated by Veblen, Mitchell and J.M. Clark, who had moved to Columbia; then, the duo Frank Knight and Jacob Viner transformed it into the counter-Institutionalism head-quarter. The Chicago School of the 1920s was a very peculiar mixture of atypical neo-classical economists of Austrian and Marshallian inclination but suspecting the efficiency of *laissez-faire*, supporting government intervention in recessions but later hostile to the Keynesian movement. It attracted mathematically trained economists such as Lange, Schultz and Paul Douglas, who were more of the Lausanne abeyance, but Knight himself did not bet on mathematics. When the Cowles people, with Marschak and Koopmans, came to the University, the tension was unavoidable. Koopmans had replaced Marschak in June 1948: it was time for the Commission to turn to developments in General Equilibrium theory and models, to the discontent of some econometricians. Koopmans generated a 'metamorphosis' of the Cowles Commission, both in structure and in research (Mirowski, 2002: 249).

Although this turn of events is quite well known, the concomitant discussions inside the econometric province are generally ignored, in spite of its importance. Oskar Morgenstern was the initiator of the critique of the standard criteria to select the Fellows, but this represented an obvious expression of malaise in relation to Koopmans' strategy of promoting abstract generalizations based on general equilibrium models. At first, Morgenstern 1953 comments were presented as simple proposals for valuation of two criteria for the choice of next Fellows: they should have

proceeded to applied work (and not only abstract modelling) and created new methods. Morgenstern himself had provided such new methods with Von Neumann, in their 1944 book on the "Theory of Games". Later, in 1959, in a project at the NBER, he developed these insights for an international monetary system, considering pairs of countries as a 2-person game, with data from four countries: France, England, USA, Germany (Morgenstern, 1959).

In any case, his proposal ignited an intense debate with pros and cons, arguing about the type of mathematical and statistical adequate to the purposes of the Society.

Morgenstern's remarks were not explicitly directed against Koopmans or the Cowles Commission, but he offered an alternative view on the role of theory, of empiric models and of the economic research. This is available by means of the correspondence deposited at the Koopmans Archive (Yale University) and at the Morgenstern Archive (Duke University).³² The correspondence reveals two methodological views on the course of econometrics, namely on the articulation between the empirical and theoretical levels. The correspondence began by an apparently naïve suggestion by Morgenstern to his Fellows, in a letter to the executive director, Rosson L. Cardwell. Cardwell elaborated a report on this proposal and the following reactions to it:

Recently a suggestion from Professor Oskar Morgenstern was circulated which said, in part, "in my view the Fellows ought to be persons who have done some econometric work in the strictest sense. That is to say, they must have been in one way or another in actual contact with data they have explored and exploited, for which purpose they may have even developed new methods."

(R.L. Cardwell to the Fellows of the Econometric Society, September 18, 1953, T. Koopmans Archive, University of Yale, Folder 105)

This proposal for selection criteria valued those who developed empirical work or applied innovative methods, and it divided the Fellows. Some supported Morgenstern's proposal and some rejected it: of the eleven economists whose answer is known, six agreed whereas five opposed. To the first group belong J. Åkerman, O. Anderson, R. C. Geary, P.C. Mahalanobis, C. F. Roos and E. Schneider,³³ and all of them interpreted Morgenstern's suggestion as a radical shift in standard practices inside the ES.

Johan Henryk Åkerman (1896-1982), a Swedish economist, worked on the theory of economic cycles. Considering Morgenstern's proposal, he favoured empirical work as a criterion for selection. Besides the relevance of mathematical expertise and knowledge in economics, Åkerman emphasised the connection with empirical work:

J. Åkerman: "My comment is that Professor Morgenstern is fundamentally right. In later years one has received the impression that the requirements for an econometrist are absolutely to be a mathematician, knowledge of economics being desirable while contact with data is not necessary. It seems to me that these tendencies are more conspicuous in the content of *Econometrica* than in the election of Fellows. As statistical data analysed by statistical methods form the basis of our science, economic theory formulating rules of interdependence and mathematics constituting a language properly expressing some of its laws, the present stress on mathematical forms as the

Society' main theme may formalize our science and estrange it from the trends of economic and social reality."

(R.L. Cardwell to the Fellows of the Econometric Society, *ibid*)

This comment was developed in a letter Åkerman sent to Morgenstern, arguing that there are two kinds of econometric work: a methodological and deductive one, and another which is based on empirical analysis. This last one should be developed:

If the content of *Econometrica* were to be divided in two parts – one including contributions of a deductive kind as well as methodological studies and the other contributions based on facts – it is possible that a new stimulus could be given to empirical work in the domain of econometrics.

(J. Åkerman to R.L. Cardwell, July 30, 1953, Archive of the University of Duke, Box 39)

Another supporter of Morgenstern's proposal was Oskar Anderson, but there is no statement explaining the reasons for this support: Cardwell just states 'O. Anderson: "I wholly agree with the suggestions of Professor Oskar Morgenstern"' (R. L. Cardwell to the Fellows of the Econometric Society, *ibid*).

Robert Charles Geary (1896-1983) was an Irish statistics expert responsible for the division of National Accounts in the United Nations, in New York, between 1957 and 1960. Geary aimed at demonstrating the relevant scientific significance of empirical work. For him, the choice should be between work in descriptive statistics, with no connection to economical theory, and the use of statistical approaches for econometric work:

R.C. Geary: "I entirely agree with Professor Morgenstern. One of the main preoccupations of directors of national statistical offices is that so large a part of the vast output of their statistics is rarely put to any scientific use (and sometimes to no use at all). In the world generally econometric work, in Morgenstern's sense, is negligible in relation to the volume of statistical raw material. There is a perfectly clear distinction between mathematical statistics in the one hand and econometrical applications on the other. The Society should accept as candidates for Fellowships only those persons who made sophisticated applications of statistical data."

(R. L. Cardwell to the Fellows of the Econometric Society, *ibid*)

In Morgenstern's archive at Duke University there is a copy of Geary's letter explaining minutely the reason for supporting his suggestion, in reply to Cardwell's circular letter of July 21 1953. To Cardwell's, Geary adds that theoretical development is much more advanced than empirical applications:

The reason why there are one hundred research economists and ten mathematical statisticians to each econometrician is mainly because econometrics is a far more stern and onerous discipline than mathematics or theoretical economics. The econometrician knows that a long series of calculation may not yield a single useful result. Even negative results have, however, a definite value in sometimes showing the invalidity or inadequacy of economic assumptions. Theory is far ahead of practice. It would be no great harm if the mathematically-minded could impose a self-denying ordinance on themselves for a period as regards theory, and apply instead existing theory to actual data.

The society should take a positive attitude on this important matter with a view generally to the development of genuine econometric work. It might, for instance, circularise the economic faculties of universities throughout the world suggesting that practical econometrics should be included in the courses of economics. This would mean that laboratory work (the laboratories being equipped with computing machines, books of reference, etc.) would become part of the training of economists.

This comment goes far beyond Professor Morgenstern's suggestion. It is submitted because it is of vital interest to national statistical offices. I have sent a copy of this letter to Professor Morgenstern.

(R.C. Geary to R.L. Cardwell, July 24, 1953, Archive of the University of Duke, Box 39)

The Indian scientist Prasanta Chandra Mahalanobis (1893-1972) devoted himself to statistical application and founded the Indian Institute of Statistics. He had studied in England with R.A. Fisher, and his work was particularly praised by Hotelling. In what concerns Morgenstern's suggestion, Mahalanobis also considered the importance of empirical work as opposed to abstract mathematical formalism:

P.C. Mahalanobis: "I agree with Professor Morgenstern's views about the importance and need of actual contact with data. Abstract work and the use of mathematics are but means to an end. In my opinion, econometric work must seek justifications in the interpretation and utilization of information in the form of statistical data so that the value of such work should be assessed not by the abstract quality of the tools but the fruitfulness of the results. Due consideration should, of course, also be given to fundamental advances in technical methods when there are good prospects of using such methods in practice."

(R. L. Cardwell to the Fellows of the Econometric Society, *ibid*)

Charles F. Roos (1901–1958) had been the director of Cowles Commission between September 1934 and January 1937, and he was a founder of the Econometric Society, to which he presided in 1948. His work was mainly developed in dynamic economics and planning, as he distanced himself from the typical econometric work, and he also supported Morgenstern.

Erich Schneider (1900-1970) was a lecturer at the University of Kiel, in Germany, and director of the journal of the Institut für Weltwirtschaft. His training in economics was later complemented by the study of mathematics and physics. In 1932 he became associate lecturer in Bonn, together with Schumpeter, devoting himself to the theory of monopoly and to different forms of market structure. In 1936 he became a lecturer at Aarhus University, in Denmark, where he was one of the founders of the faculty of economics. He eventually became close to Scandinavian economists, particularly Frisch and Haavelmo. In 1946 he accepted a position as professor at the University of Kiel. He was a member of the Econometric Society, having organised a conference of the Society, in Kiel, in 1955 (Bombach, 1971). Schneider encouraged the study of statistics in the courses in economics in Germany and this explains his approach:

E. Schneider: "I am in perfect agreement with prof. Morgenstern's remarks. A necessary condition to be able to be elected as a Fellow of the society should be that the candidate in his research work has been in actual contact with data and has done some significant econometric research work."

(R. L. Cardwell to the Fellows of the Econometric Society, *ibid*)

On the other hand, the economists disagreeing with Morgenstern argued for keeping the empirical or statistical work under the vigilance of theory – a distinctive brand of the previous debate on ‘measurement without theory’. Furthermore, they feared that, were the criterion of practical application to be applied, some valuable theoreticians would be excluded from the Society. In this group, contesting Morgenstern's proposal, counted L. Amoroso, G. C. Evans, R. Roy and essentially the tenors of the Society and the Cowles Commission, T. Koopmans and J. Marschak.

The Italian Luigi Amoroso (1886-1965) convoked his mathematical inclination to declare that:

L. Amoroso: “I, too, have more than once experienced the unpleasant feeling mentioned by Prof. Morgenstern in seeing pages crammed with formulae which, in general, apply only with limitations to the economic reality. And in this connection allow me to recall Edgeworth’s golden maxim: if discretion in the use of algebraic symbols constitutes an elegance for the mathematician, it constitutes a duty (or: a necessity) for the economist. I cannot, however, consent (or: agree) with the opposite thesis which, it seems to me, is implicit in the thought of my eminent colleague. I.e., I cannot agree that econometrics should consider as its principal task the mathematical analysis of statistical data, with the purpose of obtaining directly the empirical laws governing the phenomenon in question.

In reality econometrics (like Janus’ temple) is two-faced: at the same time an inductive and deductive science. And these two lines must be harmoniously blended, avoiding any excess in one or the other direction.

By this criterion I let myself be guided in indicating candidates for fellowship; this I do within the limits of my possibilities which are bounded by my incomplete knowledge of the papers written by the individual authors.”

(R.L. Cardwell to the Fellows of the Econometric Society, *ibid*)

Amoroso’s interpretation assumes Morgenstern’s suggestion to imply the need for more inductive statistical work. His rejection of Morgenstern's proposal is based on his non-acceptance of an exclusively inductive interpretation of econometrics practice and in this sense he quoted Edgeworth, considering the use of mathematics essential to econometrics.

The American mathematician Griffith Conrad Evans (1887-1973) was a lecturer at Rice University, Texas, who, by 1934, moved to the University of California, in Berkeley. As responsible for the department of mathematics at Berkeley he recruited some European mathematicians and statistic experts, who were crucial for the development of econometrics, such as Jerzy Neyman. He was recognised as one of the first researchers on dynamic economics (Morrey, 1983). His comment to Morgenstern's suggestion reflects his essentially mathematical work, rejecting the proposal on the grounds that it might narrow econometric work down to statistical work:

G. C. Evans: “I do not see surpassing merit in Professor Morgenstern’s proposed restriction on candidacy for fellowship in the Econometric Society. It may well be that a person’s contribution will depend fundamentally on something other than statistics and that a corresponding statistical problem will be treated by another. Perhaps I am

influenced by the fact that I have myself only a neighbourly interest in statistics, but there are problems – e.g., the theory of money, - in which statistics seems to play a secondary role.

Certainly, sound statistical investigations are very important in themselves, as well as to serve as a brake or as an incentive to speculation, and it is a matter of congratulation that government statistical work, for example, is in so much better shape than it was a few years ago. The Econometric Society has not been without influence. But I believe that part of its influence on the progress of economics is due to the broad statement of Section 1 of its Constitution; and I hope that the concepts inherent in the founding of the Society will continue to be exemplified in its choice of Fellows as well as in its membership and its publication.”

(R. L. Cardwell to the Fellows of the Econometric Society, *ibid*)

Together with François Divisia and Maurice Allais, René Roy (1894-1977) was part of the French Marginalist School. Roy was one of the heads of the Institut International de Statistique in Paris and specialised in the field of transport economics. He argued that Morgenstern's proposed restriction might deprive the Society of important theoretical analysis:

R. Roy: “S’il est désirable que les économètres ne s’orientent pas exclusivement vers la présentation d’exposés abstraits et s’il convient d’encourager les recherches et les publications reposant sur la mise en œuvre de données concrètes, il serait toutefois excessif d’exiger d’un candidat-Fellow qu’il ait à se prévaloir de semblables travaux, car une telle condition risquerait de priver la Société du concours actif de personnes ayant fait œuvre originale et parfois important dans le domaine des analyses théoriques.”

(R. L. Cardwell to the Fellows of the Econometric Society, *ibid*)

But the decisive reactions came from the dominant figures in econometrics at the time. Koopmans, the director of the Cowles Commission, did not agree with Morgenstern because he considered that work in theoretical economics was essential:

T. C. Koopmans: “I wish to register disagreement with Professor Morgenstern’s proposal that would limit nominations for Fellowship in the Society to scholars who have made studies involving actual processing of statistical data. While I would feel alarmed at the state of econometrics if not a large majority of the fellows of the Econometric Society were in this category, the progress of economic knowledge requires concurrent work in economic theory, statistical inference, and econometric work proper, and important contributions to the common purpose of these three activities have been made by individuals preferring to work in only one compartment.”

(R. L. Cardwell to the Fellows of the Econometric Society, *ibid*)

Marschak took the same stance. He had been until recently the director of CC, and he considered the proposed criteria for selection a dangerous limitation for econometrics:

J. Marschak: “I disagree with the statement of Dr. Morgenstern as circulated. If his suggestions were adopted the following men could never aspire to the fellowship: John v. Neumann, Vilfredo Pareto, and Leon Walras.”

(R.L. Cardwell to the Fellows of the Econometric Society, *ibid*)

The reference to Von Neumann is obviously a cripted critique of Morgenstern, since they had co-authored the "Theory of Games", the book establishing their authority in the province of economics. But it missed the target, since of all people Morgenstern was aware of Von Neumann's dissatisfaction in relation both to neoclassical economics and to econometrics (Louçã, 2007), since furthermore never before the theory had so closely justified the choices of the techniques, giving birth to abstract modelling in the framework of general equilibrium economics. Under the prevailing econometrics of the 1950s, economics became a branch of mathematical logic.

In any case, the front of rejection eventually came to dominate. As the argument goes, these econometricians stood for a theoretical choice: the integration of the different forms of analysis under the guidance of general equilibrium models. In the first group of reactions, those supporting Morgenstern's proposal, on the other hand, there was greater diversity of opinion. Morgenstern's suggestion was an appeal to further empirical work applied to economical reality. In that sense, Åckerman, Mahalanobis and Geary rejected the dominance of axiomatics and Geary rejected theoretical analysis without application to reality. Furthermore, the dispute turned out to be a difference of evaluation of the role of induction and deduction in econometrics.

This debate calls into question the work of econometrics around the reality of the model or the model of reality. That is, while structural econometrics, as represented by Koopmans, stood for the generalizations from theory, Morgenstern and his fellow thinkers argued for empirical work so that his peers would devote their effort to research real data (or new models of it, implying the search for new approaches) and not to constructions of abstractions.

Much later, both Frisch and Haavelmo would exhibit the same concerns about the danger of econometrics being trapped by abstraction. This was obvious in Frisch's argument against Playometrics, or the statistical exercises without regard to real economic problems (Frisch, 1970: 4). Haavelmo also argued on the limits of theoretical applications if ignorant of concrete facts:

Contrary to what many people seem to think, it is in the practical application of theories to facts, in attempts to draw conclusions on the concrete level, that the need for stringent logic and fancy mathematics really shows up.

(Haavelmo, 1957: 352)

It is not certain that, at that time, he was still understood by the fellowship.

4. Conclusion

This paper reassesses the creation of the Econometric Society concentrating on the internal discussions on the role of mathematics. In the first debate, some of the founders took opposed stances on the degree of mathematical expertise required for the participation in the Society: Roos and Fisher argued for an open society for all economists, whereas Schumpeter, Frisch and Bowley proposed a closed club of mathematical economists. The solution of the divergence was to define two levels of affiliation, that of normal members and that of Fellows, to be elected under strict requirements of demonstrated capacity in mathematics.

In any case, the very definition of mathematical expertise was pluralistic, as pluralistic was the conception of economics itself. In the first years of the Society Frisch and

Tinbergen were there, but also Keynes. Neoclassicals of different strands crossed paths with Institutionalists, the to-be-Keynesians and other economists. Schumpeter was one of the masterminds of the mathematization of the Society, although he was totally innocent of mathematics. Mills, the Institutionalist, was to be appointed associate editor of *Econometrica* in its first year and, although he was twice rejected in the election for Fellow, he always proved enthusiast about the society.

When the question was reassessed twenty years past, in 1953, the landscape of econometrics had dramatically changed. The pluralistic views of the required mathematical skills had been superseded by the programme of structural estimation in the 1940s, and this was in shambles by the 1950s. Morgenstern's attack on the criteria for the selection of Fellows can therefore be interpreted as an echo of his and eventually Von Neumann's dissatisfaction with the course of econometrics, and certainly as a critique of the current practices of the Society (asking for openness and for the treatment of real data) and of its dominant mathematical approach (asking for innovation). In that, Morgenstern was opposed by Koopmans and Maschak, the tenants of the abstract approach then prevailing. He was not heard.

B.9 S_Pov

Deprivation analysis based on Bayesian latent class models

1. INTRODUCTION

The aim of this paper is to analyse household deprivation, taking into account the multiple aspects that define the well-being of households. We use an appropriate methodology based on Bayesian latent class models, which allow us to quantify the households at risk of deprivation and to establish the dimensions of well-being that the households most lack. The analysis conducted considers Portuguese households for periods of time relative to the 1990s and the present decade, based on European Community Household Panel (ECHP) survey data, gathered by the Portuguese National Institute of Statistics (INE). The ECHP was implemented under Eurostat co-ordination in every European Union Member-State and nowadays provides the official data to analyse poverty. It is a harmonised longitudinal survey focusing on income, social conditions and the lifestyle of households and their component members (Eurostat, 2003).

Assuming that poverty is a multi-dimensional problem, we use a relative deprivation concept: “People are relatively deprived if they cannot obtain, at all or sufficiently, the conditions of life – that is, the diets, amenities, standards and services – which allow them to play the roles, participate in the relationships and follow the customary behaviour which is expected of them by virtue of their membership of society” (Townsend, 1993: p. 36). If a household lacks the resources (material and immaterial) necessary to obtain access to these conditions of life and thus fulfil its role and responsibilities in society, it may be said to be in deprivation.

Poverty as a multi-dimensional problem has been the object of much attention in both the political and scientific domains. The scientific study of poverty has been progressively centred on the observation of this phenomenon as a state of household deprivation. There have been many attempts to go beyond the bounds of the simple consideration of income resources (Dewilde, 2003; Förster et al., 2002; Jenkins and Cappellari, 2004; Layte et al., 2001; Nolan and Whelan, 1996; Pérez-Mayo, 2005; Whelan and Maître, 2005). In addition, the need to measure poverty from a multi-dimensional perspective has been recognised politically, which has in turn contributed to bring about some progress in social policies. A good example is the new European Social Agenda 2005-2010. Moreover, the European and national strategies are now being developed on the basis of a deeper knowledge of the living conditions of the population, particularly of those who are most vulnerable. This will allow us to construct, evaluate and support the global measures which are inherent in the multiple facets of this phenomenon.

Since the 1970s, some authors have measured deprivation by taking different stages into consideration, such as the choice of a set of variables considering multiple well-being dimensions, the evaluation of the household situation for each variable, the aggregation of the indicators considering, or not, a weighting structure and, finally, the determination of a threshold that separates the deprived population from the non-deprived (Nolan and Whelan, 1996).

In the present study, we use a deprivation measurement method structured into different stages to analyse deprivation from both partial (by well-being dimensions) and overall perspectives. This method allows us to select the variables that best identify the households' state of well-being, as well as to obtain composite indicators of deprivation. For this purpose, we use latent class models (LCM), similarly to Moisis (2004, 2005), Pérez-Mayo (2005) and Whelan and Maître (2005, 2006). Differing from other studies, we introduce a new selection of well-being dimensions into the analysis and choose some variables with more than two categories. In addition, with LCM it is possible to estimate a latent variable with more than two standard groups by well-being dimension (deprived and non-deprived), considering intermediate situations of deprivation (partly deprived). This perspective constitutes a substantial improvement in respect of the cross-sectional analysis of deprivation. Moreover, in contrast to the majority of the studies that have been made on deprivation measurement, we use Bayesian methods to draw inferences of interest due to the parametric complexity of the LCM and the need for a full accounting of the related uncertainty.

The remainder of the paper is structured in six further sections, as follows. Section 2 overviews the concept of deprivation and the ensuing need to regard the condition as a latent variable, then briefly describes the three stages of the deprivation measurement method. In Section 3, we describe the ECHP survey data on which our analysis is based. The Bayesian latent class model is introduced in Section 4. Next, in Section 5, we explain how Bayesian analysis by Markov Chain Monte Carlo (MCMC) methods was implemented, with regard to parameter inferences and model comparison, paying particular attention to the problems of non-identifiability. Our results are presented and discussed in Section 6. Finally, in Section 7, we draw some conclusions and consider the implications of our findings for further research.

2. THE CONCEPT OF DEPRIVATION AND ITS MEASUREMENT

In the developed countries, most studies focus on an indirect poverty approach, inducing a one-dimensional perspective of poverty based on the income level. However, a direct poverty approach, emphasising the multi-dimensional nature of the phenomenon, has been playing an increasing role in scientific studies at an international level. In this approach, the identification of a lack of resources is based on the direct observation of people's living conditions (Ringen, 1985).

In this paper, we use the relative deprivation concept, wherein the relative needs of the individuals and households are established from the observation of a certain society's way of life at a certain moment in time. A household is considered to be in deprivation when it cannot satisfy multiple standard well-being needs, or expectations, within a certain context of place and time, stemming from an insufficiency of material and immaterial resources.

With the observation of material and immaterial resources, we have a more extensive approach to identify different dimensions of well-being needs. Based on a growing concern to induce other perspectives of realities that are directly connected with a possible social disadvantage of the households, the following well-being dimensions are considered to be relevant: internal and external conditions of housing, durable goods, financial capacity and indebtedness, essential goods (food, clothing, heating

and lighting), health, transportation, education and training, labour market, safe environment, leisure activities and social relationships, among others.

Deprivation measurement demands more complex methodologies than that of income poverty¹. Since the 1970s², several methodologies have been proposed in respect of deprivation measurement, supported essentially in multivariate statistical techniques. Different methods of measuring deprivation are viewed as alternative means of gaining information on the same complex social problem. Measuring social science concepts directly by the observation of factual indicators is a very difficult (or even impossible) task, because they have diverse related factors. Furthermore, the deprivation concept presents complexities that render precise quantification highly difficult, in view of the large number of dependent well-being variables. It requires a careful selection of the observed well-being variables and the development of a statistical model which is able to capture the essential part of deprivation as an unobserved variable (or, more commonly, latent variable).

In deprivation research, it is crucial to find a latent variable that clusters different households' deprivation situations and somehow accounts for the association among the observed well-being variables. Given the categorical nature of observed variables, this is achieved by latent class models (see, e.g. Bartholomew and Knott, 1999).

We use a three-stage approach to measure deprivation. After a preliminary stage involving the prudent choice and treatment of several manifest categorical variables, the estimation of a deprivation indicator as latent categorical variable, based on Bayesian latent class models (BLCM), is conducted in two stages. The first stage entails a partial deprivation analysis by well-being dimension. The second comprises an overall analysis reliant upon values generated for the partial deprivation indicators from the preceding stage. This allows us to construct an overall deprivation indicator reliant upon all information gathered, overcoming the difficulties that stem from the large number of selected manifest variables. A more detailed description of these three stages follows below.

(i) Choice and treatment of manifest categorical variables

The selection of manifest variables was based on several criteria, namely: available information from ECHP survey data; earlier studies on deprivation measurement based on the same source (e.g. Bomba et al. (2007), Layte et al. (2001), Pérez-Mayo (2005)); well-being dimensions that are theoretically relevant; exploratory data analysis; and cultural, social and economic habits of Portuguese society.

Treatment of the variables is necessary throughout this stage. Firstly, it is necessary to ensure that all variables concern the same unit of analysis and present a categorisation that is valid and appropriate to the measurement of deprivation. Secondly, we must ensure that complete information is available on all of the units³.

(ii) Partial deprivation analysis by each well-being dimension

The second stage entails BLCM analysis by each well-being dimension, which allows us to discover the manifest variable(s) with a larger contribution to deprivation and to estimate a (latent) composite indicator of deprivation for each dimension. One of the latent categories of this latter indicator represents the group that experiences the most acute deprivation, to which we designate the probability, deprivation risk. The

remaining categories can also play useful role in the identification of other groups beyond the non-deprived group, depending on the selected model.

In this approach, it is also possible to estimate the probabilities of a household (with a specific well-being profile) belonging to different deprivations groups (latent classes). In order to move to the next stage, we construct hypothetical observed values for the deprivation indicator by well-being dimension in the household data base, according to allocation of the households' profiles to the latent classes by some criteria based on those a posteriori probabilities.

(iii) Overall deprivation analysis

Lastly, the third step provides an overall approach to the deprivation problem by bringing together the hypothetical values of deprivation indicators for the well-being dimensions constructed in the previous stage. At this point, an analysis of BLCM allows us to estimate an indicator of overall deprivation reliant upon three levels: non-deprived, partly deprived and fully deprived, as well as to identify the most vulnerable well-being dimensions of the household's standard of living.

3. DATA AND VARIABLES

Since 1994, Eurostat has conducted an annual survey in the EU Member-States with the purpose of collecting information on the incomes of households and individuals, as well as on the living conditions that characterise them. This survey is designated the European Community Household Panel (ECHP). In Portugal, the official entity responsible for gathering this information is the Portuguese National Institute of Statistics (INE).

The first advantage of this instrument is related to the possibility of establishing the harmonisation and cohesion of social statistics in order to achieve comparability at European level. The second advantage is related to an enlargement of households' well-being dimensions, making it possible to deepen multi-dimensional analyses on social and economic phenomena. Thirdly, its longitudinal nature allows individual and household trajectories to be followed, capturing social dynamics and even the impact of the social policies.

This paper uses the ECHP data for waves 2 and 8 (1995 and 2001) in order to evaluate the extent of deprivation in the middle of the 1990s and at the beginning of the 21st century.

The available variables of the ECHP are distinguished by the survey unit: persons (13,285 in 2001) or households (4,614 in 2001). Most of the living conditions variables characterise directly the household's well-being situation. For this reason, we chose the household as the unit of analysis and all the manifest variables considered were regarded accordingly.

Selection of the manifest variables is directly related to the information available in the ECHP in both years of analysis. The coverage represented by this source is not as fully embracing as would be desirable in terms of well-being dimensions and related manifest variables. Initially, we intended to examine the following seven well-being dimensions: housing, durable goods, economic strain, social relationships, education/vocational training, health and labour market (Eurostat, 2003). However, a small number of variables in the education/vocational training and health dimensions

and a too-restricted domain of application for the labour market dimension variables led us to withdraw these three dimensions. Thus, the manifest variables that we selected after a possible prior treatment by well-being dimensions (housing, durable goods, economic strain and social relationships) are displayed in Table A, shown in Appendix.

The ECHP survey is the result of a complex sampling design based on a multi-stage sampling. Eurostat constructed and displayed different weight variables that reflect the sample design and characteristics of persons and households. The weights were adjusted by correcting for non-response and by calibrating according to external control distributions in order to reflect the structure of the population. The external distributions used for calibration cover individuals by age and gender; households by region, tenureship status, size, number of economically active persons; and people aged 16 or over by activity status (see Eurostat, 2000b for more details). With the aim of understanding the effect of the weights, we compared point estimates from weighted and non-weighted data. There were no special differences between both sets of estimates, particularly when the number of latent classes was higher. In each case, the models were based on the conditions underlying the usage of the multinomial model, as explained in the next section. The results shown relate to weighted sampling data that reconstitute a sampling information frame that is closer to the population's true behaviour.

4. BAYESIAN LATENT CLASS MODEL

Let us denote the observed categorical variables for each well-being dimension by X_j with l_j categories, $j=1, \dots, J$. The contingency table formed by them have distinct cells and the corresponding vector of frequencies $n=(n_p, p=1, \dots, l)$ is assumed to follow a multinomial model with probability function (for any subset of $l-1$ components):

where θ and ϕ , with the θ_p component of θ representing the probability that each unit presents the observed profile (x_1, x_2, \dots, x_J) , represented by the p -th cell.

Our aim is to create a single indicator that in some form characterises the possible deprivation situation of households. This is achieved by the so-called latent class model (Bartholomew and Knott, 1999), which is simply a conditional independence log-linear model for the augmented table with the latent variable (Agresti, 2002; Paulino and Singer, 2006). This model imposes a finite mixture structure on θ in (1) expressed as:

where θ_y and ϕ_{jy} , $x_j=1, \dots, l_j$, $y=1, \dots, K$ stand for the probabilities of K latent classes and the conditional marginal probability of the categories of X_j given each latent class, respectively, satisfying the natural constraints $\theta_y=1, y=1, \dots, K$ and $\phi_{jy}=1, j=1, \dots, J$.

We denote by θ and ϕ , respectively, the vector of the marginal probabilities of the latent variable Y and of observed variable X_j conditionally on each class y , for each j . Other quantities of interest can be computed from the model parameters. For instance, the conditional probabilities of the latent classes, given the (x_1, \dots, x_J) profile, are expressed as:

Contrary to most studies on deprivation measurement (e.g. Moisiu, (2004, 2005); Pérez-Mayo, 2005; Whelan and Maître, 2005, 2006), we use Bayesian methods to make inferences due to LCM complexity and the need to take full advantage of

inherent uncertainties. Given that latent class models differ from one another in the number K of latent classes, we consider here a model selection approach reliant upon a set of fixed values for K . This implies that the unknown parameters for each model are exhausted in θ .

For this reason, the sample information contained in the likelihood function:

must be supplemented with initial information quantified by some prior on θ . Due to the meaning of θ -components, it is customary to resort to the Dirichlet family (Agresti and Hitchcock, 2005; Paulino et al., 2003). As no subjective prior information is available, independent multivariate uniform prior distributions are taken for θ and for every j and γ . This prior, together with the multinomial likelihood (4), define the Bayesian latent class model (BLCM). Since the analysis of each Bayesian latent class model was performed via WinBUGS software, the corresponding code used the well-known characterisation of the Dirichlet distribution $Ds-1(a_1, \dots, a_s)$, through the independent Gamma distributions $Ga(a_i, b)$, $i=1, \dots, s$.

It is necessary to exercise caution because of the possible over-parametrisation of some models. Note that the total number of free parameters in (4), given by, increases strongly with the number of latent classes, rapidly becoming larger than the $I-1$ counts. Furthermore, independent of over-parametrisation, the formula (4) shows that the likelihood is invariant under re-labelling of the mixture components, i.e., the so-called label-switching problem (see, e.g. Stephens, 2000a). The ensuing model non-identifiability obliges us to resort to strategies that overcome this type of problem, as referred to in the following section.

5. BAYESIAN INFERENCE

Examination of the posterior distribution kernel $\pi(\theta)$, defined over the parameter space, promptly shows the analysis of each BLCM demand, resorting to MCMC methods like the Gibbs sampler. Simulation from the full conditional distributions is not straightforward, requiring more sophisticated means, such as the slice-sampling algorithm (Neal, 2003).

In the preceding section, it was mentioned that the likelihood is invariant under the $K!$ permutations of the mixtures' component labels in θ . Given the symmetric nature of the prior distribution for the components of θ , the posterior distribution of θ is also invariant under permutation of its component labels. This feature often has undesirable consequences on the inferences drawn from the simulated posterior sample. Figure 2 illustrates some of the serious effects of label-switching in the posterior densities for the mixing probabilities in a BLCM with four latent classes relating to the Housing dimension. The traces display several jumps related to some forms of labelling, the mixture components making convergence quite difficult to be reached. Estimates of the marginal posterior densities exhibit multi-modality and much similarity, having the concomitant effects on inaccurate parameter point estimates.

The lack of identifiability caused by the label-switching problem in finite mixture models has been handled by resorting to several strategies. The most usual strategy (e.g., Richardson and Green, 1997; Bartolucci et al., 2003; Congdon, 2005; and Fruhwirth-Schnatter et al., 2002) consists of imposing artificial identifiability

constraints on model parameters. However, it is known that this does not always provide a satisfactory solution (Stephens, 2000a; Jasra et al., 2005).

An alternative strategy relies upon re-labelling algorithms applied to MCMC output, such as those suggested by Celeux (2000) and Stephens (2000a, b). Among them, due to its ease of implementation, we advocate Stephens' re-labelling algorithm based on a decision-theoretical approach through the Kullback-Leibler divergence measure, hereafter denoted as the KL algorithm.

In this work, we adopted the following approach:

- (i) Judicious choice of the manifest variables for each well-being dimension, in order to exhibit good characteristics of deprivation differentiation among households;
- (ii) Application of the identifiability constraint, giving rise to a modified model that we call constrained BLCM, whenever the original (unconstrained) BLCM revealed convergence or label-switching problems;
- (iii) Implementation of the KL algorithm in the event that the previous procedures proved fruitless.

Convergence analysis of a simulated long chain by standard methods (Gilks et al., 1996; Paulino et al., 2003) allowed us to specify a length of between 40,000 and 50,000 iterations for the burn-in period, according to the model at hand. The final sample size was approximately 8,000 iterates in general, specified after choosing convenient lags for autocorrelation elimination, by the customary analytical and graphical methods. The lag size was shown to be highly dependent on K and varied over the range from 15 to 180 iterations. We used WinBUGS software to carry out these analyses (Spiegelhalter et al., 2004).

Once an adequate simulated sample from the posterior distribution of θ was obtained, summaries of the marginal posterior densities for the selected model were calculated, namely the posterior means and 95% HPD credible intervals, the latter by means of BOA software (Smith, 2007).

Comparison of models was made by assessment of goodness-of-fit, complexity and predictive ability through DIC, BIC and PBF (pseudo-Bayes factor).

The DIC measure (Spiegelhalter et al., 2002) is based on the posterior distribution of the deviance:

where $g(n_1, \dots, n_{l-1})$ is referred to a proportionality constant which does not have any influence in the choice of the model. It is defined by:

where \hat{D} is a Bayesian measure of fit given by the posterior expectation of $D(\theta)$, d , and pD is a penalty measure by the model complexity (so-called effective number of parameters), regarded as the difference between the posterior mean deviance and the deviance of the parameter posterior mean, that is:

The BIC measure:

where d is the number of model parameters, was approximated by the variants suggested by Carlin and Louis (2000) and Raftery et al. (2007), by which the maximised log-likelihood is replaced by simulation-based versions of \hat{D} , only to the former, and

added to the corresponding posterior variance of the log-likelihood to the latter. They are denoted by BIC(C-L) and BIC(R), respectively.

Evaluation of PBF was based on the logarithm of the conditional predictive ordinates (CPO) computed from:

where z_{ip} , $p=1, \dots, l$, denote the observed vectors of cell indicators for unit $i=1, \dots, np$ and represent the simulated values of θ from its posterior distribution, given all observations but z_{ip} .

According to these criteria, the best models are those with the lowest values for DIC or BIC and the highest value for lnCPO. In order to avoid conflicting results, we further considered a sensible criterion related to a simpler and unambiguous interpretation of the latent classes.

6. DISCUSSION OF RESULTS

Table 1 displays results of the application of the above-mentioned model comparison criteria to BLCM for each well-being dimension and overall analysis, on the weighted samples from the two selected years. It also indicates whether the models are constrained or not in order to overcome the label-switching problems.

Taking these criteria into consideration together with the interpretative clarity of each model, we believe the best models to describe deprivation for the Housing and Economic Strain dimensions are those with four latent classes in both referenced years. In the Durable Goods dimension, we selected the BLCM with three latent classes in both years. Moreover, the best BLCM for the Social Relationships dimension displays a different number of latent classes between 1995 and 2001. We chose the BLCM with four latent classes in the first year of analysis and three classes in the second year. Interpretation of the latent classes can be achieved from examining the pattern of the conditional probabilities given latent classes. Tables 2.1 to 2.4 mark the category of each manifest variable with the highest estimated conditional probability (posterior mean) with the symbol (x).

In all dimensions, the fully-deprived class is identified when the highest estimated conditional probability in all manifest variables refers to categories related to greater deprivation. In contrast, we have a non-deprived class associated to categories that show good situations of well-being.

The intermediate deprivation situation (partly-deprived classes) has a different interpretation according to the corresponding well-being dimension. Focusing on the Housing dimension, we find two intermediate classes with a distinct differentiation: the partly-deprived class in housing external components (that is, characterised by the households living in an area with crime or vandalism and with pollution problems) and the partly-deprived class in housing internal components (households living in accommodation with shortage of space, problems of damp and darkness or insufficient light). For the Durable Goods dimension, the partly-deprived class corresponds to the group that is most deprived of domestic appliances, such as a micro-wave and a dishwasher. The Economic Strain dimension presents two intermediate classes. One of them relates to the household's ability to manage its general expenses, other than housing and clothing. The other describes the household's ability to manage its due responsibilities (housing costs and others). Lastly, for the Social Relationships

dimension, we find two intermediate classes of deprivation with distinct patterns of differentiation in 1995: the partly-deprived class on convivial frequency (friends, family and neighbours) and the partly-deprived class on tourism-related leisure. In 2001, the partly-deprived class on convivial frequency does not appear to be clearly identified and according to BIC and InCPO criteria, we selected the BLCM with three classes: non-deprived, partly-deprived (tourism) and fully-deprived.

To analyse overall deprivation, we constructed hypothetical observed values for the deprivation indicator by well-being dimension in the household data base, according to the following criteria concerning allocation of the households' profiles to latent classes:

(i) The partly-deprived classes were aggregated in a single class to reduce the dimensionality of BLCM on overall analysis and therefore, the deprivation indicator by well-being dimension became related to three groups (fully-deprived, partly-deprived and non-deprived);

(ii) The latent class y enclosed the household observed profile p if the posterior mean of the conditional probability of a household belonging to latent class y , given the profile p , was larger than the posterior mean of the conditional probability of any other class ($y=1, \dots, K$; $p=1, \dots, I$). This criterion was supported by very small Monte Carlo errors.

The application of methodology based on BLCM yielded some useful results to deepen the knowledge around the deprivation phenomenon. Starting with the overall deprivation analysis, the estimated latent class probabilities show clearly a substantial improvement in the well-being of Portuguese households between 1995 and 2001. In the mid-1990s, the deprivation risk of Portuguese households was estimated at 27 per cent, having decreased to 13 per cent by the beginning of the 21st century. In turn, the posterior mean of the probability of deprivation absence rose from 26 per cent to 39 per cent in the same period. With regard to estimates of conditional probabilities within the fully-deprived class, the dimensions that provided the highest contribution to overall deprivation risk were Economic Strain (93 per cent and 82 per cent in 1995 and 2001, respectively), followed by Social Relationships (69 per cent and 39 per cent) and Durable Goods (38 per cent and 36 per cent). See Table 3 for these and other posterior estimates.

More details on Portuguese household deprivation can be observed from the analysis of deprivation by well-being dimension. The corresponding outcomes are not shown, owing to the amount of space that this would require (they are available from the authors on request). However, we shall briefly summarise some of the main findings. They allow us to testify that the household's economic capacity is one of the greatest contributing factors to its deprivation, evident in both of the analysed years (1995 and 2001). Moreover, the estimated deprivation risk in the Economic Strain dimension is one of the highest found, standing at 30 per cent in 2001 (compared to 47 per cent in 1995). These most deprived households are characterised by a reduced capacity to possess an adequate heating system and to manage general expenses.

Durable Goods constitutes another well-being dimension which determines a higher extent of deprivation among households. The deprivation risk estimate was 18 per cent in 1995, declining to 9 per cent in 2001. The most vulnerable households revealed

a strong lack of domestic appliances (such as micro-waves and dishwashing machines), although telephone possession had a less important status after the mid-1990s.

With regard to the dimension referred to as Social Relationships, which has been disregarded by most of the earlier studies on deprivation, the deprivation risk was approximately 12 per cent in 1995 (1 percentage point higher in 2001). The households found in this situation are characterised essentially by the impossibility to spend a week away on holiday, to go out in their free time or to invite friends or family for a drink or a meal at least once a month.

Finally, the dimension that was found to have the least influence on deprivation was Housing. It was estimated that the deprivation risk in this dimension in 1995 was 10 per cent, with an improvement of two percentage points shown in 2001.

7. CONCLUDING REMARKS

This paper applies a methodology based on Bayesian Latent Class Models in order to enable the quantification of the households at risk of deprivation and to attain the well-being dimensions that the deprived households most lacked in Portuguese society in the second half of the 1990s and at the beginning of the 21st century.

Recognising poverty as being a multi-dimensional phenomenon, we opted for the relative deprivation concept. The households are considered to be in deprivation when they cannot fulfil multiple standard well-being needs, or expectations, in a certain place and time context, owing to the insufficiency of material and immaterial resources.

Deprivation measurement was accomplished by means of a stepwise approach reliant upon latent class modelling. Analyses of the latent class models followed a Bayesian paradigm in order to take full advantage of inference uncertainties, particularly for rather complex models. In order to carry out these analyses, we selected a set of manifest variables with two or more categories in four well-being dimensions: Housing, Durable Goods, Economic Strain and Social Relationships.

The application of this methodology to Portuguese households for the years 1995 and 2001 revealed a substantial improvement in their well-being. Deprivation risk appeared to undergo a strong decline, while correspondingly, the probability of the non-deprived group increased significantly. The individual analyses by well-being dimension demonstrated that the highest deprivation risk is found in Economic Strain, being followed by Social Relationships, Durable Goods and Housing, in this order. Moreover, these dimensions reveal the existence of intermediate deprivation groups, which allows us the opportunity to gain a deeper insight into the deprivation phenomenon. This approach enables the researcher to conduct partial and overall analyses of deprivation based on a substantial amount of categorical data from the ECHP. It would be of interest to apply this methodology to the other EU Member-States for purposes of comparison. A possible modification in the BLCM would be to incorporate the unknown number of latent classes as one additional parameter provided with an appropriate prior.

Rather than comparing the deprivation status from a cross-sectional analysis confined to two years, as was the case in the present paper, it would be far preferable to pursue a dynamic deprivation analysis, taking full advantage of the annual panel data for the

period 1995 – 2001. Some recent works have dealt with (classical) analyses of this kind through generalisations of latent class models, but confined to composite indices constructed from some manifest variables (e.g. Moisiso, 2004, Whelan and Maître, 2006). Direct usage of several household manifest variables (even per dimension) in this modelling framework raises serious issues on the grounds of sample size. One of the present authors is currently engaged in a research project on other modelling strategies towards a Bayesian longitudinal data analysis that supplements the study presented here.

B.10 S_CEO

Governance and CEO Pay and Performance in Non-Profit Organizations

1. Introduction

The governance of non-profit organisations is a theme that has come to prominence recently, Birchall and Simmons (2004), Cornforth (2004), Spears (2004). This development follows the research in governance of market enterprises observed elsewhere, Berglof (1997), Brickley et al. (1997).

Effective corporate systems of governance are best-practice rules regarding the behavior of the Board of Directors in an organization, and are intended to address deficiencies in corporate governance due to the principal-agent relationship (Jensen and Meckling, 1976). Governance practice aims to align the principals' (shareholders) and the agent's (CEO) interests in order to maximize shareholder value. The CEO is the highest-ranking corporate officer in any company. He/She may also be the Chairman of the Board, or Company President in a small business, but these two roles are usually separated in larger companies in order to prevent the company from becoming dominated by a single personality and to prevent a conflict of interests with the stockholders. Governance practices are based on codes of governance, which are a set of norms governing the role and composition of the Board of Directors, relationships within it, auditing and information disclosure and the selection and dismissal of directors and senior managers (Aguilera and Cuervo-Cazurra, 2004). The fundamental concerns of corporate governance are twofold: (i) to ensure that conditions apply whereby a non-profit organization (NPO) directors and manager act in the interest of the organization and of its stockholders as well as its workers; (ii) to ensure that the means exist to hold managers accountable to stockholders and employees for the use of assets (Williamson, 2002).

Pay-performance contracts are settled with the aims to align the principal-agent relationship between managers and stockholders. Unfortunately, the standard accounting measures can be manipulated to boost executive compensation. Monitoring of the Board is therefore crucial. The monitoring of the Board depends on a number of factors, the most significant of which is the independence of the external directors from the CEO and internal board-members. If the CEO influences the director election process, the independence of the Board is compromised and CEO monitoring rendered ineffective, resulting in CEO entrenchment, and the opportunity to demand compensation in excess of the market equilibrium wage.

The present paper analyzes pay and performance of Portuguese NPO managers, aiming to identify rules of good practice in Board composition in this sector. The motivation for the present research stems from several issues. First, the sole code of governance in Portugal was established by the Stock Exchange. This procedure is contrary to observed codes of governance worldwide, which are drawn up by governments, directors' associations, managers' associations, professional bodies and investors and are usually overseen by autonomous watch-dogs (Aguilera and Cuervo-Cazurra, 2004) Hence, we can assert that Portugal lacks necessary codes of practice, there being no code whatever for non-profit organizations, which may affect negatively their efficiency. Moreover, only 61% of the quoted companies respect

partially the while it is safe to assume that none of them respect the rules completely (Diário Económico, 2005). Since there is a recognized positive relationship between best-governance practices and efficiency (Ames, 2003), this putative lack of governance principles is a cause of inefficiency in the non-profit activity.

Second, we seek to examine the consequences of the concentration of the powers of the Chairman and the CEO in the NPOs. The combination of the roles of Chairman and CEO in the hands of a single individual is commonly found in Portuguese NPOs, due to the fact that they are small organizations with only a single-tier Board, and that are not normally quoted on the Stock Exchange. Traditionally, one distinguishes between the one tier and two-tier Board system. The single-tier system does not only prevail in countries based on Common Law tradition. For instance, the UK Companies Act allows for a de facto two-tier structure, while France follows the tradition of the single tier (Rose, 2006). The concentration of power in the single-tier system is based on the closed corporate tradition of the country, in which the capital of quoted companies is usually controlled by internal regulations. Therefore, the corporate tradition runs counter to the interests of efficiency. Furthermore, it goes against the stipulation of the European Commission that stresses the importance of having a non-executive Chairman of the Board (Berglof, 1997). This stipulation is mirrored in the USA by the Sarbanes-Oxley Act that stresses the importance of governance at world level. Therefore, we believe that it is necessary to analyze the present situation in Portugal and to observe its evolution in the future.

Finally, if rules and behavior do not enhance efficiency, there must be visible characteristics in the market that identify these endogenous factors of competitive advantage in the non-profit sector; since they are endogenous, observers will be unable to identify them by observing the market alone.

The present paper aims to analyze the role of enterprise performance, board composition and individual characteristics defining the governance environment of the Portuguese NPOs. In so doing, it enlarges previous research in the field, since we are not aware of any studies on non-profit organizations that focus on governance issues. Moreover, we adopt an innovative econometric procedure in this context (Borsch, Supan and Koke, 2002).

The paper is organized as follows. In the second section, the contextual setting is described. In the third section, the literature survey is presented. In the fourth section, the theoretical framework is presented. In the fifth section the data is presented. In section sixth, the results are presented. In section seventh the results are discussed. Finally, in section eight, the results are presented.

2. Contextual setting

There is a broad spectrum of non-profit organizations in Portugal. It embraces the social solidarity charities called Misericórdias, which are the oldest such bodies in the country, dating back to the 15th century. In the present day, their main function is to provide social welfare to the elderly, in the form of retirement homes, day centers and meals. The sector also includes the foundations which are the wealthiest in premises type of NPO in Portugal, and which support cultural and scientific activities. Next, there are mutual associations, which complement the public social security system for the benefit of their members (Barros, 2001). In addition, we find associations and bodies

that exist to promote and pursue activities of solidarity and aid, for example, in developing countries. Finally, there is the cooperative movement (Barros and Santos, 2003; Barros, 2006), which is particularly prominent in the agricultural sector, for example, in wine production and distribution.

Table 1 presents some information on the Portuguese NPO sector. The ratio of members to organizations varies among the different types of body, with the cooperatives displaying the lower number of this ratio, signifying that this is a market-oriented type of organization, and that cooperatives consist of small numbers of members. The largest membership number is observed in the mutual associations.

3. Literature Survey

There are several strands observed in the literature on governance. A prominent line of research concentrates on the macroeconomic impact of governance by country (La Porta et al., 1997, 1998, 1999, 2000), concluding that the corporate governance systems are linked to the legal tradition of the country. Countries with a civil-law tradition, as opposed to the common law tradition (dominant in Anglo-Saxon countries) are less effective in protecting shareholders' rights (La Porta, 1998).

Another line of research, which is of greater relevance for the present paper, adopts a microeconomic approach, analyzing pay and performance and board composition. Works include Pennathur and Shellor (2002), who measure the determinants of CEO compensation in function of firm performance, measured by the stock returns, investment and funds from operations. Core et al. (1999) analyze the relationship between CEO compensation and board composition, concluding that the board composition explains a significant amount of cross-sectional variation in CEO compensation, after controlling for standard economic determinants of pay. Their result reveals that CEOs earn greater compensation when governance structures are less efficient. Gosh and Sirmans (2005) conclude that CEO compensation depends significantly on the usual economic measures of performance, including firm size and returns on assets, as well as on board composition, being higher in firms in which monitoring is weak, owing to large dimension and elderly directors. The existence of a blockholder is adverse.

Rosenstein and Wyatt (1990) analyze the role of external directors in the governance of the firm, concluding that the greater representation of external directors enhances firm performance. Byrd and Hickman (1992) find that announcement period of abnormal return in stock market is a positive function of the proportion of seats held by external directors. Yermack (1996) demonstrates that larger boards are detrimental to performance, although Bhagat and Black (1999) fail to confirm the results. Further corroboration, however, comes from Brickley et al. (1994). External directors enhance shareholder gains in tender offers. Nevertheless, Agarwal and Knoeber (1996) and Mishra and Nielsen (2000) fail to confirm the positive effect of external directors.

The relationship between board composition and performance, which is centered on the relationship between economic variables and CEO earnings, has thus far no been conclusively established by applied research.

In order to separate the competing hypotheses of the economic variables, degree of managerial entrenchment and omitted economic variables, Core et al (1999) analyzed

the impact of excessive CEO compensation implied by board variables against future performance. Gosh and Sirmans (2005) adopted this framework in their research.

4. Theoretical Framework

Williamson (2002) present a theoretical framework to analyze governance issues at firm level. Hermalim and Weisbrod (1998) provide a theoretical framework that relates pay and performance to the board composition. These two models are the theoretical basis for the present research. In the Hermalim and Weisbrod (1998) model, board effectiveness is a function of board independence. Four elements essentially explain this relationship: first, management turnover is more related to earnings than to stock returns; second, a board tends to be less independent the longer is a CEO's tenure; third, independent directors are added to the Board following poor corporate performance. Fourth, CEO turnover is negatively related to performance, this relationship being stronger when the board is more independent.

This relationship can be seen in Figure 1 below, which presents a situation in which the CEO's remuneration is set by the corporate Board according to a market demand and supply.

The CEO's pay may be set by the Board at an optimal level, if the observed board structures are conducive to the monitoring of the CEO. In this case, pay-performance contracts are optimal and reflect the economic determinants of performance. On other hand, if the director election process is influenced by the CEO, independence of the Board is compromised and CEO monitoring is rendered ineffective, resulting in CEO entrenchment and the opportunity for the CEO to demand compensation in excess of the equilibrium wage rate. Under this scenario, Board structure variables reflect ineffectiveness of the firm's governance structure, as well as the concomitant impact on CEO compensation. The pay will be non-optimal in view of the performance observed.

4.1 Hypotheses

Three sets of explanations for CEO compensation have been suggested in the literature: non-profit organization performance indicators, which capture accountancy determinants of CEO earnings (Ames, 2003); board compensation indicators in the industry (Gosh and Sirmans, 2005) and CEO characteristics. These hypotheses are derived from the theoretical model and from the explanation indicators.

Hypothesis 1 - Non-profit organizational performance indicators: In pay performance contracts, CEO compensation is an increasing function of performance measured by accounting numbers (Ames, 2003). Organisational non-profit operational insufficient performance may restricts the CEO's opportunity to increase his earnings. Under this hypothesis, CEO compensation is a function of non-profit performance. Core et al. (1999) and Gosh and Sirmans (2005) have tested this hypothesis.

Hypothesis 2 - Board compensation indicators: The lower the number of independent external directors, the less effective is the Board. Factors that diminish board power include older directors and the dual role of CEO and Chairman (Byrd and Hickman, 1992). Board members appointed by the current CEO are more loyal and less vigilant, and therefore are expected to be more generous in CEO compensation decisions.

Hypothesis 3 - CEO individual characteristics: The CEO's educational level has a positive impact on performance, which boosts the compensation. This hypothesis has not yet been tested in this context, because the data set used by the researchers is from the Stock Exchange. However, small companies such as NPOs are usually unquoted, and there is no Stock Exchange data in existence. Since we had to obtain data from a questionnaire, we decided to include some questions on individual characteristics of the CEO, allowing us to test this hypothesis.

5. Data

We carried out the empirical study in 2002 by means of the above-mentioned questionnaire. This was mailed to a random sample of Portuguese NPO administrators, stratified by organization type (Misericórdias, foundations, mutual associations, solidarity associations, other associations and cooperatives), with the central aim of determining their socio-economic characteristics. The sample was stratified by organization in order to be representative of all social organizations, using a Portuguese Government Ministry of Labour and Solidarity database of all NPOs. Because of budgetary restrictions and the limited time available, it was decided to collect data from only 1000 professional NPO administrators. The respondents returned only 648 completed questionnaires, all from professional administrators, which were retained in the analysis, giving a response rate of 64.8%.

The rate response does not differ significantly from the sample for age variable (chi-square = 7.32, $p = 0.05$), nor for gender (chi-square = 8.26, $p = 0.05$). Hence, the 648 administrators who completed the questionnaire might be representative of all Portuguese NPO administrators.

As to the respondents' general characteristics, they were male (98.4%), with a college degree, an average age of 42, and an average monthly income of 1.396,64 euros. This profile leads to an overall definition of the responding administrators as middle-aged, middle-class Portuguese males.

The questionnaire was delivered to a randomly-selected sample of NPO managers at their headquarters, after they had been approached first by phone and then personally. We pre-tested the questionnaire on students of economics at the Instituto Superior de Economia e Gestão, Technical University of Lisbon.

The administrators were asked to complete a standard questionnaire, including questions on socio-economic conditions, as well as other issues. We present in Table 2 the variables used in this paper.

5.1 Reliability, Validity and Generalizability

Several steps were taken to ensure the validity and reliability of the data. First, the point of departure was a questionnaire already tested in other economic fields (Barros and Gomes Santos, 2003), which was adapted for the present purpose, ensuring that prior research in the field is considered and face validity established. Second, it took into consideration all relevant literature. Third, the questionnaire was pre-tested on students of economics of the Technical University of Lisbon. Following the administration of the final survey, a stratified random subset of 50 respondents was contacted a second time to check if any problem persisted/the reliability of responses received, but no problems emerged. These procedures ensure the validity of the

questionnaire, meaning that it measures what it was intended to measure. Fourth, the questionnaire opted for a national random sample, with a response rate of 64.8%, which was considered an acceptable sample of respondents (Dillman, 1978). This procedure ensures the generalizability of the data, meaning that the findings are applicable to a more general population. Fifth, the reliability of the data was examined, analyzing it extensively with alternative methods and reaching the same conclusions (Barros and Santos, 2006).

The extensive examination of the survey's validity, reliability and generalizability leads to the inference that there is nothing in the evaluation to suggest that it is either invalid or unreliable.

5.2 Testing for Non-Response

The 64.8% response rate raises the question of non-response, for which we therefore adopted a testing procedure, based on Dillman (1978). As mentioned previously, a first test for this problem involved contacting the respondent again and suggesting testing the answers. The answers maintained the declared values, ensuring the accuracy of the responses. A second test involved contacting those who had not responded, in order to understand the reasons for their non-response. As a result, several explanations emerged. The first reason was the organization's declared secrecy policy, which is a common obstacle to questionnaires. The second reason was a lack of time available to complete the questionnaire. The third reason was saturation, associated with completing too many questionnaires. From these three reasons, it can be asserted that the non-responders have the same characteristics as those who did respond, establishing the representativeness of the questionnaires that were completed by NPO managers.

6. Results

The classic specification of Shleifer, A. and Vishny, R. (1997) for the determination of the pay-performance equation is:

Equation(1)

where $\log \text{TOTCOMP}_i$ is the log of total earnings of the CEO i , Economic are the variables describing the NPO's performance, Board are the variables describing the board composition, individual are the variables describing the individual characteristics of the CEO, ϵ_i is the statistical error term and β are parameters to be estimated. Because TOTCOMP is measured in logarithms, the parameter estimates from this regression can be interpreted as reflecting the proportional change in exogenous variables when the independent variables change.

This equation is similar to the traditional earnings equation (Mincer, 1974), and, since our data was obtained by questionnaire, this reference literature should be taken into account for the present research.

Our empirical strategy consists firstly of estimating the governance model with OLS. Traditionally, eq.(1) was estimated by OLS, but Core et al. (1999) and Gosh and Sirmans (2005) found that in the pay-performance model, the board composition is endogenous to the system and as such, an OLS model is subject to mis-specification. Therefore, estimated OLS parameters obtained by the traditional approach are not consistent, either because of measurement errors in the Board variables or because

the explanatory variables, such as board composition or individual education, are correlated with the unobserved disturbances in the equation. Alternative approaches to deal with this endogeneity problem fall into three categories: (1) controls for variables, which may be correlated with board composition and which determine earnings; (2) the use of instrumental variables (IV) which affect board composition, but do not directly affect earnings; (3) estimation of a system of equations for both endogenous variables, earnings and board composition with 2SLS, and Gosh and Sirmans (2005); and (4) the use of a GMM method that overcomes the problem.

This paper adopts the IV approach for the estimation of the traditional earnings equation, taking into consideration that according to the traditional literature, Board composition and individual characteristics (particularly education) are potentially endogenous and measured with error. The IV approach, which deals with instrumental variables, chooses instruments that are not correlated with the earnings residual (Card, 1993). This approach can be conceptually described as being carried out in two stages: first, by estimating the effect of the instrumental variable on schooling and on Board composition; then by estimating the effect of schooling and Board composition on earnings. By assumption, the instrument is correlated with earnings because it influences schooling and Board composition. The instrument chosen in empirical work may not be truly independent of earning residuals.

In the IV equation, our instruments are all the variables used in the OLS model and two additional instrumental variables to account for Board composition and education. Since the quality of the proxy variables is important (Madalla, 1992), we looked to previous studies on earnings in relation to Board composition and education for guidance as to the variables for which we opted, as well as the availability of the data. For board composition, we chose the number of stockholders in the company, assuming this to have an exogenous influence on earnings. The higher the number of stockholders, the more dispersed is the capital and the higher the discretionary power of the CEO in establishing the Board.

For education, we chose the father's education, assuming this to have had an exogenous influence on the respondent's schooling, and the number of children, assuming this to have an exogenous influence on the respondent's age, since older individuals are more likely to have children.

Table 3 above shows the results for the traditional CEO earnings function in which the logarithm of the net salary plus bonus serves as a dependent variable, which is regressed on the standard controls. We find that the regression produces the theoretically expected, statistically significant results on the variables. Controls first for board dimension and finally for individual characteristics produce a small alteration in the value of the estimates and maintain the statistical significance. The OLS model, estimated for comparative purposes, is in line with the IV model, but with less explanatory power.

We considered some test results proposed by Bound, Jaegger and Baker (1995) to shed light on the quality and validity of our instruments. Instrumental quality is ensured if there is a strong correlation between the instruments and Board composition and schooling. It is well documented that an IV procedure using weak instruments may yield more inconsistent point estimates than those produced by OLS. First, the F-test of

joint significance for the respective instrument set is equal to 0.53. Secondly, instruments are valid provided they affect earnings through schooling only. We tested whether our instruments have any direct influence on earnings through schooling only with the Sargan test, a test of over-identification restrictions on the instruments, with an asymptotic χ^2 distribution and degrees of freedom equal to the number of over-identifying restrictions. With a value of 4.31 (p -value=0.53) we cannot reject the validity of over-identifying restrictions.

The models fit the cross-data well with the extended model, presenting an adjusted R square ranging from 20% to 65% and an F-statistic ranging from 4.34 to 10.12, denoting that the variables jointly explain the model.

Assuming that the lack of independence appears to be the rule rather than the exception in survey data, we conducted the Breusch-Pagan test and White's test of heteroscedasticity. The White test in the Mincerian model is 20.51 and the 95% critical value is 7.80, so the hypothesis of homoscedasticity is not rejected by the test. For the extended model, the White test is 32.12 and the 95% critical value is 15.2, so the hypothesis of homoscedasticity is not rejected either. The hypothesis of homoscedasticity of the residuals is not rejected by the test.

The Breusch-Pagan test is 6.54 for the Mincerian model and the 95% critical value is 7.52, thus the hypothesis of homoscedasticity is not rejected. The same conclusion applies to the extended model with a Breusch-Pagan test of 7.43 and the 95% critical value of 15.2.

According to the extended model, the earnings are a function of the following: (1) accountant variables. (2) Board composition variables. (3) individual characteristics, and (4) organizational identification.

6. Policy Implications of the Study

Our findings point to a significant positive association between earnings, accountancy performance variables, board composition and individual characteristics. Even after controlling for individual characteristics, we were able to confirm the earnings-board composition relation for this particular labour market.

What are the implications of these results? Firstly, the Portuguese non-profit performance accountancy variables such as ROA (Rosen, 1990) affect the CEO's earnings. The semi-elasticity of Portuguese CEOs' annual cash compensation with respect to ROA is between 1.187 and 1.013, which is in line with results obtained elsewhere (Kato and Kubo, 2006; Rosen, 1990). Other accountancy variables are also statistically important, such as sales, but whereas ROA and sales affect earnings positively, users has a negative effect, which validates Hypothesis 1.

Second, the relationship between earnings and board composition signifies that there are principal-agent problems active in the Portuguese NPOs, which validates Hypothesis 2. Therefore, a managerial policy to align the principal-agent relationships is needed. The CEOs' earnings increase with board size, number of intern directors on the Board, older managers in the Board, politically-affiliated managers, when the CEO is also the Chairman, the number of friends on the Board and finally, with the duration of the CEO on the Board. Thus, it is concluded that Board composition influences

positively the CEO's earnings. This result is in line with what has been observed elsewhere (Sierra, Talmor and Wallace, 2006).

Furthermore, individual characteristics affect CEO earnings, namely, the number of years of education, to have a performance-related contract and the CEO's age, which validates Hypothesis 3, (Barros and Santos, 2005).

Finally, all types of organization have a positive relationship with the endogenous variable, which is statistically significant for almost all of them, signifying that governance problems exist in all of them.

The general conclusion is that the CEOs' earnings in the NPO sector are defined by multiple factors, some of them being individual attributes defined by the market, such as individual characteristics, while others are defined by the performance of the NPO. However, board composition also affects the CEO's earnings positively, highlighting the existence of principal-agent problems in this market.

What is the predominant implication of the present research? The main conclusion is that adoption of governance principles related to board composition is urgently needed in the Portuguese NPO sector.

7. Conclusion

The general conclusion of the present paper is that governance is obviously important in the Portuguese non-profit sector. Several variables affect the CEO, namely, organization performance variables, board composition variables and individual variables. The present result sheds light on the governance problems in the Portuguese NPOs and highlights the urgent need for a code of governance practice to be introduced in this sector. This should prevent the organizations from becoming dominated by a single personality, in addition to preventing conflicts of interests with the stakeholders. Moreover, the paper integrates human capital in the definition of the earnings, concluding that this also contributes to earnings. A relationship is thus established between human capital models (Barros and Santos, 2003) and pay-performance governance models in NPOs.

However, the final conclusion is that the situation is more complex than our conceptualization suggests, therefore further research is needed to confirm these results.