

**Metadiscourse in Academic Writing: A Corpus-based
Study of Expert, L1 and L2 Postgraduate Student Text**

Seowon Lee

NEWCASTLE UNIVERSITY LIBRARY

207 32715 4

Thesis L9056

**Thesis submitted for the degree of Doctor of Philosophy
School of Education, Communication, and Language Sciences
Newcastle University UK**

April 2009

I certify that all material in this thesis which is not my own work, has been identified and that no material in included which has been submitted for any other ward or qualification.

Signed:

A handwritten signature consisting of a stylized, cursive letter 'L' with a horizontal stroke extending to the left.

Dated:

3 / 4 / 09

Acknowledgments

The completion of this thesis would not have been possible without the assistance and encouragement I have received from many people,

First, my supervisor Vivian Cook who has been very supportive and helpful throughout the composing of this document. I am grateful for his feedback, support in every phase of research designing and thesis writing. Without his patience, guidance, availability and devotion it would have been a 'mission impossible'. My deepest gratitude continues to my second supervisor Dr. Tony Young, who was supportive in the data collection and sincere appreciation also goes to Simon Kometa for his generosity in providing consultancy for statistical analysis, Dr. Jan Hardman and Mr. Scott Windeatt for their valuable comments and assistances on earlier versions of this work as a supervisor and a panel chair.

Second, Dr. Paul Seedhouse, the director of the doctoral programme, for his valuable guidance from the first day in this programme, and continuous assistance during the stages of the integrated PhD programme.

Third, my gratitude goes to my husband, Dr. Il-Hyun Yoon, for being a great inspiration, a strong support, and for his tremendous help with Excel and statistical analysis throughout this long and arduous process of completing the thesis in the UK.

Finally, I would like to express my deepest gratitude for the constant support, understanding and love that I received from my two children, Yongseo and Youngjun during the past years.

Thanks also, are extended to Dr. Chris Letts, who helped me with technical support in data collection and analysis with WordSmith Tools for this thesis.

Dedication

Though I fear that this work is too meagre to warrant it, I dedicate this thesis to my parents, without whom I would not be who I am today, and without whom, certainly, any successful aspects of this research could never have been written.

Abstract

This study looks at how differences of language and genre affect the pattern of the usage of metadiscourse (MD: the rhetorical resources used to organise a discourse, or the writer's point of view towards either its content or the reader) in the academic writings produced by expert and student writers in the same discipline.

The corpora consist of thirty journal articles (PRO) and fifty-five student assignments, twenty-five from native English speakers (NES) and another thirty from non-native English speakers (NNES). All texts in both the PRO and student corpora are in English, produced by a single-author, in topics of study in language, culture, and communications in the same school of Newcastle University. The research uses a typology derived from those in the literature but focussing on written texts and modified by a pilot study.

From the comparison of texts produced by the NES and NNES postgraduate students, the language variable (native English vs non-native English) plays a greater role in the use of MD. The NNES writers use statistically more textual metadiscourse (TMD) while the NES writers employ more interpersonal metadiscourse (IMD) in general and there are statistical differences in the use of sequencers, code glosses which were more used by the NNES, hedges and self references which were more employed by the NES in particular.

The finding indicates that the NNES writers are more concerned about expressions to show the logical order and relations between different parts of the text through TMD; the NES writers try more to involve the readers in the argument than do their NNES counterparts with IMD. The findings also show that learning a writing style which is acceptable in western academic life (e.g. 'writer-responsible') influences the use of MD in the NNES academic writing. Evidence of this comes from the interview data and the results of text analyses which show the statistically greater use of textual metadiscourse (e.g. sequencers, code glosses) and the significantly infrequent use of self-references in the NNES texts.

From the results in the genre/expertness variable (journal articles vs student assignments), no overall significant differences were found in the use of the main

categories (TMD and IMD), but differential purposes (effects) and frequency were found in the use of MD subcategories. The student writers do not use MD devices in the same way as the PRO writers, as reflected in the use of MD devices with a limited range of items and purposes in the student texts compared to a broad range of MD features and functions in the PRO texts. In fact, the PRO writers made more use of concessives, concluders, sources, hedges and self references with a broad range of purposes; the student writers made significantly more use of sequencers, topicalisers and more use of emphatics with limited purposes.

Thus the finding proposes that the way they use MD is influenced by the two factors in the student and PRO texts; the consideration about the readership and the goal of the argument; which lead the different pattern of MD usage in the student and expert writings. This suggests that the genre variable (student assignments vs journal articles) is also a crucial one to influence the use of MD within the same discipline.

As regards the language aspect from the comparison between the NES and NNES, the differences are mainly in the amount of features in the use of MD. When it comes to the student and journal article texts, genre variable, the differences are not only in the frequency of MD subcategories but also in the way they use the MD features.

Abbreviations

AL	Applied Linguistics
AMD	Attitudinal Metadiscourse
CARS	Create A Research Space
CBA	Corpus-Based Approach
CCC	Cross Cultural Communication
CR	Contrastive Rhetoric
DA	Discourse Analysis
DM	Discourse Markers
EAP	English for Academic Purposes
ESL	English as a Second Language
ESP	English for Special Purposes or English for Specific Purposes
H.E.S.A.	Higher Education Statistics Agency in the United Kingdom
IMD	Interpersonal Metadiscourse
L1	First Language or Mother Tongue
L2	Second Language
MA	Master of Arts
MD	Metadiscourse
MWU	Mann Whitney U test
NES	Native English Speaker
NNES	Non-Native English Speaker
POSTCORP	Postgraduate Student Corpus
PRO	Professional writers
RAs	Research Articles
RACORP	Research Article Corpus
TESOL	Teaching English to Speakers of Other Languages
TMD	Textual Metadiscourse

Table of Contents:

Chapter 1: Introduction.....	1
1.1 Introduction.....	1
1.2 Background to MD.....	2
1.3 Purpose of the Study	4
1.4 Significance of the Study	5
1.5 Methodology	6
1.6 Organisation of the Study	8
Chapter 2: Metadiscourse (MD) and Rhetoric.....	9
2.1 Introduction.....	9
2.2 Background to Rhetoric	10
2.2.1 Classical Rhetoric	10
2.2.2 Components of Classical Rhetoric.....	11
2.2.3 Modern Rhetoric.....	11
2.3 Metadiscourse and Rhetoric.....	13
2.4 Writing Style and Culture	15
2.4.1 Individualism and Collectivism.....	16
2.4.2 High Context and Low Context.....	18
2.4.3 Reader or Writer Responsibility?	19
2.5 Language, Thought and Culture.....	20
2.6 Contrastive Rhetoric (CR)	21
2.7 Summary	28
Chapter 3: A Working Definition of Metadiscourse	29
3.1 Introduction.....	29
3.2 The Concept of Metadiscourse	30
3.3 Characteristics of MD	33
3.3.1 Common Factors in MD	33
3.3.2 Contribution to Authorship.....	33
3.3.3 The Multifunctional Nature of MD	34
3.3.4 The Social Characteristics of Metadiscourse.....	36
3.4 Definitions of Metadiscourse and Parallel Concepts to Metadiscourse.....	36
3.4.1 Metadiscourse	36
3.4.2 Parallel Terms to MD	44
3.5 A Proposed Working Definition of Metadiscourse.....	49
3.6 Summary	51
Chapter 4: A Review of MD Related Studies and a Proposed Typology.....	53
4.1 Introduction.....	53
4.2 Functions of MD in Written Discourse.....	54
4.2.1 Functions of Textual MD (TMD).....	54
4.2.2 Functions of Interpersonal MD (IMD)	55
4.3 A Critical Review of Metadiscourse Typologies.....	55
4.3.1 Joseph Williams (1981).....	55
4.3.2 Typologies Used by Avon Crismore	56
4.3.3 Vande Kopple (1985)	62
4.3.4 A Typology of Metadiscourse Based on Classical Rhetoric (Nash 1992)	66
4.3.5 Paul Jude Bauvais (1986)	68

4.3.6	Typologies of Hyland (1998) and Hyland and Tse (2004).....	73
4.4	A Critical Review of Typologies of Parallel Terms Other than MD	79
4.4.1	Typologies Used by Meyer.....	79
4.4.2	A Typology of Non-topical Materials (Lautamatti 1978)	82
4.4.3	A Typology of Gambits (Keller 1979)	83
4.4.4	A Typology of Metatalk (Schiffrin 1980)	86
4.4.5	Typologies of Discourse Markers.....	87
4.4.6	A Typology of Metatext (Mauranen 1993a).....	91
4.5	The Proposed Typology of Metadiscourse	93
4.5.1	The Proposed Typology of TMD	96
4.5.2	The Proposed Typology of IMD.....	100
4.5.3	Feasibility of the Typology.....	104
4.6	Applied MD Studies.....	111
4.6.1	Genre Related Studies of MD.....	112
4.6.2	Language Related Studies of MD	118
4.7	Summary	122
Chapter 5:	Methodology and Construction of the Corpora	124
5.1	Introduction.....	124
5.2	Research Questions	124
5.3	Research Methodology	126
5.3.1	Pilot Study and the Methodologies.....	126
5.3.2	The Characteristics of the Methods adopted in the Main Study.....	127
5.4	Data Collection Methods	131
5.4.1	Pedagogical Applicability.....	131
5.4.2	The Expert Corpus.....	132
5.4.3	The Student Corpus	135
5.5	Building the Corpora.....	137
5.5.1	Expert Corpus	137
5.5.2	Student Corpus.....	137
5.5.3	The Expert (RACORP) and Student (POSCORP) Corpora	138
5.6.	Searching the Corpora.....	139
5.7.	Data Analysis Methods	141
5.7.1	Counting the MD Frequency	142
5.7.2	Statistical Analysis.....	142
5.8.	Methodological Issues.....	145
5.8.1	Ethical Considerations	145
5.8.2	Validity	146
5.8.3	Reliability	147
5.8.4	Limitations of the Method	147
5.9	Summary	154
Chapter 6:	Results.....	156
6.1	Introduction.....	156
6.2	The Quantity and Density of MD in the Whole Corpus	157
6.3	General Overview:	159
6.3.1	MD Use in the Three Group Texts	159
6.3.2	MD Use in the NES and NNES texts	163
6.3.3	MD Use in the PRO and Student texts	170
6.4	Results from Individual Subcategories in TMD	177
6.4.1	Sequencers	180
6.4.2	Topicalisers.....	184

6.4.3	Concluders	188
6.4.4	Pre/reviews	191
6.4.5	Additives	194
6.4.6	Concessives.....	198
6.4.7	Code Glosses	201
6.4.8	Rationales	205
6.4.9	Sources.....	208
6.5	Results from Individual Subcategories in IMD	212
6.5.1	Hedges	214
6.5.2	Emphatics	218
6.5.3	Evaluatives.....	223
6.5.4	Commentaries	226
6.5.5	Appeals	229
6.5.6	Self-References.....	231
6.6	Results from the Interviews	237
6.7	Summary	245
Chapter 7:	Discussion	246
7.1	Introduction.....	246
7.2	A Summary of the Answers to Research Question 1	248
7.3	Language Aspect (NES vs NNES) for Research Question 1.....	249
7.4	A Summary of the Answers to Research Question 2.....	257
7.5	Genre Aspect (Assignments versus Journal Articles) for Research Question 2 260	
7.6	A Summary of the Answers to Research Question 3.....	269
Chapter 8:	Conclusion	274
8.1	Introduction.....	274
8.2	Theoretical Contributions	275
8.3	Empirical Contributions.....	276
8.3.1	From the Language Difference	277
8.3.2	From the Genre Difference	279
8.4	Implications.....	280
8.4.1	Implications for Teaching.....	280
8.4.2	Implications for the Development of Materials.....	282
8.5	Limitations and Suggestions for Further Research	283
8.6	Concluding Remarks.....	284
References:	286
Appendices:	300

List of Tables:

Table 3- 1: Elements Mentioned in the Different Definitions of MD	43
Table 3- 2: A Proposed Synthesis of the Elements in MD Definitions	50
Table 4- 1: Hyland’s Classification of Metadiscourse in Academic Texts	75
Table 4- 2: A Summary of MD Subcategories used in the Hyland(1998) and Vande Kopple (1985).....	76
Table 4- 3: A Summary of TMD Subcategories used in the Reviewed Typologies	77
Table 4- 4: A Summary of IMD Subcategories used in the Reviewed Typologies	78
Table 4- 5: A Proposed Synthesis of Main Categories and Subcategories of MD:.....	94
Table 4- 6: A Proposed Typology of TMD	96
Table 4- 7: A Proposed Typology of IMD	100
Table 4- 8: Subcategories of TMD Used by Other Researchers under the Name of MD	105
Table 4- 9: Subcategories of IMD Used by Other Researchers under the Name of MD	106
Table 4- 10: Subcategories of TMD used by Other Researchers under the Parallel Terms other than MD.....	108
Table 4- 11: Subcategories of IMD Used by Other Researchers under the Parallel Terms other than MD.....	109
Table 5- 1: A List of MD Subcategories used in the Pilot Study	126
Table 5- 2: The Top 8 Popular Disciplines with International Students	131
Table 5- 3: Details of the RACORP and POSCORP.....	138
Table 6- 1: The Size of the Corpora	160
Table 6- 2: Raw Frequency of TMD in the Three Groups	160
Table 6- 3: Raw Frequency of IMD in the Three Groups	160
Table 6- 4: Mean Frequency for TMD per 5000 words in the Three Groups	161
Table 6- 5: Mean Frequency for IMD per 5000 words in the Three Groups	162
Table 6- 6: Mean Frequency of Total MD, TMD and IMD per 5000 words in the Three Groups.....	162
Table 6- 7: The Size of Data in the NES and NNES	164
Table 6- 8: Mean Frequencies of total TMD and TMD Percentages out of MDs per 5000 words in the NES and NNES texts	164
Table 6- 9: Mean Ranks for TMD per 5000 words in the NES and NNES	164
Table 6- 10: Mean Frequencies of TMD subcategories per 5000 words in the NES and NNES texts	165
Table 6- 11: Percentages of TMD Subcategories per 5000 words out of total TMDs in the NES and NNES texts	165
Table 6- 12: Mean Ranks for the Frequency of TMD Subcategories per 5000 words in the NES and NNES texts	166
Table 6- 13: MWU Tests for TMD Sub-category in the NES and NNES	167
Table 6- 14: Mean Frequencies of total IMD and IMD Percentages out of MDs per 5000 words in the NES and NNES texts	167
Table 6- 15: Mean Frequencies of IMD Subcategories per 5000 words in the NES and NNES texts	167
Table 6- 16: Percentages of IMD Subcategories per 5000 words out of total IMDs in the NES and NNES texts	168
Table 6- 17: Mean Ranks for IMD Subcategories per 5000 words in the NES and NNES texts	168

Table 6- 18: MWU Tests for IMD Sub-category in the NES and NNES.....	169
Table 6- 19: The Size of Data in the Student and PRO Texts	170
Table 6- 20: Mean Frequencies of total TMD and TMD Percentages out of MDs per 5000 words in the Student and PRO texts	170
Table 6- 21: Mean Frequencies of TMD Subcategories per 5000 words in the Student and PRO texts	171
Table 6- 22: Percentages of TMD Subcategories per 5000 words out of total TMDs in the Student and PRO texts	171
Table 6- 23: Mean Ranks for TMD Subcategories per 5000 words between the Student and PRO texts	172
Table 6- 24: MWU Tests for TMD Sub-category in the Student and PRO	173
Table 6- 25: Mean Frequencies of total IMD and IMD Percentages out of MDs per 5000 words in the Student and PRO texts	173
Table 6- 26: Mean Frequencies of IMD Subcategories per 5000 words in the Student and PRO texts	174
Table 6- 27: Percentages of IMD Subcategories out of Total IMDs in the Student and PRO texts.....	174
Table 6- 28: Mean Ranks for IMD Subcategories per 5000 words in the Student and PRO Texts.....	175
Table 6- 29: MWU Tests for IMD Sub-category in the Student and PRO.....	175
Table 6- 30: Mean Frequencies of TMD Subcategories per 5000 words with the indication of the Statistical Difference in the NES and NNES texts.....	178
Table 6- 31: Percentages of TMD Subcategories out of total TMDs in the NES and NNES texts	178
Table 6- 32: Mean Frequencies of TMD Subcategories per 5000 words with the indication of the Statistical Difference in the Student and PRO texts.....	179
Table 6- 33: Percentages of TMD Subcategories out of total TMDs in the Student and PRO texts	179
Table 6- 34: Mean Frequencies of IMD Subcategories per 5000 words with the Indication of the Statistical Difference in the NES and NNES texts.....	213
Table 6- 35: Percentages of IMD Subcategories out of total IMDs in the NES and NNES texts	213
Table 6- 36: Mean Frequencies of IMD Subcategories per 5000 words with the Indication of the Statistical Difference in the Student and PRO texts.....	214
Table 6- 37: Percentages of IMD Subcategories out of total IMDs in the Student and PRO texts	214
Table 6- 38: The Summary of the Statistical Differences from the Comparison between the NES and NNES Texts.....	245
Table 6- 39: The Summary of Statistical Differences from the Comparison between the Student and PRO texts	245
Table 7- 1: Research Questions and Variables	246

List of Figures:

Figure 6- 1: Mean Ranks for the Frequency of TMD subcategories per 5000 words in NES and NNES texts	166
Figure 6- 2: Mean Ranks for IMD Subcategories per 5000 words in the NES and NNES texts.....	169
Figure 6- 3: Mean Ranks for TMD Subcategories per 5000 words between the Student and PRO texts	172
Figure 6- 4: Mean Ranks for IMD Subcategories per 5000 words in the Student and PRO texts.....	175

List of Appendices:

Appendix 5- 1: Results from the Pilot Study.....	300
Appendix 5- 2: Questionnaire for Semi-Structured Interview	301
Appendix 5- 3: A List of Journal Articles used in the Professional Corpus (PRO)	303
Appendix 5- 4: A List of the Assignments used in the NES Corpus.....	305
Appendix 5- 5: A List of Assignments used in the NNES Corpus:.....	307
Appendix 5- 6: Samples of Data Analysis for the Establishment of the Typology of MD	309
Appendix 5- 7: List of Potential MDs to be Searched for in the Corpora	316
Appendix 5- 8: Samples of Counting the MD Frequency for ‘Below’ in the Subcategory of Pre/Reviews.....	318
Appendix 5- 9: Parametric versus Nonparametric Tests	327
Appendix 5- 10: Shapiro-Wilk Tests:.....	331
Appendix 5- 11: Informed Consent Form	333
Appendix 5- 12: Ethical Issues from British Association for Applied Linguistics	334
Appendix 6- 1: Raw Frequencies of TMD Subcategories in the Data	338
Appendix 6- 2: Frequency per 5000 words for TMD.....	340
Appendix 6- 3: Raw Frequencies of IMD in the data.....	342
Appendix 6- 4: Frequency per 5000 words for IMD	344
Appendix 6- 5: Raw Frequencies of Sub-categories in TMD	346
Appendix 6- 6: Raw Frequencies of Subcategories in IMD.....	364
Appendix 6-5- 1: Sequencers:	346
Appendix 6-5- 2: Topicalisers	348
Appendix 6-5- 3: Concluders.....	350
Appendix 6-5- 4: Previews/reviews.....	352
Appendix 6-5- 5: Additives	354
Appendix 6-5- 6: Concessives	356
Appendix 6-5- 7: Code Glosses	358
Appendix 6-5- 8: Rationales	360
Appendix 6-5- 9: Sources	362
Appendix 6-6- 1: Hedges.....	364
Appendix 6-6- 2: Emphatics	367
Appendix 6-6- 3: Evaluatives	371
Appendix 6-6- 4: Commentaries	375
Appendix 6-6- 5: Appeals.....	378
Appendix 6-6- 6: Self-references.....	381

Chapter 1: Introduction

1.1 Introduction

This thesis is about metadiscourse (MD) in written texts produced by native English speaking (NES) and non-native English speaking (NNES) postgraduate student writers and professional writers (PRO) from journals in the areas of Applied Linguistics and Cross Cultural Communication studies, both of which concern language, communication and culture.

Most texts that consist of more than one paragraph involve two types of discourse: a primary discourse and a secondary discourse. The primary discourse carries the core content or message that the writer intends to convey to his or her readers. The secondary discourse, metadiscourse as it will be called from now on, refers to the orientation expressions that writers use to help readers understand the core message and the writer's stance towards either the content or the reader. It has been defined as discourse about discourse, a non-content aspect of discourse that includes textual and interpersonal features of discourse or what will be called in the study, textual metadiscourse (TMD) and interpersonal metadiscourse (IMD).

1.2 Background to MD

The concept of MD is relatively new and the available literature is rather limited. At the theoretical level there are, for example, Williams (1981) and Vande Kopple (1985). These works deal with the concept of MD and provide valuable suggestions for the main categories and sub-categories that can be used for the identification and classification of MD. Several typologies have been proposed: Williams (1981), Crismore (1982; 1989), Vande Kopple (1985), Intaraprawat (1988), Crismore et al. (1993), Stainton (1993), and Hyland and Tse (2004).

Several studies have discussed the positive effects of the presence of MD in texts. With reference to Halliday's (1973 and 1994) macrofunctional theory of language, on the *interpersonal* level, Schiffrin (1980, p231) and Crismore (1989) both point out that MD allows written texts to take on some features of spoken language (e.g., personal pronouns to establish an "I-you" relationship), and thus become more "reader-friendly". On the *textual* level, Crismore & Farnsworth (1990) and Crismore (1989) note that the discourse structuring functions of MD guide readers through a text and help them to organize the content as they read, thus fostering global comprehension.

Other benefits of MD derive from its use of explanatory and persuasive elements which attest to its key rhetorical function (Crismore 1989; Hyland 1999a; Hyland 2000). In fact, writers use these devices to produce a desired effect, depending on their underlying purposes and perception of readers' expectations. For example, in expert to non-expert communication (e.g., textbooks) MD helps to present information in a clear, convincing and interesting way in an effort to promote acceptance and understanding, as well as reader-writer solidarity. It is also an important persuasive resource used to

influence readers' reactions to texts according to the values and established conventions of a given discourse community.

Descriptions of metadiscourse have shown it to feature prominently in various types of academic discourse. These include school textbooks (Crismore, 1989; Crismore and Farnsworth 1990), university textbooks (Hyland, 2000; Bondi, 1999) and doctoral dissertations (Bunton, 1999; Hyland 2004a). It has also been studied comparatively in order to understand differences across cultures (Mauranen, 1993a; Valero-Garces, 1996; Alkaff 2000; Crismore et al, 1993). Some work has focused on MD in student writing. Intaraprawat & Steffensen (1995) analysed ESL university students' essays and concluded that good writers used a greater variety of MD than poor writers. Steffensen & Cheng (1996) conducted an experiment to investigate the effect of targeted instruction in MD on the writing abilities of native-speaker university students. An experimental group that had been taught the form, function and purpose of MD learned to use it effectively and produced compositions that earned significantly higher scores than those of a control group, which had received no instruction in MD.

Mauranen (1993a) found that native English speakers used more MD than non-native speakers (Finnish writers). Another suggestion by Crismore et al. (1993) and Alkaff (2000) was that the difference in the use of MD between NES and NNES is due to cultural and language background. It has been shown that, in general, native English-speaking writers used more MD than non-native-speaking ones. MD is then affected not only by language background but also by the factor of nativeness. These are the language differences as a 'language variable' in the present study.

Considerable differences have been found in MD use between students' and journal article texts from the comparisons of different genres (Harwood 2005, Hewings and

Hewings 2002, Hyland 2002a and 2002c). This is the ‘genre variable’ in the present study.

1.3 Purpose of the Study

Metadiscourse is recognised as an important means of facilitating communication, and is important at advanced levels of academic writing as it represents writers’ attempts to present and comment on the propositional information (Mauranen 1993b) that can be manipulated by either language variable or genre aspect in a particular disciplinary community. Though ‘language’ and ‘genre’ variables can be important factors which affect the pattern of MD usage in the same discipline, little is known about how they influence the use of MD in the academic discourse in which NES and NNES postgraduate student writers participate. This study seeks to address this gap.

This research gap suggested analysing the texts produced by the NES and NNES postgraduate students and journal article writers in order to find out how language or genre factors influence the pattern of MD in academic discourse and thus led to the thesis presented here. For instance, what are the types and functions of MD used in the NNES and NES students’ writing? Are some types of MD more frequent in the student corpora (NES + NNES) than in the expert corpus, or vice versa? If there is a difference in MD use between the groups, how differently are the features of MD used in the corpora and why? In other words, does the difference stem from language difference or genre difference?

The aim of the study is thus to investigate the differences and similarities of writing style from academic texts produced by native English speakers (NES), non-native

English speakers (NNES), and professional (PRO) writers, particularly as it relates to the use of MD features with special consideration to the language and genre aspects.

1.4 Significance of the Study

Considerable research has shown the differences in using MD features between the NES and the NNES (Crismore et al. 1993; Intaraprawat 1988; Mauranen 1993a; Valero-Graces, 1996; Alkaff 2000). A number of pieces of research have been investigated in the use of MD features in NNES student and professional texts, and in NES student writing and professional academic writing. For instance, dissertations of undergraduate NNES and texts produced by professionals (Hyland 2002a), dissertations by postgraduate NNES (Hewings and Hewings 2002; Hyland 2005a), and dissertations of native English speakers (Harwood 2005) were compared with a large corpus of research articles to find the differences in the use of MD between the corpora. However, at the same time, the comparison of the features of MD has been neglected between the NES and NNES students' academic writing and postgraduate students' (NES + NNES) and professionals' academic writing, within the same discipline, to find out how the variables affect the differences in the use of MD. The factors concerning the language or genre aspect have not been studied simultaneously to find out the effects of the difference of language and genre in the use of MD in academic writing.

The present study gains its significance from several factors. As far as I am aware, this is the first attempt to compare MD use in three different groups' texts within the same discipline to find the effect of factors governing MD usage between different groups' texts as an attempt to close the research gap. The study will reveal the differences/similarities in the use of MD between groups with a special focus on

frequency and purpose of the MD usage as a reflection of underlying rhetorical strategies in their academic writing.

In addition to this, the study employs a computer-assisted corpus-based study (using WordSmith 4.0) and statistical analyses to determine whether the differences between the two groups are meaningful, which introduce a more rigorous and empirical approach to MD studies in written English. So this will hopefully be seen as a contribution with positive impacts for current and future MD researchers and users.

Since this research is involved with MD use in the written texts produced by three different groups of writers, it is anticipated that the findings will constitute a potentially fruitful extension of existing English language teaching of writing in the academic area. In other words, teachers and the people who are involved in the fields of teaching and designing materials for an academic writing style to both native English speakers and non-native English speakers can use the findings to improve the effectiveness of teaching a writing style in English. Also, the findings will hopefully play a part in improving the teaching of writing style in general and helping students to develop their academic writing with an awareness of the effects of MD use in different group texts in particular.

1.5 Methodology

The present study combines the methodological advantages of double contrastive approaches with language and genre aspects (comparisons of writings from two different variables) to analyse texts produced by NES and NNES writers and student and professional writers. The detailed comparisons of academic texts written by NES and NNES students provide a means of identifying the similarities and differences in

the use of MD features. This is for investigation of the effect of the language difference in the use of MD in academic writings. And it is hoped that the comparisons of texts produced by the journal article and postgraduate student writers will give an answer to the question of whether the genre difference affects the use of MD in their academic writing in the same discipline.

The textual analysis of the scripts in this study is structured around the objective features of MD in each corpus. Specifically, the text analysis in this study identifies and compares the frequencies of use of textual MD (TMD) and interpersonal MD (IMD) features in 85 texts from each group (NES, NNES and PRO) to find out differences and similarities between groups, using the list of MD expressions in the proposed typology (which will be explained in chapter 4).

The procedure of this study has been two-fold: first, to analyse published papers in the topics of Cross Cultural Communication studies and Applied Linguistics, exposed to both non-native and native English speaking postgraduate students, and to discover the specific divergences in the use of MD features. The features in the journal article texts will be compared with those in the student texts according to the proposed typology in the study.

The second part of the research, after the analysis of the published papers, involves the study of MD in student writing, MD use by postgraduate students (native and non-native English speakers) in the same discipline as the published papers will also be analysed and compared to see if there are significant differences in number, type, and purposes of MD features as their rhetorical strategies.

1.6 Organisation of the Study

The study is structured in eight chapters, of which this introduction is the first. The second and third chapters contain the literature review that provides the conceptual and theoretical framework that guided the study. Chapter 2 deals with the connection between MD and rhetoric in terms of writing. Also this chapter elaborates the relationship between the writing style and culture-related factors which might affect the writing style in general and the usages of MD in particular. Next, chapter 3 establishes basic concepts and a working definition of MD as a result of reviewing related literature of MD in terms of academic writing, in order to understand the proposed typology of MD in the present study. Then the fourth chapter offers a critical review of the MD related study to establish the proposed typology which links to the working definition of MD in chapter 3, and is followed by the discussion of applied MD studies which explains the differences of MD usage from language and genre aspects in the studies of MD. The fifth chapter explains the methodology and provides the data collection methods (the procedure of collection and building corpora) with characteristics of each group corpus. The results from text analyses and interviews are described in chapter six. After that, all the research questions are answered in chapter 7 with a focus on the variables, language and genre respectively. The final chapter discusses the contributions of the study, together with several related features such as the implications of this study, the limitations of the study and some suggestions for further research.

Chapter 2: Metadiscourse (MD) and Rhetoric

2.1 Introduction

This chapter is especially devoted to the background related to the basic concepts and relationship between metadiscourse (MD), the topic of this thesis, and rhetoric (i.e. the art/technique of persuasion through the spoken or written language). MD and rhetoric cannot be dissociated from each other since both involve the making of different linguistic and non-linguistic choices, recognise the importance of audience and aim at achieving effective communication. Rhetorical choices vary from culture and become distinct across languages.

Within the context of this thesis, it is possible to place the use of MD in a culture perspective since the present data was collected from a different group of writers (e.g. NES and NNES). This will help us see the rhetorical similarities and/or differences in the use of MD in their academic writing between the groups. In order to situate this in a framework of contrastive rhetoric which will be reviewed in this chapter, the chapter starts by giving a background to the definition of rhetoric and a brief introduction to classical and modern rhetoric. This is followed by a discussion of the relationships between Metadiscourse and rhetoric. Then the interpretation of culture which is broad concept of language will be described in order to explain the language-related factors which might influence the writing style in general and MD usage in particular that vary across languages and cultures, and then the relationship between language, thought and culture will be explained for the purpose of understanding contrastive rhetoric. It then moves to the review of Contrastive Rhetoric. This chapter ends with a summary and some concluding remarks.

2.2 Background to Rhetoric

To discover the relationship between rhetoric and metadiscourse, it is useful to illustrate the important aspects of classical and modern rhetoric in the following sections.

2.2.1 Classical Rhetoric

One of the important characteristics of classical rhetoric (from Greek and Roman antiquity to modern times) is its concern with audience. Greek rhetoric is concerned with public speaking, whereby one speaker, the orator, addresses the people. The basic interest is to move the audience and to win a victory over them by way of persuasion. So the place of the audience has been recognised since the early stages of the art of rhetoric. However, the fact that the speaker's goal is to win a victory over the audience implies that the latter is assigned a passive receptive role rather than being considered as an active participant in communicative situations. Another characteristic of classical rhetoric is that it considers persuasion to be the purpose of rhetoric. Accordingly, Aristotelian rhetoric is defined as the art of producing persuasive texts. Aristotle believed that an effective argument comprises three important aspects. The first is the source or means of persuasion, the three major elements of which are central to Aristotle's rhetoric. These are outlined below:

- (a) Ethos: the personal character of the speaker (the personal appeal of the sender);
- (b) Pathos: engaging the addressee in a particular frame of mind (appeals to the emotions or values of the receiver);
- (c) Logos: words and expressions used by the speaker to prove his case (appeals to reason).

The second aspect of Aristotle's effective argument refers to the language, i.e. selection of appropriate words, expressions, themes etc. The last aspect concerns the organisational structure of the various parts of the argument, i.e. introduction, argument, counter-argument (see Barnes 1984 for more details).

2.2.2 Components of Classical Rhetoric

Classical rhetoric from the Renaissance to the Enlightenment in the 16th century, (see Plett 1985) consisted of five parts: *inventio* (invention), which refers to the finding of argumentative matter; *dispositio* (disposition), the structural arrangement of the argument; *elecutio* (style), the verbal adornment of the matter; *memoria* (memory), memorising the structured and verbally adorned texts; *action* (visual: gesture, facial expression) and *pronuntiatio* (pronunciation).

According to Plett, each of the five parts represents “successive stages in the production of a text” (1985, p60), and each of them has its own set of rules. Plett further argues that “in their totality they constitute the rhetorical competence (or code) of the orator.” (ibid)

It can be said, then, that classical rhetoric is characterised by recognising an audience, but viewing them as passive recipients, focusing on persuasion as a means of winning over an audience and focusing on text production as well as being governed by restrictive rules of a conventional nature.

2.2.3 Modern Rhetoric

New or modern rhetoric (20th century) is, of course, based on classical rhetoric but differs from it in a number of ways. The domain of modern rhetoric is broader and

includes a wide range of choices as reflected in different definitions. Young et al. believe that rhetoric is:

...concerned primarily with a creative process that includes all the choices a writer makes from his earliest tentative exploration of a problem in what has been called the 'pre-writing stage' of the writing process, through choices in argument and strategy for a particular audience, to the final stage of the final draft. (1970, p.xii)

For further characterisation of choices in modern rhetoric, Purves emphasises the role of language use versus language usage and their relationships to rhetoric:

We might define rhetoric as the 'choice' of linguistic and structural aspects of discourse- chosen to produce an effect on an audience. Rhetoric, therefore, is a matter of choice with respect to the uses of language as opposed to those uses that are determined by lexical and grammatical structure (1988, p9)

The fundamental criteria for modern scientific rhetoric have been specified by Plett as "analytical perspective, generative principle, logical coherence, and practical usefulness" (1985, p59). In the first place, the hearer/reader becomes the centre of interest, thus gaining positive and interactive roles. For written texts, this suggests that the writer interacts with the reader through the text. In the second place, performing in a communicative situation requires awareness of all the elements involved in that situation, such as the participants, the subject of the communication, the channel and the 'effect' of communication. Finally, modern rhetoric, as Plett points out, is not only theoretically conclusive, but also practically applicable.

The practical applicability of modern rhetoric can be seen in the different models that have been designed and used in L1 and L2 writing research, including models of metadiscourse and models of argumentation such as those of Toulmin (1958) and Perelman (1982).

Toulmin's model consists of three obligatory elements and three optional ones. The obligatory elements are: *the claim*, which refers to the statement of the problem or argument; *the data*, which refers to the subject of the claim; and *the warrant* provides extra support for the claim, especially when the data is challenged and the relationship between the claim and the data needs to be justified.

The optional elements of the argument in Toulmin's extended model (Toulmin et al. 1979) are: (a) *backing*, which refers to generalisations about the truth value of the different kinds of experience used in the strategy of arguing; (b) *rebuttal*, which refers to "the extraordinary circumstances that might undermine the force of supporting arguments" (Toulmin et al. 1979, p70) that are used to complete conclusions successfully; (c) *qualifiers*, which refer to expressions of certainty and uncertainty (similar to hedges and emphatics in metadiscourse).

Another model of argumentation has been proposed by Perelman (1982). This model is distinguished by its focus on argument and audience. According to this model, there are three audiences: the addressees, the speaker and all humanity at large or what Perelman referred to as the "universal audience". Perelman's model suggests that the way the argument is put forward is determined by two factors: the nature of knowledge of the audience and the purpose of the argument which can be elements of the genre variable in the present study. However, it should be noted that the ways they convey the argument also vary between languages which is an important element of culture.

2.3 Metadiscourse and Rhetoric

The connection between MD, which is a non-content aspect of discourse and will be explained in detail in chapter 3, and rhetoric can partly be seen from the brief

discussion of classical and modern rhetoric presented in the previous sections. Both rhetoric (especially modern rhetoric) and MD are concerned with effective communication. They are concerned with the effectiveness and appropriateness of expressing content or conveying messages. Furthermore, they treat the audience (reader or listener) as their centre of interest. The relationships between the parts of classical rhetoric, Aristotle's major components of rhetoric as well as his elements of effective argument, and MD cannot be overlooked. MD expressions are used as markers of rhetorical moves or actions. Successful writers use them to mark their introductions, goals, explanations, summaries, conclusions, evaluations, attitudes, organisation etc. These devices, as Nash explains;

... imply a working relationship with an audience, the negotiation of an 'I' and a 'you' with the intention of establishing a 'we'. (Nash 1992, p100)

The most obvious connection between MD and rhetoric can be explained with reference to the classical rhetorician's division of oratory into 'taxis' and 'lexis'. According to Nash (1992, pp100-101), Greek rhetoricians use 'taxis' to mean "the structure of a speech, its programme or running order of 'here beneath' and 'firstly' and 'consequently' and 'on the other hand' and 'in conclusion'. As for 'lexis', the second division of oratory, Nash states that "Lexis signified the diction and style of the piece, as adapted to the orator's perception of the formality of his topic and the status of audience". Taxis and lexis have been used to establish the major categories of MD typologies. Nash calls them tactical MD and lexical MD. Other researchers have used similar terms such as textual or informational MD for taxis and attitudinal or interpersonal MD for lexis.

In view of the concepts of classical and modern rhetoric, the concern of a theory of rhetoric as well as the connection between MD and rhetoric, three relevant and

important points need to be made. The first point suggests that it is possible to work towards a rhetorical theory of metadiscourse, a rather lengthy definition of which is provided by Crismore and is in order here:

We shall define a “rhetorical theory of metadiscourse” as any statement or series of statements about metadiscourse, asserted to have some generality of reference and testable by application to relevant acts of discourse. Such a theory is an organised, consistent, coherent way of talking about metadiscourse in any of its forms or modes and involves written texts that systematically assert propositions about how metadiscourse actually works in the world, that identify some rhetorical metadiscursive tactic, strategy, device, etc., and that attempts to account for its effectiveness (or ineffectiveness) in general (1989, pp49-50).

Crismore recognises that the above definition “takes a great deal of territory” (ibid.).

Secondly, due to the observation that rhetoric, rhetorical moves and devices (including MD) vary across languages and cultures, rhetoric becomes a fertile area for cross-cultural contrastive studies, hence contrastive rhetoric (see also Kaplan’s elaboration of contrastive rhetoric in the following pages and how it points to MD). Thirdly, the above two points provide adequate justification for putting the present study in the framework of contrastive rhetoric with language variable (native and non-native English speaker).

2.4 Writing Style and Culture

For the possible explanation of the results of the present study it is necessary to refer to several culture-related factors which include an important element, language, (Hofstede 1991; Hall 1983; Hinds 1987) and might affect the writing style in general and the usages of MD in particular. These factors include characteristics of collectivity and individuality, dimensions of culture such as high and low context, and reader or writer responsibility which will be discussed in the following subsections.

In attempting to define culture, one would usually consider some kind of 'shared patterns' that characterise certain people and distinguish them from other nations. Shared patterns are a result of education and experience and are reflected in the way people live in a general sense. Not only that, but without such patterns people would find it difficult to understand each other and live together (Geertz 1973). This understanding is reflected in the anthropological concept of culture as 'a way of life'. This concept has been adopted by some linguists, like Lado who considered culture as synonymous with the 'ways of people' (1986, p52).

The definition of culture as a way of life is general and rather vague. To unpack this definition, other attempts to define culture recognise the importance of the shared patterns and include what can be considered as components or elements of culture. The following elements have been taken from several sources (The Longman Dictionary of Applied Linguistics, Encyclopaedia Britannica, and Encyclopaedia Americana) and are thought to be components of the concept of culture: language, ideas, beliefs, values, customs, taboos, codes, institutions, tools, techniques, works of art, ceremonies, behaviour, patterns or objects, attitudes, social habits, and knowledge.

In a similar way Hofstede (1991, p5) stresses the collective nature of culture: 'it is the collective programming of the mind which distinguishes the members of one group or category of people from another'.

2.4.1 Individualism and Collectivism

Hofstede (1991) observes that the vast majority of the population of the world belong to collectivist cultures where the interests of the group prevail over those of the individual and where the family structure is that of the extended family type. In

collectivist cultures the group is the source of identity and children are brought up as members of the 'we' group to which they must be loyal and on which they depend for protection against the hardships of life. In collectivist cultures there are very strong ties between the individual and the group to the extent that there is very limited division between personal time and life and work time, as well as living within a group. Relationships with people are more important than the task in collectivist cultures. Such relationships may bring about limitations in the creativity and competitiveness of the individual. Communication in the collectivist culture is of the high context type which will be described in the next section.

Individualist societies are those in which the interests of the individual prevail over those of the group. They are in a minority when compared with collectivist societies.

The family in individualist societies is nuclear and children are brought up to think of themselves as 'I', which is the source of their identity.

Individualistic cultures are low context and verbal communication is more explicit and linear. In the individualist index the USA and Great Britain scored high whereas the Arab countries and Japan scored low (Hall and Hall 1990). Linguistically speaking, this concept can be observed in sequencing ideas or elements, citing sources, putting one's ideas strongly and arguing for them, challenging other peoples' arguments, and the whole idea of putting the author in the centre.

The anthropologist Hall, who studied dimensions of culture for about 40 years, focuses on communication patterns. He suggests four dimensions: context, time, space, and information flow. I will only focus on the dimensions of context as they are more related to the present work than the others.

2.4.2 High Context and Low Context

Hall (1983) distinguishes between cultures according to the degree of context. He describes the communication system of cultures as 'high context' or 'low context'. According to Hall (1983, p210), high context or low context "refers to the amount of information that is in a given communication as a function of the context in which it occurs".

This implies that most of the meaning in high context communication is in the context and very little is in the actual message transmitted. In low context communication, on the other hand, the meaning is explicitly coded in the message. Hall compares low context communication with interacting with a computer that requires an explicit expression of meaning and strict implementation of the programme rules, otherwise the meaning is distorted. A graphic representation of messages and context and provided by Usunier (1993, p103) shows that the Japanese and the Arabs are at the top of the high context scale, the English are somewhere in the middle followed by North Americans. It can be concluded that, in general, oral cultures or cultures greatly influenced by orality e.g. Arabic and Japanese, are high context whereas visually literate cultures, e.g. Western cultures are low context. Moreover, high context messages reflect homogeneity, close and intimate social relations, collectivism, and a high degree of dependency. The communicators, consequently, assume a lot of shared knowledge with their audience and do not see the need to make their messages explicitly coded. This may be reflected in the manipulation of somewhat less use of connectors in the NNES texts in the present data.

2.4.3 Reader or Writer Responsibility?

A significant contribution to contrastive rhetoric and cross-cultural studies is Hinds' (1987) typology of languages as writer-responsible versus reader-responsible. The focal point of the typology in Hinds' words is that

...there are different expectations with regard to the degree of involvement a reader will have, and that this degree of involvement will depend on the language of the reader. (Hinds 1987, p141)

Hinds argues that some languages (e.g. English) are writer-responsible while other languages (e.g. Japanese and classical Chinese) are reader-responsible. The English writer is expected to produce well-organised and clear discourse and to use precise expressions. Commonplaces, poetry and proverbial sayings are not accepted as a means for expressing truth. Failure to meet these expectations can lead to communication breakdown. The classification of writer-responsibility versus reader-responsibility is discussed in terms of coherence or unity. English readers, Hinds further writes,

...expect and require landmarks along the way. Transition statements are very important. It is the writer's task to provide appropriate transition statements so that the reader can piece together the thread of the writer's logic which binds the composition together. (Hinds 1987, p146)

Transition in reader-responsible languages, Hinds explains, is subtle and expected to be supplied by the reader whose role is presumably more active than that of the English reader reading English texts.

Typical modern English writers follow either deductive or inductive approaches in their texts. Consequently, according to Hinds (1990), English readers are used to the deductive and inductive organisational patterns respectively. In other words they expect the statement to appear, deductively, in the initial position, but if it does not the reader

immediately assumes that it is in the final position, i.e. the essay follows an inductive style.

Oriental languages (Japanese, Chinese, and Korean) mostly follow a style whereby the statement is buried somewhere in the passage. This may suggest that the NNES writers may not feel the need to use their conclusions explicitly which leads them to less use of conclusers in the NNES texts.

In general, it can be noticed that Hinds' classification of languages as writer-responsible and reader-responsible coincides with Hall's (1983) descriptions of languages as 'high context' and 'low context', especially the fact that Japanese is at the top of the high context scale.

2.5 Language, Thought and Culture

The relationship between language, thought and culture has been a controversial issue for a long time. Sometimes it has taken the form of the chicken and egg argument. In its extreme version this relationship was expressed by Sapir, and later Whorf, in what has become known as the Sapir-Whorf hypothesis which implies that language controls thought; it therefore shapes and structures our view of the world and behaviour. This hypothesis influenced what has become known as the contrastive rhetoric hypothesis, as coined by Kaplan (1966). The counter argument held by most scholars nowadays implies that language is an important element of culture and it reflects and explains culture. This can be seen, for example, in the use of metaphors, a wide range of vocabulary, ever-increasing technological terms, and a variety of choices made by language users in the organisation of texts, ways of expressing ideas and attitudes and

ways of coding messages in general. It can be concluded that language, thought and culture are interrelated, as put by Valdes (1986, p1)

‘... the current consensus is that the three aspects are three parts of a whole, and cannot operate independently, regardless of which one most influences the other two’.

2.6 Contrastive Rhetoric (CR)

I believe that such culture-related factors are important for the understanding of different writing styles, e.g. contrastive rhetoric, for several reasons derived from Grabe and Kaplan (1996). Firstly, they help us understand the social contexts and attitudes by which writing and writers, and indeed a theory of writing, are influenced. Secondly, they help achieve the goal of contrastive rhetoric “to describe ways in which written texts operate in larger cultural contexts” (Grabe and Kaplan 1996, p179). Thirdly, there is considerable evidence regarding discourse variations and cultural rhetorical preferences.

... The issue should centre on ways to understand differences revealed by contrastive rhetoric, rather than ask whether contrastive rhetoric is a fruitful avenue for research (Grabe and Kaplan 1996, p197)

In discussing early contrastive rhetoric, Connor claimed that “the Sapir-Whorf hypothesis of linguistic relativity is basic to contrastive rhetoric because it suggests that different languages affect perception and thought in different ways” (1996, p10). This weak version of the hypothesis (i.e., that language influences thought), rather than the once-dominant strong version (i.e., that language controls thought and perception), is regaining respectability in linguistics, psychology, and composition studies, resulting in a renewed interest in the study of cultural differences (Gumperz & Levinson 1996).

The term 'contrastive rhetoric' (CR), as coined by Kaplan (1966), is based on the understanding that logic and rhetoric are not universal and that they vary across cultures. Kaplan notes that the linguistic and cultural differences between languages affect the way L2 writers organise their prose. He was particularly referring to international students learning English in the United States. In fact, the basic assumption of CR suggests that, in organising their paragraphs, L2 writers are influenced by the rhetorical preferences of their L1. Elaborating on the concept of CR, Kaplan (1982, pp140-141) states that it refers to;

...the notion that speakers of different languages use different devices to present information, to establish the relationships among ideas, to show the centrality of one idea as opposed to another, to select the most effective means of presentation.

Kaplan's initial interest was to describe those differences in order to suggest solutions to the pedagogical problems of L2 writers. This explains why some of his studies and those of the researchers who followed him focused on compositions written by L2 learners. It is well recognised that Kaplan's ideas regarding contrastive rhetoric came in time and represented a breakthrough in the face of conventional approaches such as the structural approach, the audio lingual method, and the grammar translation method. According to Kaplan, the rhetorical differences between languages cannot be attributed to grammatical and semantic influences only, but one needs to consider the phenomenological differences between the cultures involved. However, an adequate account of cultural differences with respect to students' compositions requires consideration of the educational contexts in which the composition writers were brought up (Liebman 1992; Mohan and Lo 1985). Those contexts may form rival hypotheses (Huck and Sandler 1979) to the very cultural components of contrastive rhetoric.

Kaplan's early conclusions about CR do not seem to have taken educational aspects into consideration. More specifically, Kaplan's conclusions were not accompanied by an investigation or survey of how composition was taught to the international students whose compositions were studied. However, his later publications with Grabe (Grabe and Kaplan 1996) clearly emphasise this issue which may be reflected in the less use of MD features by the student writers in the present data because of their educational contexts (e.g. self references).

Kaplan further describes the L2 writing as 'out of focus' and lacking in coherence and cohesion. The other controversial issue that Kaplan (1966) raises in his seminal article results from the crude graphic forms that he uses to represent the rhetorical structures of paragraphs in five language groups: English, Semitic languages, Oriental languages, Romance languages, and Russian.

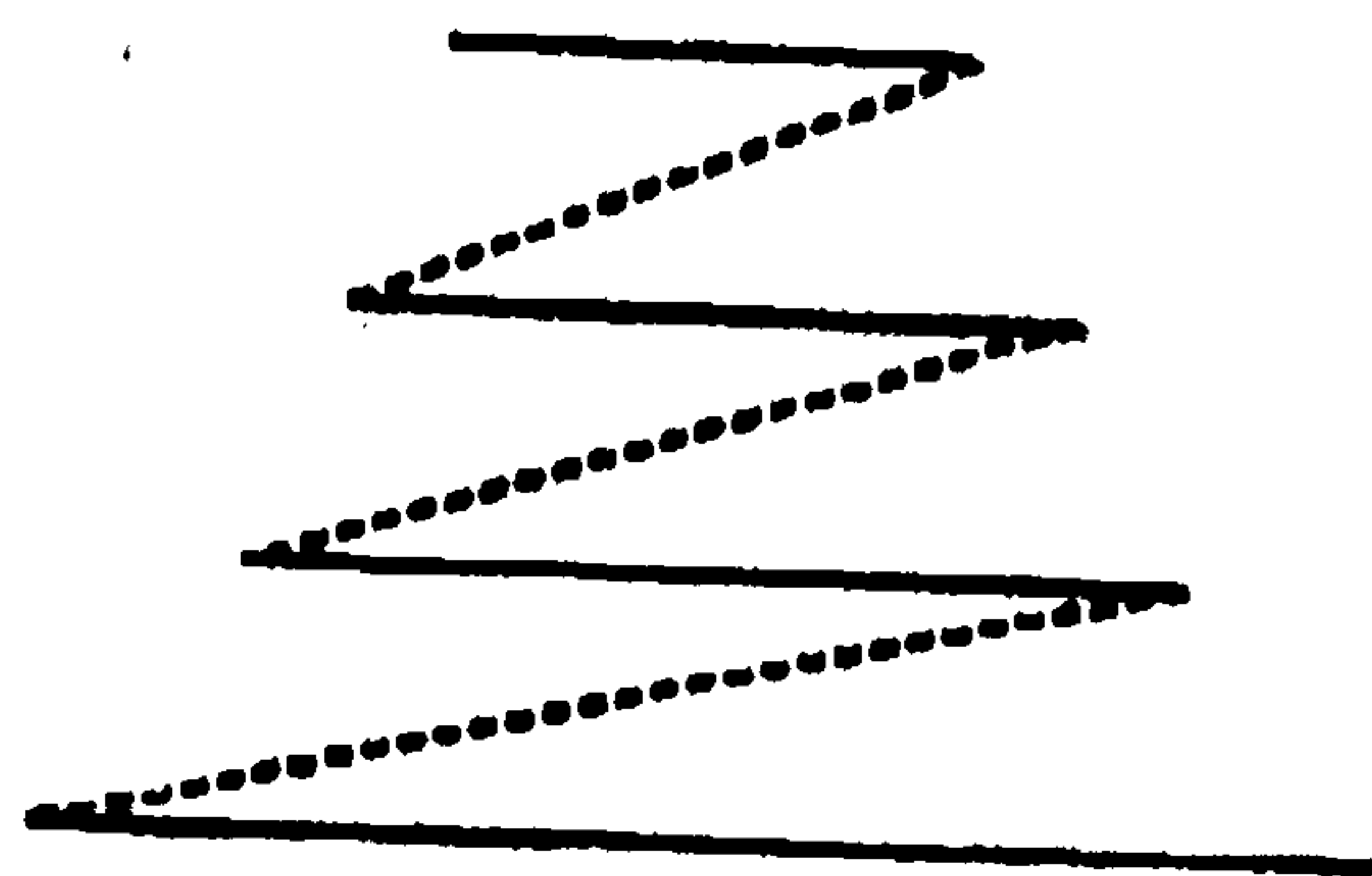
Linear (Anglo-Europeans e.g. English)

The paragraph development in this group follows a linear progression. That is, the paragraph starts with a topic sentence which carries the main idea. The sentences that follow provide elaboration and support for the main idea. Kaplan represented the linear paragraph by a straight vertical line:



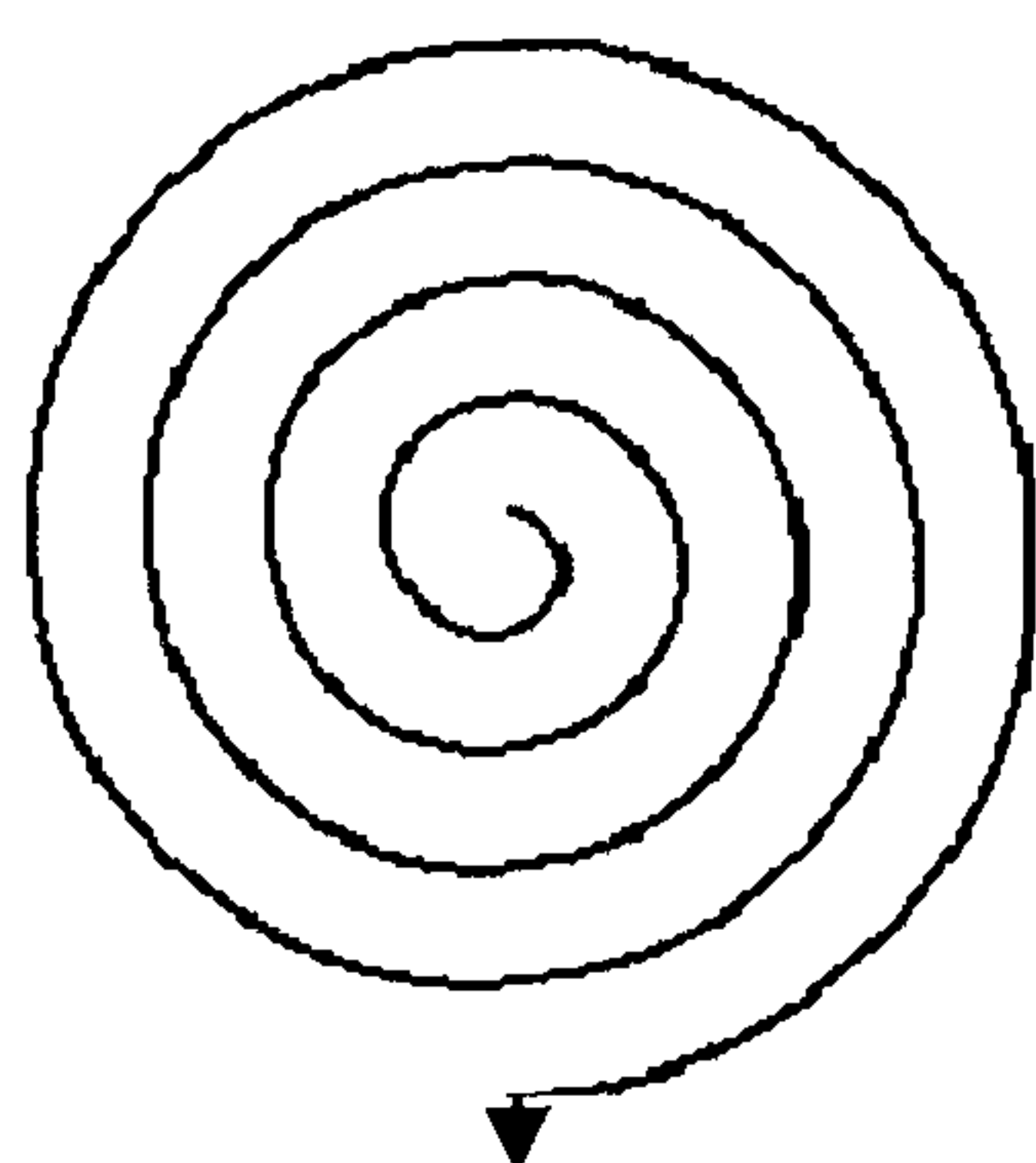
Parallel Construction (Semitic languages e.g. Arabic)

Arab students writing in English use coordination far more than subordination. This results in parallel constructions being used throughout the paragraph. Accordingly the graphic Semitic languages look like parallel lines:



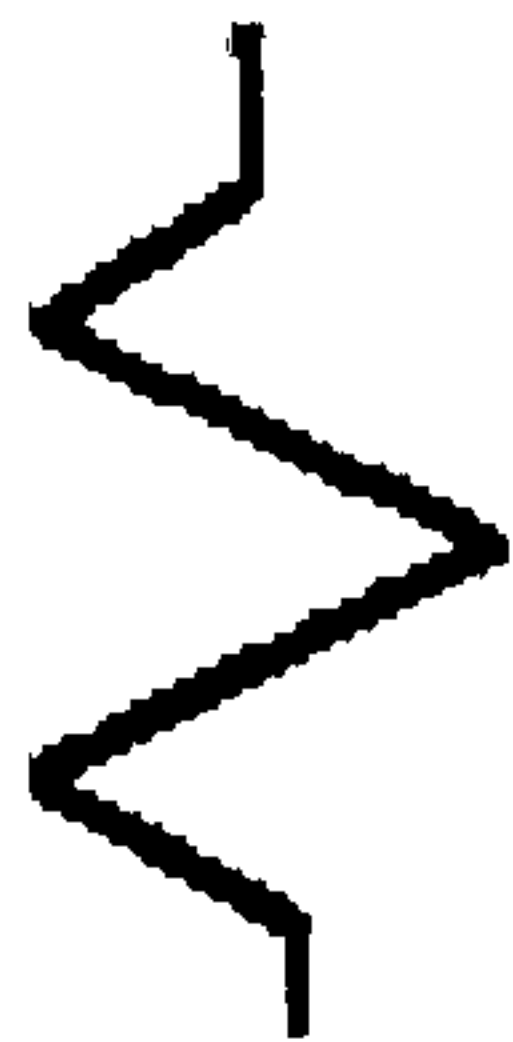
Spiral/Circular (Oriental languages e.g. Korean)

The graphic representation of oriental languages as drawn by Kaplan, took a spiral or circular shape, thus reflecting an indirect way of stating and addressing the main idea of the paragraph. This makes it difficult for a reader from the linear group to locate the main idea and follow the argument. The development of the paragraph in oriental languages is as follows:



Digressive (Romance e.g. French)

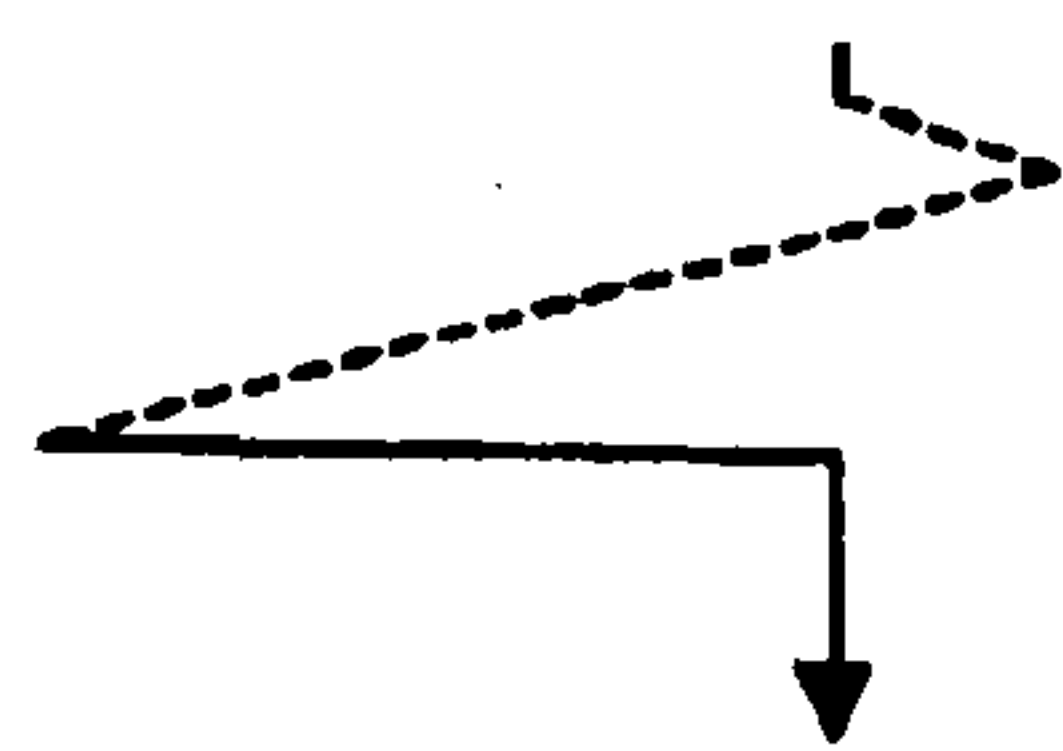
Digression in this group results from the addition of extraneous material which distracts the English reader and does not help develop the paragraph in a linear way. This type of digression takes the following shape:



The digression of Romance languages can be described as ‘weak’ in comparison to the digression of Russian.

Strong Form of Digression (Russian):

This kind of digression is stronger than that of French. It comes as a consequence of attaching many subordinate ideas to the main one to the extent that the latter is buried by the former. It is a kind of open-ended digression as its graphic representation shows:



It has been rightly argued (Kachru 1995 and Matalene1985) that by so doing Kaplan implied that writing like native speakers of English is both necessary and desirable. He also seemed to devalue the patterns that do not conform to the expectations of native speakers of English. The idea of reducing language, cultures and the related thought patterns to diagrammatic structures is hard to digest and has been strongly rejected (Harder 1984; Kachru 1995; Ostler1987; Scollon1997).

In order to assign certain characteristics or behaviour to a particular culture one has to check if those characteristics take the form of a pattern or whether they prevail in other aspects of that culture such as music, arts, architecture etc. and not only in one aspect like language (Scollon 1997).

Furthermore, to imply that, say, all oriental languages are the same is an unjustified overgeneralisation because there are variations within these groups. It is worth noting that Kaplan himself recognises that he overstated the differences and made the case too strong:

... in fact, it now my opinion that all the various rhetorical modes identified in the 'doodles article' are possible in any language i.e. in any language which has written text. (1987, p.10)

He goes on to say:

My contention is that any native speaker of any particular language has at his disposal literally hundreds of different mechanisms to signify the same meaning (ibid).

It can be noticed that most recent CR studies agree that the differences are in cognitive styles rather than cognitive abilities (Purves 1988).

In fact, criticism has been directed at several issues in contrastive rhetoric in general and Kaplan's views in particular. Harder (1984) argues that in addition to reflecting the native culture, L2 compositions exhibit a mixture of the logic and rhetoric of L2 and L1 cultures. He adds that this could be a result of composition learning methods or the conflict between culture-linguistic systems. On the basis of his studies of compositions

written by Chinese and Japanese students, Harder suggests an extension of the model of interlanguage ¹in ESL and TESOL (Selinker 1972; 1992).

It can be argued that the mixture of logic and rhetoric observed by Harder might have resulted from what Goodenough (1971) called 'operational culture'. That is, when one finds oneself in a situation different from one's cultural background, one tries to adjust even temporarily, to the new situation. This can be found from the non-native speakers' awareness of MD use from the interviews in the present study.

For a long time, and under the influence of Kaplan's early thesis, research in CR focused on one main aspect, namely the organisational one, at the expense of other important aspects of written discourse that vary between languages and cultures. In fact, organisation corresponds to 'arrangement' which is only one of the five elements of rhetoric as defined by Aristotle. The five elements are invention, memory, arrangement, style, and delivery (Liebman 1992). In addition, Purves (1988, pp11-12) includes the following areas: discourse functions, organisation or reorganisation of material that takes different forms like shopping lists, letters, articles etc., the generation of content and/or form, and the pragmatics of discourse. Focusing on one aspect only might be one of the reasons that led Kaplan and his followers, for some time, to deal with compositions as products and neglect the process of writing which takes into account various aspects. Many of Kaplan's critics pointed to this negligence as a very negative aspect that leads to prescription, especially when the findings of CR are applied to L2 writing (Leki 1991). The transfer that is manifested in L2 writing, in addition to organisation, reflects other aspects such as planning (Jones and Tetroe 1987), skills (Mohan and Lo 1985), metadiscourse (Crismore et al. 1993; Mauranen 1993a) etc.

¹ Selinker (1972) proposed and elaborated the term 'interlanguage' to explain the unique utterance of L2 learners. It is regarded as a separate linguistic system which results from a learner's attempts to produce a target language norm.

The situation is that there are cultural similarities and differences between different languages, not only at the organisational level, but at other levels such as cohesion, coherence, style, modes of argumentation, persuasion, voice, tone etc. Ideally, every language has its own distinct characteristics and preferences, which are clear when theory-oriented contrastive research is conducted, but it seems difficult to establish borders given the present status of research in the world of today, especially the pedagogy-oriented aspect in L2 writing.

2.7 Summary

The link between metadiscourse and rhetoric was found from the review of the background to classical and modern rhetoric, their components and definitions in the present chapter. As this study is in the research paradigm of comparative academic writing, the chapter discussed cross-cultural aspects of writing, particularly culture-related factors and contrastive rhetoric, to explain the similarities and/or differences in writing style between languages (native and non-native English speaker). The culture, thought and language are interrelated. So this chapter reviewed the culture-related factors which might affect the writing style in general and MD usage in particular, and which it is hoped will be useful for the possible explanation of the results of the present study.

In the next chapter the literature related to MD will be considered in order to understand the definition of MD proposed by other scholars. Also the working definition of MD in the study will be presented as a result of carrying out a review of the literature related to MD definition.

Chapter 3: A Working Definition of Metadiscourse

3.1 Introduction

This chapter is concerned with the issues related to the basic concepts and definitions of MD. It reviews and analyses different aspects of MD, such as the concept of MD in relation to discourse, definitions, and characteristics. The chapter also discusses definitions of concepts similar to MD and points out the differences between them.

Details of examples of MD and similar terms will be presented in the next chapter. The chapter will also provide a working definition of MD to be used with the typology to be developed in the next chapter.

3.2 The Concept of Metadiscourse

The aim of this section is to bring into focus the concept of ‘metadiscourse’ and to show its place in the context of the general term ‘discourse’. The term ‘metadiscourse’ points to two levels of discourse: the first level is known as the ‘primary’ discourse and the second level is known as MD. The primary level carries the core message, the propositional content. Metadiscourse, on the other hand, consists of the linguistic expressions used by the writer to ‘signpost’ the reader to understand his or her message, attitudes and the overall organisation of his or her text. It is considered as the non-content part of discourse. These two levels are two complementary parts of good discourse that leads to successful communication.

The concept of two levels has been used by several scholars. Lautamatti (1978) discusses the elements that create coherent texts and differentiates between two types of elements: ‘the topical elements’ that refer to the content or topic of the discourse and ‘the non-topical elements’ that help readers understand the text, and can be counted as MD.

There are similar ideas in the work of Vande Kopple (1980), an author who speaks of the first level, ‘the primary level’, and explains that it refers to the propositional content, and the second, ‘the secondary level’, which is MD “does not add propositional information but signals the presence of the author” (Vande Kopple 1980 in Beauvais 1986 p50).

Williams (1981) defines the first level as ‘the primary discourse’ which represents the content or subject matter. The second level, ‘metadiscourse’, refers to the different

expressions that are related to the act of discoursing, 'discourse about discourse', and that add nothing to the content of the primary discourse.

Like Williams, composition theorist Dillon (1986) suggests that the first level is concerned with information about the content of the text, i.e. its subject. The second level consists of expressions that help the reader understand the organisation and evaluation of the text and lead to reaction to the content. This level also involves what has become known as the presence of the author and the author/reader relationships.

The relation of the above concepts to the macro-functions of language should not be overlooked. Halliday suggested the functions of language into three broad macro-functions: ideational, textual and interpersonal. Each is concerned with different aspects of the world and the 'modes of meaning' that are identifiable in language in social contexts (Halliday 1994, p27). The ideational function focuses on the natural world and events, including our own consciousness. He describes this as the 'content function of language' (Halliday 1994, p27). The textual function focuses on the organization of text. It enables language to be, 'operational in a context of situation' (Halliday 1994, p27). The interpersonal function focuses on the social world. It expresses attitudes, judgments, and relationships: the, 'participatory function of language' (Halliday 1994, p27).

Halliday (1973) indicates that the non-content elements (MD) are concerned not with the message (the content) as such but with how that message is put to the target audience, i.e. the devices used to put the message across, such that the target audiences can understand it unambiguously.

Halliday's ideational function is related to what we referred to above as the propositional content which is not part of MD. Thus Halliday's interpersonal and textual functions will be elaborated in the following paragraphs.

The interpersonal function, according to Halliday (1973), is related to role mediation both at the individual level, expressing personal aspects and feelings, and at the social level, social interaction between participants in the communication. This role of language gives them the opportunity to explicitly reflect their personal characteristics, attitudes, viewpoints, subjective evaluation of the propositional content and to establish a working social relationship with each other. This applies both to oral as well as written communication, as reflected in the views of writer/reader relationships (Tierney et al 1979; Tierney and Raphael 1981; Bruce 1981) and writer-text-reader interaction (Nystrand 1986).

Halliday (1973) sees the textual function as enabling us to create a text and organise its content so that it makes sense to the other participants (readers) and qualifies it as a message. With respect to MD, especially in written texts, this function is important for writers as well as readers. It helps writers relate various portions and sections of the text to each other and readers to see that relation and interpret it according to the writer's intentions.

These two MD related macrofunctions have influenced various aspects of MD as seen in the works of several scholars. Williams (1981 and 1990) includes connectives and advance organisers under the textual function and hedges, attitude markers under the interpersonal functions (details will be explained in the next chapter). Vande Kopple (1985) also identifies two main kinds of MD, the interpersonal MD and the textual MD, with reference to Halliday's typology of meaning. Recently, Crismore (1990) suggests

that, unlike propositional and referential elements, MD features do not refer to the external reality but to text/discourse organisation and writer/reader relationships.

3.3 Characteristics of MD

3.3.1 Common Factors in MD

While it is necessary to make a distinction between the two levels of discourse, there are common factors that unify all types of MD. The first common factor is that “they do not expand the propositional information of a text” (Vande Kopple 1985 p85). In other words they do not carry a message that refers to the real world (Halliday 1973) or external reality (Crismore 1990). The second common factor is derived from Nash (1992, p104) and concerns “the writer’s consciousness of what he is writing and his perception of the need for further elucidation, for qualification, for emphasis, for evaluation, for the involvement of the reader in the assertions of the text”. The third factor concerns the common primary function of metadiscursive expressions which is for “helping the reader interpret, not syntactic” (Crismore et al.1993 p49).

3.3.2 Contribution to Authorship

Authors use MD in nearly every move in their text. In their attempt, for example, to express certain ideas, authors borrow some parts from other sources that they respectfully acknowledge through the MD expressions of attribution and/or narration (e.g. *according to X...*, *X has maintained that*, *X suggested*). They comment on it, criticise and evaluate using commentary and evaluatives (e.g. *inappropriately*, *disappointingly*). They also use pronouns to pursue the points that they themselves add to draw special attention to their point (*I mentioned*, *I promise to*). Moreover,

attitudinal or interpersonal MD is used to express their attitudes to the content and to the reader (*surprisingly, it is more significant*). Authors also guide readers step by step from the introduction to conclusions with their goals.

3.3.3 The Multifunctional Nature of MD

Multifunctionality refers to the simultaneous performance of more than one function by the same MD features. As a result of this, MD definitions and typologies have been criticised as being imprecise, multifunctional and fuzzy (Beauvais 1986; Crismore 1989; Nash 1992; Vande Kopple 1985). Indeed, most of the scholars who have written about MD maintain that fuzziness and multifunctionality are natural characteristics of MD. Features like ‘I hypothesise’ and ‘I think’ can be classified as illocutionary markers, attributers, narrators (i.e. *I* is used to mark the speech act as an illocutionary marker; they also signal the source of what is presented as narrators and attributers). Such a classification is therefore highly subjective and influenced by factors such as the person who interprets the discourse and analyses it as well as the context in which the discourse is produced. In this part, the multifunctional nature of MD will be highlighted and explored with various attitudes from different scholars.

Multifunctionality is also found at the level of the macrofunctions. The lexico-grammatical expressions that are used to express the macrofunctions themselves are multifunctional, i.e. they can serve more than one macrofunction simultaneously. Crismore and Vande Kopple (1997) presented an example of this, whether the agent is realized or deleted in a passive construction; there are both functions, an interpersonal one as in the personalization (‘I argue that smoking causes cancer’) and a textual as in the depersonalization (‘the findings suggest that smoking causes cancer’). Crismore argues that MD is confusing and problematic because it involves confusing elements

and “almost all properties of spoken and written discourse may implicitly or explicitly signal various dimensions of the communicative situation” (1989, p49).

Moreover, as Crismore argues, some MD features behave as primary discourse in some situations and as content in other cases. Take the function of ‘really’ in the following examples from Crismore et al.(1993 p41) a) *Really it was terrible.* b) *It is really terrible.*

In the first ‘really’ functions as MD and in the second it emphasises the adjective. This multifunctional nature of MD can be found from the function of ‘in spite of’ in the following examples from my corpora a) *In spite of this, more recent versions of the curriculum begin to...* (NES15). b) *Thus, conjunctions (e.g. and, but) adverbs (e.g. nevertheless, conversely) and prepositional expressions (e.g. in spite of) can all function as connectives* (NES 7). In the first, ‘in spite of’ functions as MD and in the latter it constitutes propositional information of text. It is therefore not possible to maintain precision and complete systematisation about it. While convinced that in essence MD is ‘extremely dishevelled’ and will contain non-discrete and fuzzy categories, Crismore et al. (ibid) assert that these characteristics must be reflected in any theory or typology of MD. Consequently, the same MD elements may be judged differently by different people. This makes the boundaries of MD difficult to establish (Swales 1990).

Multifunctionality is also a characteristic of other labels that are similar to MD such as discourse markers, metatext and gambits as can be seen in the next section (3.4.2).

3.3.4 The Social Characteristics of Metadiscourse

The social dimension is an important feature of MD that enables it to play a marked role in shaping social interaction. This can be seen from its relation to rhetoric, its role in establishing writer/reader relationships, its emphasis on other relevant factors such as the presence of the author, commentary and above all the recognition of an audience who deserve to be guided in order to grasp the intended message. MD derives its social character from the nature of the language as a social phenomenon and indeed as a social instrument that people use to shape their social world (Bruner 1984) and to use language to do things in that world (Austin 1962).

3.4 Definitions of Metadiscourse and Parallel Concepts to Metadiscourse

There are several terms, used by some researchers in the field of discourse analysis, which have similar meanings to that of MD, for example, ‘signalling’ (Meyer et al 1980), ‘Metatalk’ (Schiffrin 1980; Ragan and Hooper 1981), ‘alignment talk’ (Hooper 1981), ‘gambits’ (Keller 1979), ‘discourse marker’ (Schiffrin 1987; McCarthy 1990; Fraser 1990; Redeker 1990). These terms are considered by Crismore (1989) as MD for at least two reasons. The first is that they belong to the secondary level, i.e. the non-content aspect of discourse. The second is that their main function is to guide the reader/listener as to how to understand the author/speaker, and his message.

3.4.1 Metadiscourse

As regards the coinage of the term, there seems no agreement among researchers when MD as a term was first used and who first introduced it. Beauvais, in his Speech Act

Theory of Metadiscourse (1986), claims that the term MD was coined by Harris in 1959 who viewed it as a subset of metalanguage. This claim challenges Intaraprawat's claim that it was Williams (1981) who first originated the term. Strangely enough Intaraprawat and Steffensen (1995) claimed that the term was coined by Harris in 1970 rather than in 1959.

When Harris (1959) dealt with MD kernels, he viewed them as 'retrieval status of kernels' (p944). In fact, MD appeared as the fourth type of kernels² identified by Harris that is "MD kernels which talk about the main materials and referring to a main kernel" (p944). Harris' main interest was to develop a system for extracting important information from scientific articles in order to construct abstracts of those articles. Here also we find reference to the two levels of discourse and an attempt to describe MD as 'statement about the discourse' and 'sentences which speak about discourse'. These issues have been picked up by MD researchers who came after Harris and tried to define MD.

Twenty years after Harris, Williams (1979 and 1981) emphasises the importance of MD. His dissatisfaction with reliance on the grammar of the sentence as the first level of analysis and his contention that discourse gives sentences their right shape lead him to suggest a shift of attention to the rhetorical structure of the sentence. This shift, as we shall see in different aspects of MD, influenced the way people viewed writing and reading. It resulted in a change of attitude from the text as autonomous or writer-centred to writer-text-reader interaction models and social interaction models of reading (Bruce 1981; Nystrand 1986; Dillon 1986). At the heart of these models and processes,

² The kernel is the set of elementary sentences and combinators, such that all sentences of the language are obtained from one or more kernel sentences (with combinators) by means of one or more transformations (see Harris 1959).

Williams sees a significant role for MD which is, according to him, needed in sentence analysis. He describes the MD category as

.... all the elements in a sentence that refer to the process of discoursing, as opposed to the specific reference of the discourse (William 1979 p33)

To make his concept clear Williams gives the following example:

I believe that in regard to the American pharmaceutical industry, we can say that there seems to be excessive federal government regulation.

The italicised parts are MD which Williams describes as

..... discourse about discourse, elements referring not to the referents external to the discourse but to the act of discoursing, to how we should take the truth value of probability of the proposition about those external referents (Williams 1979 p33)

Williams further stresses the non-propositional character of MD, its interactive role and the opportunities it provides to establish writer/reader relationships describing it as:

.... discourse about discourse, words and phrases and clauses... even sentences... that refer not to the subject 'out there' but to the act of discoursing, to the speech event that the discourse and its reader create (Williams 1981 p195)

Williams considers MD as a stylistic variable and in fact all his ideas about MD appeared in his works on style.

Having clearly made a distinction between MD and the primary discourse, discussed previously, Vande Kopple defines MD as

.... discourse about the act of discoursing, discourse which does not add propositional information but which signals the presence of the author (1980 pp50-51 in Beauvais 1986)

He goes on to say that

This kind of discourse calls attention to the speech act itself, often marking stages in the development of the primary discourse, displaying the author's position on

the primary discourse, or moulding the reader's attitude about the primary discourse (ibid).

In this definition, there is an explicit emphasis on the author's presence, link with speech act and the rhetorical characteristic of exerting effects on the reader.

In another influential article, Vande Kopple puts further emphasis on writer-text-reader relationships and interactions. Not only that, but he also provides hints for establishing a typology of MD. He confirms that at the level of MD

...we do not add propositional material but help our readers organise, classify, interpret, evaluate and react to such material. Metadiscourse, therefore, is discourse about discourse or communication about communication (Vande Kopple 1985 p83)

This definition suggests that in order to be able to identify MD features in a text, one needs a typology that consists of categories and subcategories of some kind. Such categories would help one find out whether or not the writer helps the reader to organise, classify, interpret, evaluate and react to the propositional material.

Like the above mentioned definitions, two main points seem to characterise the definitions of MD by Crismore. The first is the observation that it adds nothing to the propositional content, but directs the reader to understand the primary discourse, writer-text-reader relationships. In one of her early definitions, Crismore states that MD is

... the author's intrusion into the discourse, either explicitly or non-explicitly to direct the reader rather than inform (Crismore 1983 p2).

She then elaborates on this definition describing MD as

... the directives given to readers so they will understand what is said and meant in the primary discourse and how to take the author (ibid).

What may be new in this definition is that Crismore uses expressions that point to textual metadiscourse and attitudinal metadiscourse. The first is referred to by the

expression 'discourse plan', whereas the second is referred to by the author's attitude, and confidence and the interpersonal part. In this definition Crismore seems to confine MD to those elements that 'comment'. However, it is not clear whether there are types that do not comment and whether all comments are 'contentless writing about writing'. Although it is admitted that MD is a non-content aspect of discourse, Crismore's description of MD as 'contentless' is strong. It can be argued that some MD features like *it seems possible* and *you may like to* have content on their own. Furthermore, like 'explicitness' (Crismore 1983), 'commentary' in her definition needs to be explained in detail. It is worth noting that Crismore repeated the same constituents, used in the above-mentioned definitions, to formulate definitions for other studies (Crismore 1985; 1990; Crismore and Farnsworth 1990).

Nash argues that MD is an ordinary stylistic symptom that is taken for granted in lectures as well as different types of writing (scholarly writing, expository prose etc.).

He further explains that

... we assume from the outset that the expository text will be mediated to us, that the author will assume the role of guide, regularly coming between us and the text, taking his distance from his writing, on the one hand and his audience on the other (Nash 1992 pp106-107).

The key word in this concept is 'mediating' which according to Nash, refers to:

.... any form of language which can be interpreted either as commentary on the process of making a text or a negotiation of relationships with an audience (Nash 1992 p114).

Nash obviously stresses the textual and interpersonal importance of MD, the writer/reader relationships and interaction and the presence of the author as reflected in the negotiation of relationships with an audience, thus considering MD as a rhetorical

plane of discourse. Although Nash does not refer to previous work on MD or other relevant literature, his ideas are very similar to the definitions discussed above.

Beauvais criticises those definitions on the grounds that they all lack precision and involve undefined elements such as propositional information, explicit/non-explicit, overt/non-overt etc. While Beauvais' observations are valid, one needs to take two points into account. The first is that the concept of MD is relatively new and much research is needed to clarify it. The second point is that many MD features are multifunctional and it is impossible to eliminate this characteristic. As Crismore puts it "...we cannot be precise and completely systematic about metadiscourse" (1989 p49).

Beauvais located MD within the framework of Speech Act Theory and defined it as

"... illocutionary force indicators that identify expositive illocutionary acts" (Beauvais 1986 p112). His theory is based on two main sources. First, the ideas of Searle (1976) for the distinction he makes between illocutionary force indicators (that show how a proposition is to be taken) and propositional indicators (that convey statements with truth properties). The second is Austin's (1962) category of expositive illocutionary acts that clarify communication. Beauvais claims that "these terms specify that metadiscourse consists of structures that identify the communicative functions served by propositions in a passage of discourse" (Beauvais 1986 p113). In addition to these two, he is influenced by the link between MD and speech acts mentioned by Vandekopple (1980) and Crismore (1985) who suggest that MD points to the speech act. Beauvais' definition is too brief and the illocutionary acts are one aspect of several elements included in other definitions and typologies of metadiscourse.

Chafe and Nicholas (1986) and Barton (1993) have recently been devoted to what have become known as *evidentials*, which refer to words and phrases that express attitudes toward knowledge. Evidentials include two MD elements: validity markers and attitude markers. They belong to the secondary level of discourse (non-propositional) as Barton asserts and take the characteristics of MD

....an evidential is a non-propositional word or phrase used to express an attitude toward knowledge (Barton 1993 p746)

Hyland defines written metadiscourse as those “aspects of a text which explicitly organise the discourse, engage the audience and signal the writer’s attitude” (1998 p437). He further states that

Based on a view of writing as a social and communicative engagement between writer and reader, metadiscourse focuses our attention on the ways writers project themselves into their work to signal their communicative intentions. It is a central pragmatic construct which allows us to see how writers seek to influence readers’ understandings of both the text and their attitude towards its content and the audience (p 437).

He later explains that as ‘a system of meanings realised by an open-ended set of language items’ and provides the definition:

MD is the cover term for the self-reflective expressions used to negotiate interactional meanings in a text assisting the writer (or speaker) to express a viewpoint and engage with readers as members of a particular community (Hyland 2005b p37)

A summary of the elements mentioned in the different definitions of MD mentioned above is: ‘secondary discourse’, ‘discourse about discourse’, ‘non-content’, ‘does not add propositional information’, ‘author presence’, ‘writer reader relations and interaction’, ‘help readers organise, classify, interpret, evaluate’, ‘directing readers rather than informing’, ‘referring to the process of discourse’, ‘stylistic variable’, and ‘rhetorical act’, as shown in the table below.

Table 3- 1: Elements Mentioned in the Different Definitions of MD

Researcher	Secondary discourse	Discourse about discourse	Non-content	Does not add propositional information	Author presence	Writer reader relations and interactions	Help readers organise, classify, interpret, evaluate	Directs rather than informs	Referring to the process of discourse	Stylistic variable	Rhetorical act
Harris 66	+	+				+					
Williams 81		+	+			+			+	+	
Vande Kopple 80	+	+	+	+	+	+					
Vande Kopple 85		+		+			+				
Crismore 83					+	+		+			+
Crismore 84			+	+	+	+					
Crismore 85					+	+					
Crismore 90		+			+	+	+				
Crismore and Farnsworth 90					+		+				+
Nash 92					+	+		+		+	
Crismore et al 93				+			+				
Beauvais 86											
Hyland 98					+	+		+			
Total	2	5	3	4	8	9	4	5	1	2	2

3.4.2 Parallel Terms to MD

As mentioned above, several terms have been used in a way similar to that of MD. In this section a brief discussion of those terms will be presented.

a. Non-topical Materials

Lautamatti (1978) uses ‘topical linguistic materials’, which relate directly or indirectly to the topic of the discourse or text, and ‘non-topical linguistic materials’, which are separate from the topical material (the primary discourse or core content/message), essentially the same distinction as that made by Williams (1981) and Vande Kopple (1985).

b. Metatext

Some use the term ‘metatext’ in a sense similar to metadiscourse and restrict it to textual MD (e.g. Mauranen 1993a), while others (e.g. Popovic 1976) take a broader view and consider it as a whole complete text about the original text. Popovic puts metatext in the wide framework of a literary communication model where two types of communication exist.

The first is the sphere of primary communication represented by the original literary text. The second type is metacommunication which

.... refers to all types of processing (manipulation) of the original literary text whether it is done by other authors, readers, critics, translators etc. This processing is manifested in the form of further texts which are about the original text (Popovic 1976 p226)

He calls the original text ‘prototext’ and the text that results from processing he calls ‘metatext’, which “constitutes the so-called direct mediation of the original text” (ibid).

In more precise terms he considers metatext “a model of the prototext, the way in which two texts are linked” (p226). He later explains that it “calls attention to another work” (p233) and that “is a meta-sign of a work which is already in existence” (ibid).

In a recent contrastive study Mauranen (1993a) uses metatext as an equivalent to TMD as defined by Vande Kopple (1985), Crismore and Farnsworth (1990). She states that

.... metatext is essentially text about the text itself, it comprises those elements in text which at least in their primary function go beyond the propositional content (Mauranen 1993 pp7-8).

Examples of Mauranen’s metatext are: *however, for example, so far we have, we show below* etc. All these examples belong to textual metadiscourse.

c. Gambits

Keller (1979) used the term for the verbally represented strategies that conversation participants use to structure the content as well as the procedure of their conversation. These are signals the general function of which is to preface what a conversation participant is going to say. In Keller’s words

A certain set of signals in the conversationalist’s speech used to introduce level shifts within the conversation, or to prepare the listener for the next turn in the logical argument (Keller 1979 p220)

e.g. to begin with, this reminds me of, my guess is etc.

d. Meta-talk

In Schiffrin’s (1980) analysis of conversations, she differentiates between two types of metalinguistic expressions. The first comprises those expressions that focus on an individual’s own talk, e.g. *that’s what I meant, I’m telling you*; these act as evaluative and organisational brackets. The second includes expressions that focus on an

interlocutor's talk like *what do you mean by that? that's your opinion*. They act as evaluative brackets rather than organisational ones, whereas *you said* acts as a renewal bracket. Schiffrin describes such expressions as meta-talk, talk about talk, language talking about itself, the metalinguistic referents, the metalinguistic operators, and the metalinguistic verbs.

The term meta-talk is also used by Ragan and Hooper (1981) to describe explicit forms of metacommunication. They claim that it helps manage conversational repair such as *I mean, you meant*, and to disambiguate participants' utterances.

e. Discourse Markers

It has been reported (Fraser 1990) that research on discourse markers in English started in the eighties. Levinson (1983) was among the first to attract attention to discourse markers. He describes them as those words and phrases "that indicate the relationship between an utterance and the prior discourse" (Fraser 1990 p384).

In 1985 Schourup studied conversational discourse markers such as 'well', 'you know', which he calls particles and describes them as 'evincives' that indicate a process of thinking (by the speaker) at the time of speaking or just prior to it but do not state its content in detail. In comparison, Fraser (1980) claims that some discourse markers in conversations like 'well' signal "some forthcoming dissonance" (p389).

A more detailed and focused study of discourse markers was carried out by Schiffrin (1987). According to Schiffrin, discourse markers are sequential and are used as backward and forward indicators. In her words, discourse markers are "sequentially dependent elements which bracket units of talk" (p31). They include markers such as *and, but, because, I mean, now, oh, or so, then, well, y'know*, used in unstructured

interview conversations. Schiffrin defines brackets as “devices which are both cataphoric and anaphoric and whether they are in initial or terminal position” (ibid).

Like Schiffrin and Fraser, Redeker (1990) recognises the sequential dependency of markers and defines a discourse marker as “a linguistic expression that is used to signal the relation of an utterance to the immediate context” (p372). Redeker focuses on the markers of ideational structure and markers of pragmatic structures which language use usually involves and which, as he claims, are equally important for an account of discourse coherence. Although Redeker’s classification shows great resemblance to MD, his definition is vague and does not indicate the level of discourse. Besides, the phrase ‘immediate context’ requires explanation and precision. Otherwise it may be confused with either Halliday and Hasan’s (1976) concept of cohesion and/or Winter’s (1977) and Hoey’s (1979; 1983) concept of signalling.

It is worth noting that spoken discourse markers (*well/and/so*) are found in written discourse and written discourse markers (*finally, in sum...etc*) are also used in spoken discourse. McCarthy (1990) studies discourse markers ‘Spoken Discourse Markers in Written Text’. He defines them as

... they signal to the receiver independent of content what is happening, where the discourse is, where it is going, whether it has finished, whether utterances follow smoothly from what has been uttered before or whether some kind of disjunction is occurring. (McCarthy 1990 p172)

This concept of discourse markers is broader than the previous ones and relates to discourse organisation and structure and textual MD. However, it is still narrower than MD because it does not accommodate both the textual and interpersonal metadiscourse.

f. Signalling

Meyer (1975; 1982) and Meyer et al. (1980) used the term to refer to the non-propositional content, the presence of the author and the relationships the author creates with his readers. In an article on 'Use of top-level structure in text' Meyer et al (1980) adopted Meyer's former definition of signalling as "a non-content aspect of prose". This definition, as Meyer explains, gives emphasis to certain aspects of the semantic content or points out aspects of the structure of the content and which

shows an author's perspective on the relative importance of the content related in his passage (Meyer 1975 p313).

Although Meyer and her colleagues do not use the term MD, they do make a clear distinction between the content and non-content levels of discourse. Also some of the examples they use are similar to MD e.g. *in short, unfortunately, this is an important point*. They also point out that specifying the structure of text by way of signals helps identify variations between the text and the reader's understanding of it. Not only that, but it also permits "theorising about how readers process and understand text" (Meyer et al 1980 p75). Among the interesting things that Meyer and her colleagues point out is the effect of signalling on the reader as writer-text-reader interactions. They claim that signalling helps establish the structure strategy (an effective way of reading and understanding based on using the signals used by writers to signal text structure), which in turn helps readers understand and focus on major and logical text-based relationships among propositions and helps reduce the cognitive load. This load may result from the reader's intensive efforts to infer and figure out text structure because of the near or total absence of signals.

The suggestions made by Meyer and Meyer et al.(1982) cover the essence of MD: the non-propositional content, the author's presence, and the social character of MD as reflected in the work of Williams, Vande Kopple, Crismore and others, reference to which has already been made.

3.5 A Proposed Working Definition of Metadiscourse

The previous sections dealt with the basic concept of MD, its definitions, and characteristics as well as the definitions of parallel terms to MD. It has been noted that MD is fuzzy and there is overlapping between the different subcategories. Nevertheless, it is believed that the fuzziness and overlapping can be reduced with a degree of compatibility. In other words, a good definition of MD needs to reflect the essence of the concept as a non-content or non-propositional aspect of discourse, a view of its functions and a link to the typology on the basis of which the textual analysis can be carried out. The definition proposed here has been formulated with these issues in mind, drawing on the MD definitions proposed by other scholars, especially those that have been used frequently (see table 3-1 above).

In order to decide which elements to include in the definition, the elements that appear in the table above (Table 3-1) have been synthesised as shown in the table below (Table 3-2). This table consists of two columns. The first comprises the elements that will be retained in the proposed definition, and the second comprises the elements that can be put under the umbrella of the retained ones.

Table 3- 2: A Proposed Synthesis of the Elements in MD Definitions

Proposed elements	Considered elements to the proposed one
Non-content elements	Secondary material, discourse about discourse, does not add to the propositional content, does not refer to the subject
Helps readers organise, classify, interpret, evaluate and react to the primary discourse	Marks stages in the development of the primary discourse
Author's attitude and intrusion	Writer/reader relationships, directs rather than informs the reader, how to take the author, how to take the truth value, stylistic variable, commentaries, calls attention to speech act.

On the basis of the above MD can be defined as:

A non-content aspect of discourse which represents the writer's point of view and attitude to the content of text in order to help the reader react to the content of the text and the writer's point of view/perspective/attitude.

The elements mentioned in the definition will be taken into consideration in establishing the typology of MD, which will be described in the next chapter. The review of the different typologies (in the next chapter) arrives at two main categories, textual MD and interpersonal MD as some studies call them. In the present context, I propose a taxonomy consisting of two main categories of MD. The first, called textual MD, comprises textual elements which help the reader to refer, organise and classify the content of text. The second category is interpersonal MD, comprising interpersonal elements which express the writer's point of view to the content of the text in order to guide the reader to evaluate, interpret and react to the content of the text and the writer. The working definition considers not only these two main categories but also their subcategories in the proposed typology.

3.6 Summary

In this chapter, the basic concept of MD has been explained through the two levels of discourse: primary discourse which carried the core message or content and MD that consists of the orientation expressions that writers use to guide their readers to understand the message and the writer's attitude to the text and the reader. This was followed by characteristics of MD.

MD is multifunctional and fuzzy, so it is difficult to arrive at a consensus for the identification and classification of its features. Two important characteristics have been highlighted. The first is the social characteristic which can be observed from the relationships it creates between the writer and the reader. The second is the contribution it makes to the theories of authorship by putting the writer in the centre of his/her text expressing his/her views and attitudes by helping the reader to understand him/her as a writer.

In addition to the above, an important part of the present chapter was spent reviewing the common definitions of MD as well as the definitions of the concepts parallel to MD. The main difference between MD and the other terms is that the former is broader as it can accommodate all of those terms and includes textual and interpersonal MD while the latter concepts are narrower with only textual elements.

A crucial part of this chapter is the proposed definition of MD. The elements of this have been selected from those used in the different definitions reviewed and are thought to reflect most of them. Besides, the proposed definition hints at the functions of MD and possible ways of identifying and classifying various types of MD. This

definition hopefully will help reduce the fuzziness of MD and provide a basis for identifying and classifying types of MD.

Chapter 4: A Review of MD Related Studies and a Proposed Typology

4.1 Introduction

Throughout the relatively short period since the coinage of MD, several typologies have been suggested and applied. They vary in their structure and components; some contain major categories and subcategories while others contain only subcategories. The two major categories have been named as textual metadiscourse ‘TMD’ (sometimes called informational or interactive metadiscourse), and ‘IMD’ interpersonal metadiscourse (sometimes called attitude metadiscourse). The subcategories under each major category vary from one typology to the other and show considerable overlapping because of the difficulty of arriving at a consensus among analysts as to how to define, identify and classify MD features. Before reviewing the different typologies, a short discussion of the main functions of MD in general and the two major categories of TMD and IMD in particular would seem to be helpful for setting up the proposed typology. Also the different applied MD studies will be discussed to find the research gap in terms of two variables (language and genre differences).

4.2 Functions of MD in Written Discourse

The main function of metadiscourse is that it helps writers to make explicit to the reader what they are doing at different points of the discourse, whether introducing a whole text or a part of it, introducing or changing topics, expressing attitudes etc. MD can also be used to show whether what the writers are asserting is certain or not, to define terms, to acknowledge difficulties and to note the existence of a reader.

Commenting on the functions of MD, Nash states:

It helps the reader or listener to form a mental map of the discursive country I propose to travel through. (Nash 1992, p100)

From a writer's perspective, Intaraprawat and Steffensen (1995) suggest that MD helps writers identify infelicities in their developing text when writers fully understand the meaning and rhetorical function of MD markers (p254), and this allows them to increase the clarity of their writing, text readability and the likelihood of readers understanding the message. It also helps them consider their prose objectively, comment on it in various ways and reflect their convictions more accurately in a logically connected manner. This network of relationships and interactions has been expressed by several sources describing MD as a way of talking with the reader (Williams 1981; Vande Kopple 1985; Crismore 1989) involving him/her in an implicit dialogue and moving him/her through different MD expressions (Intaraprawat and Steffensen 1995).

4.2.1 Functions of Textual MD (TMD)

The major function of Textual MD (TMD) is to show how individual propositions are linked to form cohesive and coherent text. As for the other functions, it has been

pointed out (Crismore et al., 1993) that TMD signals the structure of the text (sequencers, topicaliser, pre/review), defines concepts (code glosses), and helps readers classify and organise the text in order to construct its textual meaning. From a psycholinguistic point of view, this type of MD aids readers' limited memory capacity by highlighting to the reader the areas that are more important (Meyer and Rice 1982, p156). Similarly, Crismore (1982, p4) refers to this function as a system of textual cues for readers through which they discover the degree of importance of the discourse property and understand what is meant by what is said.

4.2.2 Functions of Interpersonal MD (IMD)

Interpersonal metadiscourse (IMD) carries truth values that do not relate to facts about the world but to people's real personalities, their true evaluation of the propositional material, their understanding of their role as participants in the communication process, and their hopes for the sort of reaction that readers might show (e.g. hedges, emphatics, evaluatives). The essential function of IMD is to help writers assess and readers grasp that assessment of certainty or uncertainty of the propositional content. Furthermore, they disclose the writer's attitudes towards the content and the reader and through their comments create a dialogic atmosphere.

4.3 A Critical Review of Metadiscourse Typologies.

4.3.1 Joseph Williams (1981)

Williams (1981) proposed an early typology of MD which was later used as the basis for other typologies. Williams' typology comprises six subcategories attached to three categories as follows:

- Hedges and Emphatics:

They express the writer's certainty about the truth of the content, e.g. hedges: *might, seem, possibly* and emphatics: *of course, it is obvious that*.

- Sequencers and Topicalisers:

They signal textual relationships showing how texts are organised, e.g. sequencers: *next, first, the second point is, however, therefore*, and topicalisers: *in regard to, turning now to*.

- Narrators and Attributors:

They signal the origin or source of what is presented. Williams makes a clear distinction between narrators which reflect the author as a source and attributors that refer to the third person, e.g. narrators: *I was concerned, I have concluded, I think, I believe*, and attributors: *have been observed, have been determined, according to..., X has maintained that*.

From the point of view of terminology, Williams does not use the two major categories of MD, namely TMD and IMD, though some subcategories of these types are included. In addition, some important categories are missing such as code glosses, evaluatives and illocutionary markers. In spite of that one can say that the typology partly matches the elements of discoursing, taking the truth value of propositions, and the writer/reader relationships and interactions which are mentioned in Williams' definition of MD (refer to the previous chapter).

4.3.2 Typologies Used by Avon Crismore

Crismore (1983) suggested a typology consisting of two main categories, each of which comprises several subcategories. The first main category she calls informational metadiscourse which 'gives information about the primary discourse' (Crismore 1983,

p12). She mentioned that this kind of MD is used by a writer to signal meaning explicitly or implicitly.

- Subcategories of Informational MD

(a) At a global level

- Goals: *The purpose of this unit is... (these are also previews)*
- Reviews: *we have in this work attempted...*

Crismore also uses the term 'pre-plans' (previews) for expressions like: *this chapter is about* and 'post-plans' (review) or expressions like *we have looked so far in this chapter at xx*. In the typologies that are used on subsequent occasions, she adopts previews and reviews rather than pre-plans and post-plans.

(b) At a local level:

- Topicalisers: (are used to include local shifts in topic) *let us now turn to...*

The second main category is attitudinal metadiscourse (AMD) which, in a general sense, is a writer's explicit or implicit signal of 'his attitude toward the content or structures of the discourse and toward the reader' (Crismore 1983, p13).

- Subcategories of Attitudinal MD

(a) Saliency markers: *equally important, still more important*. They show the importance of what they refer to.

(b) Emphatics: *of course, in fact*. They reflect the writer's certainty.

(c) Hedges: *perhaps, probably*. They reflect the writer's uncertainty.

(d) Evaluatives: *unfortunately, luckily, I think it is interesting*. They show the writer's attitude towards a fact or idea.

A distinct feature of this typology is that it explicitly distinguishes the types of MD that operate at a local level from those that operate at a global level. In this typology, we find one of the essential characteristics of MD, particularly the role of the author and the writer/reader relationships. The author's role is reflected in at least two ways: in the way he or she makes the structure of his or her discourse clear to the reader and how he or she expresses his or her attitudes and subjective evaluation of the text. To a great extent the typology matches Crismore's definition of MD, given in the same work, since it reflects the author's intervention to guide the reader and help him understand the message. However, the typology is very concise and so lacks several other subcategories such as code glosses, sources, illocutionary markers etc. In addition to that some of the subcategories overlap. The examples show that, for example, pre-plans may be classified as goals and topicalisers may be classified as pre-plans.

Crismore and Farnsworth (1990) developed a typology for their study of 'Metadiscourse in Popular and Professional Scientific Discourse' using Halliday's textual and interpersonal macrofunctions of language and Vande Kopple's (1985) typology of MD (to be mentioned in the following section). It consists of two broad categories, the textual and the interpersonal, and four broad ones; code glosses, modality markers, attitude/evaluative markers, and commentary. With respect to terminology, they called announcements previews, illocutionary force indicators, action markers and attributer sources. They also introduced new elements such as graphics, captions and Latin terminology. The typology is as follows:

(a) Code glosses (textual): *I will call them x, in short,*

(b) Modality markers (interpersonal)

- Hedges: *probably, suggest that*
- Emphatics: *obviously, undoubtedly*

(c) Attitude/evaluative markers: *it is regrettable that, we are struck by*

(d) Commentaries (interpersonal and textual): *The distribution of each ally may now be considered in turn, we will present these data in two ways, see discussion above*

- *General commentary:*

- Informative: *except for colour, varying from island to island*
- Previews; *we will discuss distribution in the next section*
- Reviews: *as suggested above*
- Action markers: *I now report that*

- *Scientific Commentary:*

- Quantitative: *measure 19*
- Source: *Gould, Woodruff 1974*
- Graphics: *table 5*
- Latin terminology: *c. bandalli*
- Captions: *converted from original data in micrometer units*

Crismore and Farnsworth mentioned that they reduced modality markers to two only (emphatics and hedges) and expanded the scope of commentary to suit the genre of scientific writing.

The structure of this typology differs from the former ones in a number of ways. Firstly, the subcategories of TMD and AMD are combined. Secondly, the scope of commentary has been expanded to include general commentary and scientific commentary. Thirdly, with the exception of sources, all other subcategories of scientific commentary are novel, that is they have not been used in other typologies.

According to the authors, the modifications are made to suit the genre and allow them to refute the myth that views 'professional scientific writing as the impersonal statement of facts that all add up to the truth' (p.118). Looking more thoroughly at the typology, one can say that it is a good reflection of the authors' definition of MD, especially the point of the presence of the author and guiding the reader rather than informing him. Although they include a subcategory of informatives, the examples they use do not differ from those of the other categories in the sense that they reflect MD properties. However, as in other studies by Crismore, the issue of saying that MD does not inform and at the same time including informatives as subcategories needs to be justified.

In Crismore's (1990) study of MD and discourse processes, MD is classified into two major categories: textual (informational) and interpersonal (attitudinal). The textual category serves the informational referential function of MD that relates to how the primary message should be taken in terms of its content, structure and the author's purposes and goals. Under this main category Crismore includes the following subcategories:

- Announcements of main ideas: *my main idea is*
- Rationales: *the reason for x is that*
- Purposes: *my purpose for you is*
- Strategies: *in this section I will trace the history of x*

Interpersonal MD, on the other hand, serves expressive attitudinal function by indicating the writer's perspective and attitude towards the primary discourse and the reader. Under this main category the following subcategories are included:

- Hedges: *probably, it is possible that*
- Emphatics: *surely, it is certain that*
- Evaluatives: *fortunately, most important*

The terms used in this typology serve functions similar to those indicated by her previous terms. For example announcements are similar to topicalisers, strategies are similar to previews. Furthermore, strategies and purposes can be seen as expressing the writer's goals or pre-plans. This clearly shows the fluid state of MD categories and subcategories. As regards the relevance of the terms, it can be mentioned that the subcategories serve the specific purpose of the study to investigate the effects of the inclusion of TMD, and IMD on retention of information from social studies passages and the students' attitudes towards the subject matter. It can be noted that, to a great extent, the subcategories match the elements of the definition adopted in this study such as discourse about discourse, guiding the reader and focusing on organisational aspects of the discourse.

A couple of studies (Crismore 1983; Crismore et al 1993), base their typology on Vandekopple's (1985) typology (see following section) but make some modifications. While retaining the two main categories of TMD and IMD they reorganise the subcategories and rename some of them. The TMD category has been split into two categories: textual markers and interpretive markers. In the IMD, the attitude markers and commentary have been retained, whereas the validity markers have been subdivided into three subcategories: hedges, emphatics and attributors.

Textual MD (used for logical and ethical appeals)

- *Textual markers*: logical connectives, sequencers, reminders and topicalisers
- *Interpretive markers*: code glosses, illocution markers and announcements

Interpersonal MD (used for emotional and ethical appeals):

- *Hedges* (epistemic certainty markers), *certainty markers* (epistemic emphatics), *attributors*, *attitude markers and commentary*.

The structure of the typology shows great harmony with the definition adopted by the authors, since the categories and subcategories do not add anything to the propositional content, but aid the reader to interpret, evaluate and understand the structure of the information given. As can be seen, the categories and subcategories are not accompanied by examples. The authors, it seems, rely on the examples given by Vande Kopple in the typology they followed.

The above discussion has revealed that different studies use different MD typologies. There is no consensus on one standardised typology and there does not seem to be any tendency to establish one. Each study uses a typology thought to be relevant to its own purposes and context. The disparity of the subcategories is very clear. As mentioned elsewhere, different subcategories have been labelled differently although, in essence, they refer to the same thing. Furthermore, whereas some subcategories like logical connectives can be realised by somewhat fixed expressions such as some of the cohesive ties, many others do not presuppose one-to-one correspondence with their realisations, i.e. we cannot predict with confidence the expression that may be used to realise any subcategory.

4.3.3 Vande Kopple (1985)

Among the widely used typologies is the one suggested by Vande Kopple (1985) in his very famous article ‘Some Exploratory Discourse about Metadiscourse’. Drawing on

the typologies suggested by Lautamatti (1978) and Williams (1981), Vande Kopple suggests the following typology:

- Connectives:

They smoothly guide the readers through the text and ‘help them to construct appropriate representation in memory’ (p83). This category includes most of what other typologies group under textual metadiscourse:

(a) Sequencers: These include items like *first, second, in the second place*, as well as items indicating logical and temporal relations such as *however, nevertheless, as a consequence, at the same time*

(b) Reminders: They ‘remind’ the reader of what has been previously stated, such as *I noted in chapter one*, and announce what will follow as in *as we shall see in the next chapter, what I wish to do now*. It can be seen that these examples can be classified as review and previews. Besides, most people would expect to be reminded of something they already know about (review) but not something they will know about. It looks as though the writer is reminding himself/herself rather than the reader.

(c) Topicalisers: They attract attention to particular parts such as the main topic of a sentence, paragraph or a section e.g. *there are, as for, in regard to*.

- Code Glosses:

They help the reader to comprehend the meaning of terms, words, elements, etc. e.g. *x is, x can be defined as, in other words*. While other scholars consider code glosses as a subcategory of TMD, Vande Kopple treats them as a separate category although it does not have subcategories. The same applies to the following subcategory.

- Illocutionary Markers:

These explicitly mark the speech act that is being performed e.g. *I hypothesise that, to sum up, we claim that, I promise to*.

- Validity Markers:

These express the writer's view of the validity of the propositional content and include the following subcategories:

- (a) Hedges: They help express necessary doubts: *perhaps, may, might, seem*.
- (b) Emphatics: They help state what the writer really believes and what he wants his readers to believe: *clearly, certainly, obviously, it's obvious that*
- (c) Attributors: They belong to the validity markers on condition that 'we use them to try to lead readers to judge or respect the truth value of our propositional content as we wish them to' (p.84) e.g. *according to Einstein*
- (d) Narrators: They show readers who said or wrote what: *according to James (xxxx), Mr. Wilson announced that, the principal reported that*.

The last two subcategories 'attributors' and 'narrators' essentially refer to the source of the information provided and can be included under one sub-category namely 'sources' in the proposed typology, similar to the 'evidentials' in Hyland and Tse's (2004)

typology. However, Vande Kopple, following Williams (1981), considers them as two different subcategories.

- Attitude Markers:

These reveal the writer's attitude towards the propositional content, e.g. *surprisingly, I find it interesting that, It is alarming*. These examples signal a kind of evaluation on the part of the writer (evaluatives in the proposed typology) and can be called 'attitude markers' in Hyland and Tse's (ibid) typology.

- Commentaries:

These address the readers directly and attempt to involve them in an implicit dialogue with the writer. They also comment on the reader's probable moods, views, or reactions to the writer's propositional content: *most of you will oppose the idea that*, or recommend a mode of procedures: *you might wish to read this chapter first*, or let readers know what to expect: *you will probably find this material difficult at first*, or comment on the relationships to the writer: *my friend*. Commentary is a very broad category and almost all features of MD can be considered as some kind of commentary, whereas the examples used by Vande Kopple, and indeed by several other researchers, represent direct address to a certain audience, and can be called 'engagement markers' in Hyland and Tse's typology.

Vande Kopple first lists the above categories together. He then mentions that five of them may tentatively be considered as belonging to the main interpersonal category of interpersonal or attitudinal MD (IMD or AMD). The five categories are: illocutionary markers, validity markers, narrators, attitude markers and bits of commentary. However,

Crismore and Farnsworth (1990), and Crismore et al. (1993) group the seven categories into two main categories:

- Textual metadiscourse (TMD) that includes text connectives, code glosses, illocutionary markers, and narrators.
- Interpersonal metadiscourse (IMD) which includes validity markers, attitude markers and commentary.

It can be seen that Vande Kopple's typology outlines a real representation of his definition of MD and its overall functions. They help the reader understand the structure of the text, evaluate and react to it. Moreover, they signal the author's presence and his attitude towards the content and the reader as well as his relationships with them. Nevertheless, like other typologies, some overlap among the categories can be noticed, e.g. attributors and narrators can be grouped as sources, reminders and illocutionary markers which can be named as self-references in the present study and self-mentions in Hyland and Tse's (2004).

4.3.4 A Typology of Metadiscourse Based on Classical Rhetoric (Nash 1992)

The relationship between MD and rhetoric was explained in chapter two where a special reference was made to Nash's theory that MD can be related to the division, by Greek rhetoricians, of oratory into 'taxis' and 'lexis'. On the basis of that division, Nash suggests a typology of MD consisting of two major categories with several subcategories under each of them. He calls the first group (the taxis) 'tactical MD' and the second group (the lexis) 'lexical MD'. These two main categories correspond to TMD and IMD respectively. Elaborating on TMD, Nash states that it 'tells audiences

what point they have reached, in the narrative or exposition, how they got there and perhaps – most important of all – what they are to expect’ (p.101). He also points out that an example of a tactic that takes different stylistic versions on different occasions and modes is the use of reminders and announcements of themes to ‘buttonhole the audience’ as in *here’s one I bet you haven’t heard* (when introducing a joke) and *you know that book you lent me last week*. The typology proposed by Nash is as follows:

- Tactical Metadiscourse

- Preview: *in this lecture I propose to discuss....*
- Topic numeration: *first we shall..... then we proceed to..*
- Topic shift: *now as to, turning to another matter....*
- Review: *we have seen how..., up to this point I have....*
- Conclusion: *to sum up..., let me end by saying...*
- Forecast: *in my next chapter..., when we resume...*

- Lexical Metadiscourse

- Limiters: *to this extent only..., up to a point..., historically speaking, in this particular instance...,*
- Hedges: *admittedly..., as far as I know..., by all accounts..., most people would agree...,*
- Emphatics: *of course..., quite obvious..., it must be stressed that..., without the least doubt...,*
- Evaluatives: *strangely enough..., by a stroke of good fortune..., I am happy to say..., naturally...*
- Formulators: *so to speak..., what might be called..., for which I propose the name...,*
- Appeals: *how are we to read this? One might ask what sort of response is expected.*
- Directives: *consider this..., let the reader be advised..., we ought not to accept...,*

- Asides:

- Explications: *by which I mean..., in other words...,*
- Parenthetical comments: *by the way..., I might add..., incidentally...,*
- Promissory asides: *as we shall presently see..., in due course..., for the time being...*

The labels used by Nash for his subcategories are somewhat different from those used by other researchers. However, his typology covers most of the MD types, with the exception of attributors and narrators. But Nash's typology does show internal and external disagreements. On the one hand, there is overlap between the subcategories themselves as in the examples of the previews/forecast (which alert the reader as to what the writer is going to do in the following paragraph and can be grouped as pre/reviews in the present study, e.g. *in this lecture, in my next chapter*), as well as of formulators/explications in asides (which can be named as code glosses, e.g. *For which I proposed the name, in other words*), show. On the other hand, there is an external disagreement with terms used by other writers like Williams (1981) and Vande Kopple (1985). Although Nash's typology is extended and includes most of the MD categories it needs refinement and precision, a point that he seems to recognise as 'the terminology is purely descriptive and does not purport to represent a standard model' (p.114)

4.3.5 Paul Jude Bauvais (1986)

A different typology of MD was established by Beauvais (1986) on the basis of the work of Austin (1962) and Searle (1969, 1976) regarding speech acts and illocutionary markers.

In his typology Beauvais distinguishes between primary acts and secondary acts.

Primary Expositive Illocutionary Acts:

They “can be expressed in canonical form with explicit performative structures that use first person subject pronouns” (p113). These acts include the following subcategories:

- (i) Simple expositive acts that Beauvais considers as the most basic primary expositive illocutionary acts the function of which is to state. They consist of:
 - (a) explicitly performed acts: *I state/note/assert/I must note/I should state*
 - (b) partly explicit indicators: they “identify an illocutionary act without referring to the person who performed the act” (p114). They include clauses as in *it is notable*, adverbs as *notably*, and other words, clauses or phrases that do not identify an attributive subject.
- (ii) Complex expositive acts: “they convey features of the expositives that supplement the basic information that a speaker/writer is stating something” (p115).
 - (a) Relational expositive act “which identifies sequential and causal links among passages of propositional discourse” (p115) e.g. sequential: *I first/also state*, causal: *I therefore assert*, other possible explicit forms are: *I first would state, I also should note, I therefore must assert, my first point is...*

Within this subcategory, Beauvais considers the following forms partially explicit because the attributive subject is not stated but can be inferred by the listener/speaker e.g. *first she does not like lobster, also it rained all day, therefore we should not go fishing*

- (b) Evaluative expositive acts: they include “those illocutionary markers that indicate the speaker/writer’s assessment of propositional material”

(p117). Also, they express the speaker/writer's attitude concerning the validity of propositions, and they also may express other judgmental reactions to propositions" (ibid). e.g. *I believe/doubt/like/dislike*. It is also possible to express them by near-performatives: e.g. *I am convinced, I have no doubt, my belief is*; examples of other reactions are: *I am surprised/amused/concerned*.

Beauvais mentioned that "the range of evaluative expositive acts is better suggested by partially explicit forms" (p118). Examples are:

- Validity clauses: *it is certain/possible/doubtful*
- Validity phrases: *in fact, without a doubt*
- Validity adverbs: *certainly, possibly, arguably*
- Validity modal verbs: *must, may, might*
- Other reactions partially explicit forms are: (clauses) *it is confusing, it is surprising*.
- Adverbs: *amusingly, disturbingly*

(iii) Commissive expositive acts: they "indicate that the speaker/writer is committed to performing a specific expositive act concerning specific propositional material in the discourse and or text that follow the commissive act" For example, *I will state/note*.

(a) Relational commissive expositive acts: in addition to the above-mentioned general function of the commissive expositive acts they "also link the act and material to other passages of the discourse and text"

(p120) Examples are:

- Sequential: *I first/also will/note/state/consider*
- Causal: *I therefore will/note/state/consider*

(b) Possessive pronominal clause (explicit): *my first subject will be, my next topic is*, (partially explicit) *the next topic to be considered is..., the next subject will be...* (adverbial) *first, the causes of this growing problem....*

(c) Reiterative expositive act: it "relates an expositive act and its corresponding propositional or referential material" (p121), examples are:

- Near performative clause with verb possessing a past aspect: *I stated that the Red Sox won the pennant in 1975, I have noted the causes of the civil war.*
- Relational forms: *I first stated, I also have noted, I therefore asserted*
- Partially explicit: *having first considered, also having noted.*

Secondary Expositive Illocutionary Acts

They differ from the primary acts in that “they attribute the act of stating to someone other than the speaker/writer of the discourse or text” (p122). They consist of two broad categories, the third person category and the second person category.

Third Person Category:

- (a) Third person simple expositive acts: *she stated, she noted, Smith has observed.*
They also imply third person, *it is believed, it is considered unfortunate,* or use third person relational forms: *she also believes, he therefore dislikes.*
- (b) Third person commissive expositives: Beauvais mentions that although it seems impossible to commit other people to do things, these acts do exist: *she will state, he will note, Smith will consider.*
- (c) Third person relational commissives: “They express the entire range of sequential and causal relations that are indicated by their first person counterparts” (p128) e.g. *she first will state, he also will note, Smith therefore will consider.*
- (d) Third person reiterative expositives: Beauvais considers all third person expositive acts as reiterative “because they all identify illocutionary acts that first were performed prior to the discourse or text in which they are restated...” (p129) e.g. *she/he/Smith/stated/noted/has observed, she first stated, he next observed, Smith consequently noted.*

Second Person Category:

- (a) Second person attribute statements: the speaker/writer attributes a statement to listener/reader through the use of the second person, e.g. *you note, you assert, you may not tone/assert.*
- (b) Second person partial explicit which takes the imperative form: (you) *note/consider.*
- (c) Second person relational expositive complex acts: *you first/also/therefore/may, note/consider/assert.*
- (d) Second person evaluative expositive: they predicate the listener/speaker's assessment of propositional material and reflect either attitudinal or judgmental reactions: (validity) *you may believe/doubt*, (reaction) *you may like/dislike*, (relational forms) *you first may believe, you also may doubt, you therefore may like.*
- (e) Second person commissives: *you will note the cause of the civil war, you must consider.*
- (f) Second person relational commissives: (explicit) *you first will note, you also must consider, you therefore will note*, (partially explicit) *first note, next consider, therefore note.*
- (g) Second person reiteratives: *you noted, you have noted*
- (h) Second person relational reiteratives: *you first noted, you also observed, you therefore have noted.*

Several points can be mentioned about Beauvais' typology. An initial observation is that the typology can be theoretical in the sense that it is based on 'made up' examples which have not been applied, nor have they been taken from samples of authentic discourse. They are oriented towards 'what' may be said. One can also note that very few studies have mentioned this typology and none of them have applied it. This makes an important property of the typology, namely generalisability, a matter to be seen. Another point is that, whereas Beauvais claims precision for his typology, it is difficult to avoid multifunctionality and confusion in interpreting the subcategories. The

category of ‘evaluative expositives’, for example, performs the act of stating as in *I believe*, but they are also emphatics: *certainly, certain, must* and evaluatives: *I am amused/surprised, surprisingly*. However, as Crismore observes, Beauvais’ typology remains a useful contribution to a theory of MD.

4.3.6 Typologies of Hyland (1998) and Hyland and Tse (2004)

Hyland’s (1998) typology, a refined extension of Crismore et al (1993), takes up the distinction between two types of MD and further subdivides the functions of these.

In his work on academic texts, Hyland(1998), following Crismore et al. (1993), classifies metadiscourse into two types, textual metadiscourse, - “devices which allow the recovery of the writer’s intentions by explicitly establishing preferred interpretations of propositional meanings” (p442) - and interpersonal metadiscourse – which “alerts readers to the author’s perspective towards both the propositional information and the readers themselves” and as such is “essentially interactional and evaluative” (p443).

But he modifies their content to better reflect the object of his study. In his typology, textual metadiscourse contains logical connectives (expressing semantic relations between main clauses, renamed ‘transitions’ in their 2004 paper), frame markers (explicitly referring to discourse acts or text stages), endophoric markers (referring to information in other parts of the text), evidentials (referring to sources of information from other texts), and code glosses (helping readers grasp meanings of propositional material). Interpersonal metadiscourse covers the categories of hedges (withholding writer’s full commitment to statements), emphatics (emphasizing force or writer’s certainty in message, renamed ‘boosters’ in 2004), attitude markers (expressing writer’s

attitude to propositional content), relational markers (explicitly referring to or building relationship with reader, renamed ‘engagement markers’ in their 2004), and person markers (explicitly referencing to author, renamed ‘self mentions’ in 2004).

Hyland’s taxonomy of textual and interpersonal types of written metadiscourse seems to be very useful when analysing academic written text produced by postgraduate students and professional writers. However there are overlaps between the subcategories as in other typologies. For example, the two subcategories, a subclass (‘sequencing’) of the frame markers and endophoric markers refer to the text of the information provided or that will be provided (e.g. *in chapter x, in section x, in page x*) and can be included under one sub category namely ‘pre/review’ in the present typology. However, Hyland and Tse (2004) consider them as two different subcategories.

All the subcategories in their typology are as shown below:

Table 4- 1: Hyland’s Classification of Metadiscourse in Academic Texts

<i>Category</i>	<i>Function</i>	<i>Examples</i>
Textual resources	Help to guide reader through the text	
Logical connectives/Transitions	Express semantic relation between main clauses	In addition/ but/therefore/thus/and
Frame markers	Refer to discourse acts, sequences, or text stages	Finally/to repeat/our aim here/we try
Endophoric markers	Refer to information in other parts of the text	Noted above/see Fig 1/table 2/below
Evidentials	Refer to source of information from other texts	According to X/Y,1990/ Z states
Code glosses	Help readers grasp functions of ideational material	Namely/eg/in other words/ such as
Interpersonal resources	Involve the reader in the argument	
Hedges	Withhold writer’s full commitment to proposition	Might/perhaps/it is possible/about
Emphatics/Boosters	Emphasize force or writer’s certainty in proposition	In fact/definitely/it is clear/obvious
Relational markers/Engagement markers	Explicitly refer to or build relationship with reader	Frankly/note that/you can see
Attitude markers	Express writer’s attitude to proposition	Surprisingly/I agree/X claims
Self mentions	Explicit reference to author(s)	I/we/my/mine/our

Source adapted from Hyland (1998 p442) and Hyland and Tse (2004)

In addition to the overlapping of subcategories in Hyland’s typology, some of Hyland’s subcategories are much more diverse functionally. For example, ‘frame markers’ including announcing goals such as *our aim here*, sequencers such as *first, then*, and topic shift such as *with regard to, return to*; it also comprises labelling stages such as *at this point, at this stage*. And the category of ‘logical connectives’ (‘transitions’ in his 2005b book), for example, refers to items, additives (*also, furthermore, moreover*),

concessives/contrastive relations (*however, on the other hand*) and resultives/rationales (*the reason for this is, as a result of this*) while they are separated as independent subcategories of MD in other typologies as shown below:

Table 4- 2: A Summary of MD Subcategories used in the Hyland(1998) and Vande Kopple (1985)

Hyland (1998)	Vande Kopple (1985)	Subcategories used in Others	Examples
Logical connectives	Text connectives	Enumerators/additives/concessives	In addition/but/therefore/thus/and/because
Frame markers	Action markers, text connectives/sequencers	Logical connectives/topicalisers/theme indicators	Finally/to repeat/our aim here/in sum/overall
Endophoric markers	–	Preview/review	Noted above/see Fig. 1/below/the following
Evidentials	Narrators, Attributors	Sources	According to/Fraser 1990/X states
Code glosses	Code glosses (defining, explaining, limiting)	Formulators	In other words/such as/namely
Hedges	Modality markers: hedges	Validity markers	Might/perhaps/it is possible/would/seem/tends to
Emphatics	Modality markers: emphatics	Validity markers	Indeed/in fact/it is clear/obvious
Attitude markers	Attitude markers, evaluatives	Stylistic disjuncts	Surprisingly/I agree/X claims/it is fortunate that
Relational markers	Attitude markers, Commentary		Frankly/note that/ you can see
Person markers	–	Illocutionary markers	I/we/my/mine/our

The following section ends with summary tables which demonstrate the different subcategories, and those that are similar to them, with corresponding examples to provide a summary of the review of MD typologies. Also the tables are intended to show the massive overlap among the different subcategories. It is hardly surprising to see considerable repetitions across the subcategories as well as their realisations.

Table 4- 3: A Summary of TMD Subcategories used in the Reviewed Typologies

Subcategories	Similar to column 1	Examples
Sequencers	Enumerators, logical connectives, frame markers	First/next/the third point is/however/nevertheless/ as a consequence
Topicalisers	Theme indicators, topic shift, frame markers	In regard to/turning now to/there are/as for
Summarisers	Asides, conclusion, frame markers	In other words/to summarise/in sum/ briefly
Concluders	Conclusion, frame markers	The conclusion is/to sum up
Logical connectives	Sequencers, transitions	So/therefore/thus
Superstructure signals	Goals, purposes, preview, review, announcements, reminders, strategies, conclusion, frame markers	Our premises are/this chapter is about/in this section we attempted to/in conclusion
Limiters		Only/alone/to this extent/mainly/historically speaking
Additives	Transitions	Also/and/moreover
Antithetics	Transitions	Instead/on the contrary
Concessives	Transitions	Yet/in spite of/however/but
Replicatives	Transitions	Alternatively/on the other hand
Preview	Forecast, promissory asides, pre-plan, frame markers, endophoric markers	This chapter is about/in this section I intend to
Review	Post-plan, reminder, announcement, frame markers, endophoric markers	We have in this section suggested/up to this point I have
Goal	Preview, purpose, strategy, frame markers	The purpose of this unit
Announcements	Superstructure signals, previews, reviews, reminders, strategies, topic shift, frame markers	In this section I am going/we now turn to/this part is about/the purpose of this unit
Code Glosses	Formulators	X is, Y may be defined as, what might be called
Rationales		The reason for this is
Topic shift	Topicalisers, theme indicators, frame markers	Now as to/turning to another matter
Narrators	Sources	The principal reported that/Mr. Wilson announced that
Attributers	Sources/evidentials	According to X/X suggested that

Table 4- 4: A Summary of IMD Subcategories used in the Reviewed Typologies

Subcategory	Similar to column 1	Examples
Hedges	Validity/modality marker	Might/seem/possibly
Emphatics	Validity/modality markers/boosters	Of course/certainly/obviously
Direct relevance	Saliency/attitude markers	Important/this is more significant/equally important
Transitional	Formulators/asides/parenthetical comments/engagement markers	By the way/incidentally/so to speak
Inferentials		Otherwise/in that case
Formulators	Transitional/asides/comments/engagement markers	So to speak/as it were/better rather
Pronouns of self reference and audience reference	Directives/commentaries/engagement markers	Consider this/you may like to/it is my contention
Evaluatives	Stylistic disjuncts, attitude markers	Unfortunately/strangely enough/seriously
Appeals	Rhetorical questions/engagement markers	How are we to do this?/what can be expected from?
Illocutionary force indication	Action markers, illocutionary markers, self mentions	I warn you/I assume/you are advised to

4.4 A Critical Review of Typologies of Parallel Terms Other than MD

This section presents a critical review of the typologies that appear in the studies that use terms which are either similar to MD or can be counted as such. On the one hand, this will enable us to see the similarities and/or differences between these neighbouring areas and, on the other, will help describe and characterise MD with more details. It is also hoped that the social and interactive nature of MD may be further illuminated.

4.4.1 Typologies Used by Meyer

a. A Typology Based on Signalling

Meyer's (1975) is one of the early studies that used the term 'signalling', identifying four major types:

1) Specification of structure of relation in the content structure: words that explicitly signal discourse patterns: *problem, solution*.

Linking words: they show the structure of the text e.g. *one, two, the other*.

Meyer notes that these two types of words are sometimes used together as in *Two problems exist. One is the problem of money and the other is the problem of motivation*

2) Prospectively revealed information abstracted from the content occurring later in the text.

They represent abstracts of information to be presented later in the text. Examples of these are usually found in titles and sentences that introduce paragraphs, sections or passages:

Topic enumerators: *These three levels of analysis are syntax, text, information processing and social context.* A premature statement of issues that will appear in a later stage in the text: *These problems must be resolved within the next six months.*

These correspond to previews in MD. One may observe that the examples used by Meyer show that ‘linking words’ in the first category seem to overlap with ‘enumerators’ in this category.

3) Summary statements: these are also abstracted information but they also function as summary statements placed at the end of paragraphs or texts: *in short, a wonderful vacation is available for those with time, money, and the proper equipment.*

4) Pointer words: they signal how writers view an issue or idea, and signal their attitudes and evaluation. In MD these are categorised as ‘evaluatives’ or ‘attitude markers’ e.g. *This is an important point. Unfortunately not many legislators agree with me that the ERA is a dead issue. Unfortunately a number of linguists agree with Chomsky that language reveals most clearly that it is rule-governed and not arbitrary.*

These four major categories match Meyer’s definitions of signalling as they give emphasis to certain aspects of the text, point out the structure of the content and reflect the author’s idea of the importance of the content expressed in his text. Furthermore,

the first three categories can be considered as TMD, while the fourth one, 'pointer words', may well qualify as IMD.

b. A Typology of Rhetorical Relationships (Meyer and Rice 1981)

For their study to examine the reader's comprehension and identification of the overall logical organisation of text, Meyer and Rice (1981) use five basic but general groups of rhetorical relationships. These groups collapse those of Grimes (1975) and the previous relevant work of Meyer (1975, 1981). They (Meyer and Rice 1981) mentioned that the five groups reflect the organisational plan of texts and can be used for the top-level and lower-level structures. The five groups are:

Collection:

Relations which show how ideas or events are related together forming a group on the basis of some commonality.

Causal:

Relations which show causal relationships (covariance) between ideas, where one idea is the antecedent and the other is the consequence or effect.

Response or problem and solution:

(also remark and reply and question and answer formats): this is similar to the problem and solution in that the problem- is an antecedent for the solution.

Comparison:

Points out differences and similarities between two or more topics.

Description:

Gives more information about a topic by presenting attributes, specifics, manners, or settings.

As mentioned above, these groups are general and the perception of these plans by the reader as Meyer and Rice suggest "depends on both his strategies and on the cues and signals given by the writer" (p57). To them the cues are represented by the four types

of signalling outlined in Meyer (1975) and the definition of ‘signalling’ that accompanies them. The five categories stated above correspond to their definition since they do not add new content, but point out aspects of the structure of the content.

4.4.2 A Typology of Non-topical Materials (Lautamatti 1978)

As mentioned previously, in chapter 3, Lautamatti differentiates between two types of linguistic materials: the first type is ‘topical linguistic material’ which relates directly or indirectly to the topic of the discourse or text, and the second is the ‘non-topical linguistic material’ which is separate from the topical material and consists of different types (they relate to the secondary level of discourse and can be counted as MD).

Lautamatti identifies five types of non-topical material:

Discourse connectives (metatextual markers) *consequently, however, next, we shall discuss.*

Illocutionary markers: *to illustrate the point.*

Modality markers: *obviously, it seems possible.*

Attitude markers: *I would like to, it seems futile to*

Commentary markers: *dear reader, you may like*

These categories correspond directly to MD. The first two types can be classified as TMD while the last three clearly correspond to IMD. The examples given by Lautamatti seem to confuse modality and attitude markers that are considered as evaluatives, emphatics and hedges in most MD studies. In other words, according to Lautamatti’s examples, modality equals attitude markers and they look similar to

hedges and emphatics. However, in general terms the categories and their examples agree with Lautamatti's understanding of non-topical material.

4.4.3 A Typology of Gambits (Keller 1979)

For his psycholinguistic analysis of conversations, Keller (1979) uses a typology of seven major categories of semantic framing gambits, each of which comprises other subcategories as shown below:

Major semantic field indicators:

Congruent major semantic field: *I have a question on that, We'll first take up the questions, one answer to that would be.*

Incongruent major semantic field (a digression): *To get back to what I was saying.*

Initiation: *this reminds me..., speaking of..., before I forget.*

Returning to main topic: *in any case, to get back to, going back to.*

Various aspects of a topic:

A list:

Beginning: *first, to begin with, first of all.*

Middle: *second, another thing is, next*

End: *and finally, and the last thing*

A main aspect: *the main thing is, most of all, the real problem is*

A surprising aspect: *believe it or not, strangely enough, you may not believe this, but*

An unpleasant aspect: *to be realistic, let's face it, the catch is*

An emphasised aspect: *the main thing is, it bears emphasising, the most important thing is*

Opinion

Guessing: *my guess is,*

An opinion: *I am pretty sure that, I have reason to believe that.*

A conviction: *I honestly feel, I'm positive, without doubt.*

A personal viewpoint: *in my personal estimation, I personally believe, to my mind, the way I look at it*

A personal evaluation: *as far as I can tell, as I see it, it appears to me, to the best of my knowledge*

Personal circumstance: *in my case, what I'm concerned, for my own part*

Confidential information: *just between me and you, rumour has it, I hear from the grapevine.*

Action strategy:

A suggestion: *why don't you do the following, here's what you can do.*

A plan: *what we have in mind is, here's what we'll do.*

Subject expansion:

Expanding a point: *when it comes to, as far as that is concerned, in a case like this.*

Adding items: *another thing, what's more, I might add, and furthermore.*

Giving a reason: *the reason why, seeing as how, on account of this, for this reason.*

Explaining a result: *as a result, consequently.*

Positive contingency: *in case of, if and when, as soon as, by the time.*

Negative contingency: *barring the possibility, unless.*

Restatement: *what you're saying is, if I read you right, what I mean is, what I meant to say is.*

Appearance and reality: *you may think... but in fact, it may seem but actually, on the surface it appears as if... but the truth of the matter is.*

Subject evaluation:

Reservations: *yes, but consider, but don't forget, that's fine but, but the problem is*

Taking into account: *seeing as how, keeping in mind, allowing for the fact.*

Seeing the other side: *all the same, yet on the other hand, mind you though, but then again*

Argumentation:

Generalisation:

High frequency: *most of the time, again and again, time and again*

As a rule: *in general, ordinarily, as a rule*

Low frequency: *once in a while, every so often, every now and then*

Exceptions: *as an exception*

Examples: *as an example, for one thing, to give you an idea*

Summarising: *summing up, in short, to cut a long story short, in a nutshell*

The detailed list of categories and subcategories used by Keller coincides with his definition of gambits as “a certain set of signals in the conversationalist’s speech used to introduce level shifts within the conversation, or to prepare listeners for the next turn in the logical argument”(p220).

Moreover, the examples he uses for the subcategories may well fit into different subcategories of MD. However, there are some problems with the labels. Firstly, to match them with the subcategories, they have to be collapsed, because the examples show that several of Keller’s subcategories can be put under one of the main categories (see for example the subcategory, ‘opinion’ can be put under one of main category ‘opinion’). Secondly, he differentiates between two types of evaluation ‘personal evaluation’ in ‘opinion’ and ‘subject evaluation’, but the examples he gives show a different understanding from that of MD (e.g. the subject evaluation is different from the ‘evaluatives’ in the MD). The examples he uses for subject evaluation, especially reservation, are similar to claims and counter claims rather than ‘evaluation’. So, there is overlap and multifunctionality here. However, among other things, the extensive list of categories and subcategories does make it possible to identify them with most of the TMD and IMD subcategories, although they are originally used for conversational analysis.

4.4.4 A Typology of Metatalk (Schiffrin 1980)

The typology of “Metatalk” as proposed by Schiffrin (1980) consists of two main types of multifunctional-linguistic expressions.

The first type performs organisational functions and includes the following categories:

Explanation: I’ll answer it this way

Discourse brackets:

Paired discourse brackets (at the beginning and end of segments of talk): *there's only one thing, that's the one thing*

Initial brackets: *let me tell you*

Terminal brackets: *that's my opinion*

Reasons: *that's why*

Supporting evidence: *for instance, another thing*

Assertion: *the point is*

The second type performs evaluative but sometimes also organisational functions.

Evaluative brackets:

Argument: *I don't agree with that, that's not the point, that's my opinion*

Renewal brackets: *you said, what the hell do you mean?*

Repair: *I mean, I'm not arguing*

It can be noted that 'metatalk', as discussed by Schiffrin, and the examples taken from conversations show that metatalk is of little relevance to written discourse. In this sense it is different from Keller's 'gambits', some of the subcategories may be applied to both spoken and written discourse. In spite of that, the term 'supporting evidence' may be assigned to 'code glosses' in TMD while some of the subcategories in 'evaluative brackets', 'argument' and 'repair', may be assigned to IMD.

4.4.5 Typologies of Discourse Markers

a. Fraser (1990)

Fraser's (1990) approach to discourse markers (DM) restricts markers to those that signal "a sequential relationship between the current basic message and the previous

discourse” (p383), thus giving them anaphoric function. Fraser considers DM as “a well defined pragmatic category within the grammar of language” (ibid). He divides the sentence meaning into two main parts: the content meaning and the pragmatic meaning. He further describes the pragmatic meaning as consisting of three types of markers: basic pragmatic markers, commentary pragmatic markers, and parallel pragmatic markers.

Basic pragmatic markers:

They signal the force of the basic message e.g. *please*, and the performative expressions such as *I claim, I promise*.

Commentary pragmatic markers:

They encode an entire message – both the force and content – which constitute a comment on the basic message itself. e.g. *Frankly we are lost*.

He mentions that some of the examples in this category are quite complex, such as *to return to my original point*. The examples in this type are included in some MD subcategories as ‘evaluatives/attitude markers’ and ‘topicalisers/frame markers’.

Parallel pragmatic markers:

These also encode an entire message but cue separately from and in addition to the basic and/or commentary message(s) e.g. ‘damn’ in *take your damn shoes off the table*. They also include vocatives like *sir, my friend*.

In MD terms these may be included under attitude markers or engagement markers in Hyland’s classification which reflect interpersonal aspects and interactive writer/reader

relationships. To show that the three types of pragmatic markers may appear in the same sentence, Fraser provides this example: *Frankly, my friend, I admit that we are lost*. Fraser's approach to DM unambiguously corresponds to MD in a general sense. The markers he uses correspond to the following subcategories in MD. The first type, the basic pragmatic markers, corresponds to illocutionary markers (self mentions in Hyland's). The second (commentary pragmatic markers) includes evaluatives (attitude markers in Hyland's) from IMD and topicaliser (topic shift from frame markers) from TMD. The third (parallel pragmatic markers) can be matched with commentary (relational marker/engagement markers) in IMD. However, there are some confusing aspects to Fraser's classification in that some of the examples he put under 'commentary pragmatic markers' namely *to return to my original point* can be considered, in MD, as a topic shift in the topicaliser from TMD.

b. Redeker (1990)

Redeker proposes a typology with two main categories: markers of ideational structure, and markers of pragmatic structure, with subcategories under each of them. This is based on his central claim that “content structure and pragmatic structure are two complementary aspects of one paradigm of discourse coherence” (Redeker 1990, p369) and fit with his definition of DMs (discourse markers) as those that signal the relation between an utterance and its context.

Markers of ideational structure

They consist of the following subcategories:

Simple connectives: they include the simple superordinate 'that', the simple relative pronouns *that, who, which, whose, of which*, (and, or, are excluded).

Semantically rich connectives (clause initial connectives): this category includes conjunctions and adverbial connectives which signal semantic relations such as *but* (adversative conjunction), question words introducing embedded questions such as *what, how, why* etc.; temporal connectives like *when, as, while, meanwhile, then, next, now, before, after*; causal conjunctions like *because, so*.

Markers of pragmatic structure

They consist of the following subcategories:

Pragmatic uses of conjunction: conjunctions are considered pragmatic “if the semantic relation between the conjoined utterances did not correspond to the propositional meaning of the conjunction” (p372). Examples:

‘And so’ as in between successive elements... *and so...*, or to mark the speaker’s summing up or conclusion: ... *and he says you are gonna have to leave here. So he/he... kind of uhm kicks the guy out.*

‘Because’:... and uh it’s the next day’ cause the note says “tomorrow”.

‘But’: used to signal a return to the main topic after an aside or digression: ... *but uh they have little kids outside.*

Interjection: (that signals pragmatic relations)

Utterance-initial: *oh, all right, okay, anyway, well: A-an he says well I don’t want to make a profit on it.*

Utterance final: *okay, right... and he was a kind of outsider type person, okay, and so.*

Comment Clauses:

The marker used in these clauses signals the transition to quoted speech, like *you know* in: *he says, you know pack and go! That's it! you're out o'here*. And like 'I mean', 'mind you' (to introduce an explanation) *I mean mind you there are hardly an there were like... maybe FIVE subtitles*.

The two kinds of markers reflect Redeker's definition because they both signal the relation between utterances and their immediate context. And the comment clauses can be considered as MD and can be attached either to TMD and classified as code glosses as in *I mean* or to IMD as the commentary (engagement markers) *you know, mind you*. The markers of pragmatic structure can be attached to TMD and classified as additives, e.g. *and*; or concessives e.g. *but*.

4.4.6 A Typology of Metatext (Mauranen 1993a)

In 1993, Mauranen carried out a contrastive study to investigate 'metatext' in English and Finnish economic discourse (see the applied MD studies in the following section 4.6.2). She collapsed Vande Kopple's (1985) typology to four categories to suit her concept of metatext as a text-linguistic aspect that serves the purpose of textual organisation. Although Mauranen used the term metatext her definition and subcategories are typical examples of TMD.

- **Connectives:** conjunctions, adverbials, and propositional phrases that include relationships between propositions in texts: *however, for example, as a result*.
- **Reviews:** clauses that contain an explicit indicator that an earlier stage of the text is being repeated or summarised *so far we have*.

- **Previews**: clauses that contain an explicit indicator that a later stage of the text is being anticipated, *we show below that*.

- **Action Markers**: indicators of discourse acts performed in the text *e.g. the explanation is, to express this argument in notation, to illustrate the size of this distortion*.

In general terms, the categories and subcategories used by Mauranen represent text about text but include only four of the TMD subcategories. While limiting the scope of the study, this may be justified if the sample texts do not show instantiation of the missing subcategories. Mauranen also confines her subcategories to TMD only, whereas IMD is totally ignored. This is perhaps due to her understanding of metatext as serving textual organisation.

Furthermore, it can be argued that in Mauranen's very concise typology two problems may cause confusion to analysts. The first is a result of considering different subcategories as connectives (which is a very broad subcategory that can be divided into subcategories such as concessives, code glosses and rationales) as the above-cited examples show. The second is a consequence of classifying code glosses (Hyland 2004b) as connectives *e.g. for example*, and action markers as in *the explanation is*, and *to illustrate the size of this distortion*.

The reviewed typologies demonstrate the different subcategories and those that are similar to them with corresponding examples. In addition to this the typologies show considerable overlap among the different subcategories. It is not a surprise, therefore, to see many repetitions across the subcategories as well as their realisations.

An important aspect of the application of a theory of MD is to establish typologies for the identification and classification of MD. To this end several typologies have been reviewed in this chapter to assess the feasibility of the subcategories selected for the proposed typology. The results of this are that all the subcategories chosen for the proposed typology exist at least once.

4.5 The Proposed Typology of Metadiscourse

The above review of the different typologies shows they rely on two main categories: textual metadiscourse (TMD) or informational MD, and interpersonal metadiscourse (IMD) or attitudinal MD, as some studies call it. Here the term TMD is preferred to informational MD, because TMD highlights the textual analysis of written texts as used in this study and helps avoid any confusion resulting from the understanding that MD is a non-content aspect of discourse that does not add new information but guides the reader. The term IMD is preferred to attitudinal metadiscourse because it deals with two sides of communication, with the interactional relations between the writer and the reader. This includes not only how the writer refers to himself and attracts the reader's attention, but also how he adds his personal belief and attitude towards the content of proposition as stated in the next section, proposed typology (Steffensen and Cheng 1996).

Because MD has not been applied to English academic texts in the way proposed here, the present study cannot simply be based on existing typologies but requires a typology of MD to suit its own needs. The proposed typology reflects the comparison of the subcategories of TMD and IMD (dealt with in the next section) and the proposed definition of MD (in the previous chapter). Devising the typology depends on reviewing both typologies using MD and typologies other than MD. First, a summary

table of the subcategories for the two major categories of TMD and IMD will be presented. Secondly, a synthesis will be based on this, taking in the main categories and subcategories from the different typologies.

The review revealed that many subcategories have been given different labels, when they actually refer to the same thing. Examples of these were presented in section 4.3 ‘a critical review of MD typologies’ (e.g pre-plans and topicalisers in Crismore’s (1983) typology, attributors and narrators in Vande Kopple’s (1985), and announcements are similar to topicalisers, strategies are similar to previews in the typology of Crismore (1990)). The proposed typology aims to reduce such overlaps. The new subcategories will be checked against the typologies using the term MD and those using other terms before being adopted in the proposed typology.

Table 4- 5: A Proposed Synthesis of Main Categories and Subcategories of MD:

Main categories	Subcategories
Textual Metadiscourse	sequencers, topicalisers, concluders, previews/reviews, additives, concessives, code glosses, rationales, sources
Interpersonal Metadiscourse	hedges, emphatics, evaluatives, commentaries, appeals, self references

The working definition of MD was presented in the previous chapter; *A non-content aspect of discourse which represents the writer’s point of view and attitude to the content of text in order to help the reader to react to the content of the text and the writer’s point of view/perspective/attitude.*

The connection between the definition of MD and the proposed typology in the study can be explained as follows: the writer’s attitude to the text and the reader are signalled through the different subcategories included under IMD. As for the writer’s assessment

(point of view) of the truth of the content and his/her epistemological certainty/uncertainty, they are signalled by emphatics, commentaries and self references. Hedges are used to protect the writer from the reader's attack. In other words, hedging allows the researcher to withhold responsibility to mitigate the reader's reaction to the writer. While evaluatives reflect his/her attitude towards the content, appeals help create an implicit dialogue between the reader and the writer. As far as TMD is concerned, the subcategories of sequencers, topicalisers, concluders and previews/reviews help the reader organise and classify the text. The other subcategories – additives, concessives, code glosses, and rationales – guide the reader to interpret and react to the text. Together with sources, they also help the reader refer to the origin of ideas and concepts. The justification for these subcategories will emerge in the detailed discussion later. The actual typology of subcategories appears in the table below

4.5.1 The Proposed Typology of TMD

Table 4- 6: A Proposed Typology of TMD

<i>Subcategories adapted in this typology</i>	<i>Subcategories used in other typologies and covered by the subcategories in column 1</i>	<i>Examples</i>
Sequencers	Enumerators (M1975), logical-connectives (H1998), frame markers (H&T)	First/firstly/finally/first of all /lastly/listing(a,b,c,)/next/second, secondly/subsequently/third/ thirdly
Topicalisers	Topic shift (N1992), frame markers (H1998),	In regard to/based on/move on/ turn to/with regard to/as for/ bearing in mind/regarding/in relation to/in respect to
Concluders	Transitions (H&T)	In conclusion/in sum/in summary /on the whole/overall/so far/thus far/to conclude/to sum up/to summarize
Previews/reviews	Forecast(N1992), pre-plans(C1983), promissory-asides(N1992), purposes(C1990), goals(C1983), strategies(C1983), announcements(C1990), post-plans(C1983), reminders(V1985), frame/endophoric markers, (H1998, H&T)	Previously/earlier/(in) part X/(in) section X/(in) the section x/(in) this chapter/(in) this part/(in) this section/page X/ figure X/table X/above/below/this work(paper)
Additives	Transitions(H&T)	Accordingly/also/furthermore/ in addition/in the same way/ moreover/therefore/likewise
Concessives	Transitions(H&T)	Although/but/by contrast/ however/in spite of/on the contrary/on the other hand /yet/whereas
Code glosses	Formulators(N1992)	Defined as/e.g./for example/for instance/I mean/i.e./in other words/known as/namely/such as /that means/which means
Rationales	Transitions(H&T)	The reason for this/as a result of/because (of)/ as a consequence /consequently/thereby
Sources	Evidentials(H&T), attributors and narrators(W1981,N1985))	According to X/X suggests/X argues/ X notes/X states/X proposes

(C1983): Crismore 1983, (C1990): Crismore 1990, (H1998): Hyland 1998, (H&T): Hyland and Tse 2004, (M1975): Meyer 1975, (N1992): Nash 1992, (V1985): Vande Kopple 1985, (W1981): Williams 1981

While the right-hand column is headed ‘examples’, they do in fact reflect the sum total checklist of items that were used for assigning sentences to subcategories; i.e. each list is more of a defining characteristic for each of the subcategories than simple examples with the exception of ‘appeals’ in Table 4-7 (which relies on a syntactic category).

Textual Metadiscourse (TMD):

They are the devices which mainly act the role of organising and classifying the text for the reader.

Definitions of TMD Subcategories

1. Sequencers: the main function of such elements is to signal how propositions or pieces of information are ordered. Hyland and Tse (2004) classify it as a ‘sequencing’ which is a subclass of their frame marker category, such as the italicised parts of examples below:

Firstly, a brief background into this research will clarify terminology.....

Finally, grammar teaching can be conducted by means of corrective feedback on learner errors when these arise in the context of performing some communicative task.

2. Topicalisers: these guide the reader to what is happening in different parts of the text such as topic shift which signals a changing of a subject to another, and main topic (a subject which will be discussed) of sentences, paragraphs, or sections, as in these examples:

Drawing from the above-mentioned theories and the comparative analysis of the above two episodes of conversation, we now *turn to* critical review and adaptation...

Regarding the frequency of these words, it is needless to say that the most frequent words in a language are the most valuable for language learners

3. Concluders: they indicate that the writer is about to arrive at a conclusion or a summary of what he has been writing, such as the italicised parts of the examples below:

In conclusion, I perceive the aims of LT as having three alternative approaches:

In summary, the results of the study show that regularity markers are more frequent than power markers

4. Pre/reviews: they alert the reader as to what the writer is going to do in the following paragraph, section (s), chapter, book or article or provide guidance for the reader in connection with what has been/will be going on in the previous or following parts of the text.

The example *below* contains a metalinguistic sequence in which the NS interlocutor is the one who initiates the lengthy metalinguistic discussion.

In the next step, this residual signal is reconstructed by adding the same prediction as was subtracted *earlier* in the encoding process.

5. Additives: they include features which writers use to add information or reinforce it.

The natural content facilitates language acquisition by provision of an abundance of authentic interaction. *In addition*, Cook maintains that such exposure to authentic material can help students to fill the language gap when put in a similar situation as that portrayed. *Moreover*, this particular extract involves L2 users in authentic communicative interaction which is rare in EFL material.

6. Concessives: they mark a change of direction in the running discourse or a modification of preceding ideas.

If the students are at a lower level of L2 learning, it is inefficient and probably detrimental to expect students to use monolingual. *However*, as with use of the L2 in the classroom, as the student's level increases use of a monolingual dictionary should be encouraged, no matter how daunting a prospect it may seem.

..... readers are strongly recommended to apply various techniques and strategies while reading in order to improve their reading comprehension skills. Teachers, *on the other hand*, need to employ a wide variety of teaching strategies to make the reading class interesting and to provide the appropriate help and guidance for

7. Code Glosses: they are used to guide the reader towards the definition, meaning of terms and concepts or to elaborate or illustrate and label the issues in a discourse.

In other words, the social interactional phenomena in which ZPD values place so prominently in the role of early social and cognitive development are presented as a deficient representation of what children ultimately produce in various communicative endeavours

..... L2 learners need to acquire a substantial vocabulary to achieve competencies in practically all L2 skills, *such as* reading, writing, listening, and speaking

8. Rationales: they refer to markers of cause and effect relations, justifications and results.

The primary focus of most comprehensive models of production is on the verbal component of the social actor's behaviour. *Consequently*, they typically accord a central role to abstract propositional representations that serve as input to lower-level, language-based systems where conceptual entities are mapped onto word strings.'

However, there is no absolute agreement to what is frequent and what is not. *The reason for this* is that different people use different criteria for deciding on what is frequent and what is not.

9. Sources: they refer to the features that signal the origin of ideas, concepts or content in general.

According to Krashen (1982), for language acquisition to occur there has to be a considerably large input of authentic language which is presented in context and at a level which is linguistically challenging to the learner.

Biber et al. (1999) *suggest* that "speakers and writers have a variety of

4.5.2 The Proposed Typology of IMD

Table 4- 7: A Proposed Typology of IMD

<i>Subcategories adapted in this typology</i>	<i>Subcategories used in other typologies and covered by the subcategories in column 1</i>	<i>Examples</i>
Hedges	Validity(V1985), modality markers(V1985)	Approximately/as far as/ assumed/broadly/might/perhaps/possibly/ presumably/probably/seems/tend to/uncertain/unclear
Emphatics	Validity(V1985), modality markers(V1985), boosters(H&T)	Actually/already/always/certainly/clearly/ commonly/constantly/definitely/doubtless/ easily/especially/in fact/no doubt/obviously/ of course/really/simplely/surely/truly/ undeniably/undoubtedly
Evaluatives	Attitude markers (C1983;1993), (V1985), (H&T)	Admittedly/appropriately/astonishingly/ correctly/desirable/desirably/disappointingly /dramatically/equally important/more importantly/essentially/fortunately/hopefully / importantly/inappropriately/interestingly/it is safe to say/meaningful/preferably/ remarkably/strikingly/surprisingly/ understandably/expectedly/unfortunately/ unusual /worth/worthy
Commentaries	Engagement markers (H&T), second person category (B1986)	Allow me/you might be missing/you might add/you might consider/you might be call/imagine/you will/(the) reader's/let us/ you/your
Appeals	Rhetorical questions(M&R), engagement markers(H&T)	Questions asked by the writers; e.g. What should I ignore at this time?/How should I respond?/Will they learn from my commentary for future writing?/ Will each student understand?/Do you...?/Does it...?/How does it...?/How do people....?/What circumstances....?/(When) is it....?/Do we....?
Self References	Self- mentions(H1998), (H&T), Illocutionary markers (C1983)	I/we/me/my/our/us/author/researcher

(B1986): Bauvais 1986, (C1983): Crismore 1983, (C1993): Crismore et al. 1993, (H1998): Hyland 1998, (H&T): Hyland and Tse 2004, (M&R): Meyer and Rice 1981, (V1985): Vande Kopple 1985.

Interpersonal Metadiscourse (IMD):

This is used to interact with the reader about the propositional content.

Definitions of IMD Subcategories

1. Hedges: they are used to express uncertainty or doubts towards the content to mitigate the reader's reaction to the writer, as in the examples below.

With the ESL students, changes in grade level represent differences in age and *presumably* cognitive development.

Perhaps one of the more obvious advantages of the visual channel, then, is that it communicates the shapes, movements or relative positions of objects rather than the audio channel or the written word.

2. Emphatics: unlike hedges they are used to express certainty and what writers believe to be true.

This was *clearly* a serious flaw in a lesson that was intended to increase the teacher's understanding of her students!

Ensuring that language learners get frequent opportunities for internalizing prefabricated word groups is not the only task of the language teacher, but *it is surely* one of the most neglected

3. Evaluatives: they are used to signal the writer's attitude and point of view on the propositional content, ideas, concepts, people, things etc.

Again, *not surprisingly*, a large majority of students answered that they understood the Correction response and said they knew how to correct their errors after receiving it...

...the additional reading of what L2 literacies are needed to use the technologies, *an equally important* focus for L2 education because knowing the literacies required to use the new technologies is often critical for learners to meet their social, personal

4. Commentaries: they are used to address the audience directly either by using 2nd person pronouns such as *you, your*, or by using other forms like *imagine, allow me, let us*.

....you might consider what some psychologists have said about the impact of the situation on Jessica

One can imagine a similar kind of process taking place with a victim of bullying.

5. Appeals: the questions asked by the writers to either direct readers to certain people or institutions or which aim to attract the reader's attention and involve him/her in the issues under discussion.

Will they learn from my commentary for future writing?

How do we take into account these teacher and student factors?

6. Self-References: they are pronouns or terms of reference for the writer to indicate the force of the discourse act performed by the writer in the text.

In these interviews *I* asked questions about classroom practices *I* had observed, clarified information gathered.....

While *I* see the inclusion of cultural awareness teaching as a means for personal and intellectual growth, rather than transformation, their general point is accepted both by *myself* and many (more significant) others.

This typology differs from the other typologies in general and Hyland & Tse's (2004) and Crismore et al's (1993) typologies in particular in several ways. The first is that it extends the textual MD as a result of the pilot study and the two comparisons described in the following section. Various textual MD features were found from the pilot study. They were used by the writers in order to help the readers organise, classify, and understand the propositional contents. So the textual MD features were classified according to the writer's different intentions after the test of the feasibility of the proposed typology from the comparisons of other MD typologies.

The second is the subcategory of 'transitions' from Hyland and Tse (2004). According to Dahl (2004), this is a 'diverse subcategory', to which many features can be assigned. This has been divided up into 'concluders', 'additives', 'concessives', and 'rationales' to better describe the different functions of their use. They are used to signal the writer's various intentions to the reader, such as concluders, summarising what he has been writing; additives, adding or reinforcing the previous information; concessive, modifying of preceding ideas; and rationales, justifying the previous information with cause and effect relations. Also 'frame markers' from Hyland and Tse (2004) has been divided into 'sequencers', 'pre/reviews', and 'topicalisers' separately because they are used to help readers with textual elements which account for different functions of operations (signalling how propositions are ordered with sequencers, alerting the reader in connection with what has been going on in the previous/following parts through pre/reviews, and guiding the reader to the main topic of sentences by using topicalisers).

The third difference concerns the engagement markers, a subcategory of IMD from Hyland and Tse (ibid). This has been divided into 'commentaries' and 'appeals' in the present typology because they have different purposes, the former addressing the reader and the latter attracting their involvement.

The final difference is the 'illocutionary markers' subcategory of Crismore et al (1993). This has been renamed as 'self references' because these indicate simply pronouns or terms of reference for the writers rather than indicating a broad function as illocutionary markers mean.

4.5.3 Feasibility of the Typology

The proposed typology was tested in two ways. First the subcategories were checked against the other MD typologies, as follows:

4.5.3.1 Comparison with other MD typologies

Table 4- 8: Subcategories of TMD Used by Other Researchers under the Name of MD

Person/year	Sequencers (frame markers)	Topicalisers/ (frame markers)	Concluders	Previews/ (endopho- ric/frame markers)	Reviews/ (endophoric/ frame markers)	Additives/ (transitions)	Concessives/ (transitions)	Code glosses (code glosses)	Rationales/ (transition)	Sources/ (eviden- tials)
Williams 1981	+	+								+
Crismore 1982	+	+	+	+		+	+			
Crismore 1983		+		+	+					
Crismore 1984		+	+	+	+				+	
Vande Kopple 1985	+	+		+	+			+		+
Beauvais 1986	+								+	+
Crismore et al 1990				+	+			+		+
Crismore 1990				+	+				+	
Nash 1992	+	+	+	+	+					
Crismore et al 1993	+	+		+	+			+		+
Hyland & Tse 2004	+	+		+	+	+	+	+	+	+
Total	7	8	3	9	8	2	2	4	4	6

• Subcategories with parentheses are from Hyland and Tse (2004)

Table 4- 9: Subcategories of IMD Used by Other Researchers under the Name of MD

Person/year	Hedges (hedges)	Emphatics/ (Boosters)	Evaluatives / (Attitude markers)	Commentaries/ (Engagement markers)	Appeals/ (Engagement markers)	Self- references/(Self- mentions)
Williams 1981	+	+				
Crismore 1982	+	+				+
Crismore 1983	+	+	+			
Crismore 1984			+	+		
Vande kopple 1985	+	+	+	+		+
Beauvais 1986	+	+	+	+		+
Crismore et al 1990	+	+	+	+		
Crismore 1990	+	+	+			
Nash 1992	+	+	+	+	+	
Crismore et al 1993	+	+	+	+		+
Hyland & Tse 2004	+	+	+	+	+	+
Total	10	10	9	7	2	5

• Subcategories with parentheses are from Hyland and Tse (2004)

Each of the subcategories exists in at least one other typology. For the TMD the subcategories can be ordered on the basis of how many of the researchers used them: 9 previews, 8 reviews and topicalisers, 7 sequencers, 6 sources, 4 code glosses and rationales, 3 concluders, 2 additives and concessives. This reflects the relative importance of the roles played by each of these categories in the organisation and structure of discourse at local and global levels. For IMD, there are 10 people who use hedges and emphatics, 9 evaluatives, 7 commentaries, 5 self-references and 2 appeals. Again each of the subcategories exists at least once. The subcategories that have been used by several writers therefore represent the essence of IMD. Appeals/ engagement markers seem to be less common in discourse in general, and are only used by Nash (1992) and Hyland & Tse (2004).

In the second comparison, the synthetic categories and subcategories have also been checked against the typologies used by writers who adopted parallel terms other than MD and are summarised below.

4.5.3.2 Comparison with the Parallel Terms to MD

Table 4- 10: Subcategories of TMD used by Other Researchers under the Parallel Terms other than MD

Person/year	Sequencers/ frame markers	Topicalisers/ (frame markers)	Concluders	Previews/ (endophoric/ frame makers)	Reviews/ (endopho- ric/frame markers)	Additives/ (transi- tions)	Concessives/ (transitions)	Code glosses	Rationales/ (transition)	Sources/ (evidentials)
Meyer 1975/Signal	+		+	+						
Lautamatti 1978/Non- topical	+			+						
Keller 1979/Gambits	+	+	+	+		+	+	+	+	
Schiffrin 1980/ Metatalk			+	+				+	+	
Meyer & Rice 1981/ Rhetorical Relationships	+						+		+	
Fraser 1990/Discourse Markers		+								
Redeker 1990/Discourse Markers	+					+	+	+	+	
Mauranen 1993a/ Metatext				+	+		+	+	+	
Total	5	2	3	5	1	2	4	4	5	0

• Subcategories with parentheses are from Hyland and Tse (2004)

Table 4- 11: Subcategories of IMD Used by Other Researchers under the Parallel Terms other than MD

Person/year	Hedges/ (hedges)	Emphatics/ (Boosters)	Evaluatives/ (Attitude markers)	Commentaries/ (Engagement markers)	Appeals/ (engagement markers)	Self- references/(Self- mentions)
Meyer 1975/Signal			+			
Lautamatti 1978/Non- topical	+	+	+	+		+
Keller 1979/Gambits	+	+	+	+		
Schiffirin 1980/ Metatalk			+	+		
Meyer & Rice 1981/ Rhetorical Relationships						
Fraser 1990/Discourse Markers			+	+		+
Redeker 1990/Discourse Markers				+		
Mauranen 1993a/ Metatext						+
Total	2	2	5	5	0	3

• Subcategories with parentheses are from Hyland and Tse (2004)

The Subcategories of TMD used by Other Researchers in the Parallel Terms to MD

Each of the subcategories occurs in at least one other MD typology, showing their feasibility and usability. In terms of number of researchers using a category: sequencers, preview and rationales have 5 each (similar to the same subcategories in comparison one), concessives and code glosses 4 each, concluders 3, topicalisers and additives 2 each, reviews 1. Although few researchers use the subcategories in this part of checking, the feasibility can be attributed to two reasons: firstly, that they cannot be slotted into the subcategories that represent the focus of the used typologies; secondly, the typologies that use terms other than MD are much narrower than MD itself. This means the typologies in the name of MD include subcategories of both textual and interpersonal MD while those in the name of parallel terms other than MD have mainly interpersonal functions.

The Subcategories of IMD used by Other Researchers in the Parallel Terms to MD

Table 4-11 shows that commentaries and evaluatives are used more as terms by researchers with five occurrences for each. Next to these come 3 self-references, hedges and emphatics, each of which occurs only twice. Unlike other subcategories, appeals (Nash 1992) did not appear as such in any of the other typologies.

As commentaries and evaluatives may be considered typical representatives of MD in general and IMD in particular, they are likely to be used by more researchers. The low frequency of hedges, emphatics and self-references may be attributed to the focus and

interest as well as the context and nature of the analysed discourse. Appeals as suggested previously are not commonly used.

Generally speaking, more researchers used the subcategories in comparison one than in comparison two. It may be that some typologies, but not all, especially those in comparison two, are designed for conversational analysis (e.g. typologies of gambits by Meyer and Rice 1981, metatalk by Schiffrin 1980, and discourse markers by Fraser 1990 and Redeker 1990). Nevertheless, the essential subcategories do exist in the surveyed typologies in the two comparisons. Also many of the TMD subcategories like topicalisers, sequencers, previews, code glosses, and IMD subcategories like hedges, emphatics, commentaries, evaluatives and self references appeared in the typologies which were applied to written discourse as well as conversations. This finding leads to the conclusion that the feasibility of the proposed typology can be supported by the two comparisons.

4.6 Applied MD Studies

This section will review applied MD studies. These cover different groups of writing: English native speaker and English non-native speaker students', environmental report and chairman's statement, and professionals' and students' in the studies of metadiscourse. The results show that MD exists in all the investigated groups of writing. However, researchers differ in their interpretations of the differences. While some attribute the differences to language backgrounds, others conclude that they are from genre differences.

4.6.1 Genre Related Studies of MD

The use of metadiscourse in introductory sections of a new genre

The Skulstad (2005) study compares the use of metadiscourse in introductory sections of the emerging (new) genre of 20 environmental reports with the chairman's statement in the established genre of 19 corporate annual reports by using four categories of MD produced by Mauranen (1993a). The texts in both corpora were issued by British companies. The study concludes that writers of the emerging genre use metadiscourse to guide the readers with different frequency compared to established genres.

The differences in the use of categories *action markers* (self-references in the present study) and *previews* (*local* and *global*) are particularly meaningful when comparing established and emerging genres. The use of action markers and global previews in the new environmental genre is more frequent than in their established counterparts and the function of MD in the two corpora is also different. For example the use of previews in the new genre informs and directs the readers as to the aims and global functions of the documents, whereas in established genres this category marks a deviation from what the writer uses in the new genre.

The study concludes that the difference in the use of MD between the two genres is due to the fact that the aims and objectives as well as generic conventions of these reports are different since the audiences of the two genres are different.

Person Markers & Interpersonal Metadiscourse in Academic Writing.

Harwood (2003) investigates the use of 'person markers' (self references in the present study) in experts' and students' writing. The expert corpus consists of a total of forty journal articles from four different disciplines, Business & Management, Economics, Computing Science, and Physics. The student corpus consists of a total of ten Master's dissertations that were awarded a Distinction grade, five from Business & Management, and five from Computing Science.

He found that the most frequent person markers in his corpora are the personal pronouns *I* and *we* which were used to help the writer organise the text and guide the reader through the argument. He argued that the most striking thing about person marker use in the student corpus is its 'lack of resemblance' to that of the experts in the same field. For example while the Business journal corpus contains hundreds of instances of *I*, the Business students' writing is largely free of all person markers, and while computing journal writers prefer *we* over *I*, the computing students prefer *I* to *we*.

He also found differences in the function of person markers between student and journal writing from the qualitative analysis. 'Procedural' *I/we* (that is how *I* and *we* are used in descriptions of the research process) was the most popular function in the Management journal articles whereas there was not a single case of Procedural *I/we* in the Management students' writing. In addition to this, while procedural *I/we* was only the fourth most common person marker function in the computing journal articles, it was the most common in the computing students' texts. In summary he found that there are 'massive quantitative and qualitative differences' between student and journal writing in equivalent fields in the use of person markers while there are broad

commonalities in the journal writing between disciplines as the choice of person marker is concerned.

Thus he concludes that the different use of person markers reflects the writers' different status as students and experts in the same field (which can be genre aspect) rather than the different discipline factor.

Evidentials as Metadiscourse in Professionals' and Students' Academic Writing

Within the concern with MD in academic writing, Barton (1993) studied evidentials as MD in academic writing. Evidentials refer to expressions that reflect attitudes toward knowledge (narrators from Vandepol's typology). Under this term linguists, as Barton states, include validity and attitude markers (e.g. hedges, emphatics, and evaluatives in the proposed typology of the study) which are important subcategories of IMD. Barton argues that in academic texts evidentials reflect argumentative rhetorical strategies. More specifically, she argues that

...evidentials of contrast highlight the rhetorical strategy of problematisation, evidentials of beliefs function in the construction of a persona, evidentials of citation function to place a perspective on the literature and evidentials of all types function to mark claims and counterclaims in the development of specific arguments. (Barton 1993, p747)

The data used in her study consist of two main types: 100 Point of View essays that appeared in the *Chronicle of Higher Education*, written by experienced academic writers from a variety of academic fields; 100 essays written by university students as a requirement for writing proficiency from a variety of majors and programmes.

The results from the study were as follows:

Problematization- Barton's investigation shows that experienced academic writers frequently use evidentials of contrast to problematise their topics (to show that it needs reconsideration). The most frequently used evidentials of contrast are *but*, *however*, and *unfortunately* respectively. In contrast, more than half of the students problematise their topics by identifying a problem without using evidentials of contrast, e.g. *this question remains the topic of much debate* (p756). Barton also noticed that students start with generalisations (surrounding framework), which leads to a kind of problematisation, e.g. *the courtroom in America is something that is truly an experience that all citizens should experience* (p756) and this strategy is exhibited in about half of the student essays. According to Barton, this (linguistic) phenomenon 'makes the problematisation more of a general statement than a contrastive one' (Barton 1993, p756) and can be seen in the extracts from her sample below:

e.g.) a. There has in recent times been **confusion** about the true nature and objectives of sport

b. American newspaper reporters are always trying to **pry** into the private lives of public people.

c. Today, there is a lot of **controversy** on whether or not a woman should use her maiden name.

(Barton 1993, p756)

The words in bold text in the above examples signal a problem (Winter 1977); they anticipate the clause relation or signal the discourse pattern that follows.

Persona: like Crismore and Farnsworth (1990), Barton points out that the vast majority of the professional writers use the first-person singular/plural references. Expressions like *I think that, I believe that*, show a combination of first-person reference with evidentials of belief, and both establish an authoritative academic persona which is maintained by repetition of those expressions.

Generality is also found in the way students use persona as in *we as adults... We* here refers to members of society or culture and not a specific group of people like *we in academe, Let us as scholars*. Barton observes that whereas professionals use credentialed persona, students use an uncredentialed first person (lacking any background information that would give authority).

Citations: In the essays of the professionals (experienced writers), citations are the second consistent evidential expressed in different ways: *say, report, show, urge, some....., the argument as presented byetc.*

In the student essays only 68 use citations uncritically and reflecting a neutral stance. They do not use other evidentials to distinguish their thesis from those cited. Barton also noticed that the students formulate generalisations from the reading passages and cite an unnamed source: *It is argued that, most will agree.*

Arguments: The professional writers use evidentials of deduction like *thus, as a result* and degree-of-reliability, like *in general, undeniable*. Such evidentials call attention to the claim and specify its nature. Counterclaiming within arguments was consistently expressed by evidentials of contrast: (e.g. *but* and *yet*).

Unlike professionals, students' essays used little counter-argumentation: 24 out of 100. Barton also mentions that students used arguments based on agreement rather than

disagreement. To establish their agreement students rely heavily on statements of generalisations. In fact, generalisations dominate all parts of students' essays, e.g. *politicians usually have track records, all kids of today are.*

The common strategy used by students to develop their arguments is the use of examples from the real world, personal experience and self-made examples. This strategy was used by 81 students and was represented by non-specific figures.

Epistemological stance: Barton noticed that experienced writers use an epistemological stance that implicitly defines knowledge in a conventional manner as the result of contrast and competition. Student writers rarely adopt this competitive and contrastive epistemological stance; in most cases they

.....assume an epistemological stance that privileges knowledge defined as a product of shared social agreement and that characterise the knowledge-maker as one who articulates his agreement through generalisations. (Barton 1993, p765)

Several useful conclusions can be drawn from Barton's study. Firstly, MD is an important factor for expressing writers' ideas and attitudes explicitly with the least ambiguity and generality. This was clear in the discourse of the experienced writers who, in contrast with student writers, used more MD and put their arguments and ideas in perspective. In other words, MD plays a role in making discourse explicit, specific and coherent. Secondly, within the same culture, mature writing exhibits more appropriate use of MD than immature writing (student essays). Thirdly, there is some evidence that, in part, qualitative and quantitative use of MD is due to the different genre (expert and student) rather than other factors. Fourthly, it follows from the second and third points that maturation and genre seem to be decisive factors in the appropriate use of MD.

4.6.2 Language Related Studies of MD

The number of studies carried out to investigate MD across languages is limited. Two studies appeared in 1993, both of them comparing English and Finnish texts which will be explored below. Then, MD in English texts produced by Yemeni/Arab writers will be discussed. Together they provide an insight into MD across languages.

MD in American English and Finnish

This study was conducted by Crismore, Steffensen and Markkanen (1993). The purpose was to investigate cultural and gender variation in the use of MD in persuasive texts written by English (USA) and Finnish writers in their native languages. The subjects were forty American and Finnish university students; twenty from each country, subdivided into ten females and ten males. For the analysis of MD the researchers used a modified version of Vande Kopple's (1985) taxonomy, retaining the two main categories of interpersonal and textual MD. They used 'line density' as the unit of measurement. The results can be summarised as follows:

- Both major groups use both textual and interpersonal MD (average one item per line)
- More interpersonal MD is used by both groups.
- The descending order of the subcategories is: text markers, hedges, attitude markers, commentary (addresses), interpretives (code glosses and pre/reviews), certainty markers (epistemic emphatics, e.g. *surely, it is certain that*) and attributers (sources)
- The Finnish students use more MD per line and more hedges than their US counterparts.
- More text markers are used by the US students.

- The two groups are similar in the major categories and in the subcategories of commentary, interpretives, certainty markers and attributors. The main differences are in the subcategories of hedges, attitude markers and text markers.
- As regards gender findings, the research indicates that whereas females write more lines than males the two groups are nearly equal in the quantity of MD per line. They are also equal in the level of density of interpersonal and textual MD. However, the males use more text markers than females. As for the attitude markers, the females use them more than the males. The males use more hedges, certainty markers and attributors while females use more commentary and interpretives. There are gender and cultural differences with respect to the two main categories of interpersonal and textual MD: the Finnish males use interpersonal MD most, while the US males use it the least. Both female groups use similar quantities of interpersonal MD. The situation is different regarding textual MD where the US males use the most and the Finnish males use the least. In fact, for all groups, the US males use the most and the US females use the least. The central differences are that in general terms the Finnish use more attitude markers and the US fewer. It is observable that the differences between the subcategories are much more than between the main categories.

The authors tried to explain why the Finnish students regarded hedging the propositional content as significant and assigned more importance to expressing their attitudes to it compared with the US students, who gave more importance to expressing certainty and attributing ideas to their original sources. They speculated that the Finnish situation came as a result of the historical fact that Finland was long dominated by Russia and Sweden, the two most powerful neighbouring countries. The US students, on the other hand, may consider expressing certainty as a sign of strength, assertiveness, and self confidence and hedging a sign of weakness. In general, the research provided evidence for the universality of MD because subjects from both countries used all categories and subcategories.

Textual MD in Economic Texts Written in English by Native English Speakers and Finnish Speakers

In a small scale study Mauranen (1993a) analysed two pairs of economics texts that she considered suitable for generic comparison. Two of the texts are on forest economics models and two on taxation models. In both pairs, one text is written by a Finnish economist in English and the other by a native speaker of English who is an economist. For the analysis of the texts, Mauranen adapted Vandekopple's (1985) taxonomy. The adapted model is confined to the main category of textual MD and comprises four subcategories only: connectors, reviews, previews, and action markers.

The results were that the two groups exhibited certain different rhetorical preferences. This is evidenced by the examples of the different subcategories that represent rhetorical means for the preferred rhetorical strategies. The Finnish texts employ relatively little MD for the explicit organisation of the text and for orienting the reader. This is shown by the absence of previews (e.g. *this chapter is about...*) and reviews (e.g. *we have looked so far in this chapter at...*) as well as the in-explicit presence of the writer. So the reader has to infer the main thesis and conclusions. According to Mauranen, such features make the rhetorical strategies of the Finnish texts implicit. Moreover, it reflects a preference for a poetic type of discourse in which interpretation depends to an enormous degree on the reader.

Native speakers of English, on the other hand, use a variety of rhetorical devices to make the organisation of their discourse explicit to guide the reader and justify their claims clearly. Mauranen argues that they condition the reader's interpretation and that their awareness of the reader is high as if he or she is present in the mind of Anglo-

American writers as they write. In the Anglo-American discourse, the writer and the reader tour the text together.

In an attempt to explain the differences, Mauranen argues that the two contrasted groups reflect different notions of politeness. The Finns think of the reader as an intelligent person who can understand without too much help but, at the same time, an implicit and obscure writer can also be taken as arrogant and unconcerned, and can be viewed as superior to the reader. Whether these notions are positive or negative is relative, varying from one culture to another.

Cultures also differ in their assumptions about shared knowledge. In homogeneous cultures, like Finland, Mauranen explains, a large amount of shared knowledge can be reasonably assumed. She also speculates that

...in a homogenous context like the Finnish one, it is natural for writing conventions to remain relatively implicit, whereas in culturally heterogeneous contexts, like those in dominant English speaking countries, it becomes imperative to develop writing habits which are more explicit and leave less room for interpretations which are taken for granted. (Mauranen 1993a, p18)

MD in English Texts Produced by Yemeni/Arab Writers

Alkaff (2000) presents a cross-cultural analysis of metadiscourse in letters to the editor as samples of texts produced in English by Yemeni/Arab writers. This analysis sets out to find the extent to which MD features used by Yemeni writers meet or do not meet the expectations of the native speakers of English. In general, the edited texts by English native speakers included more words and more MD. Both (original and edited) types of text also used more textual MD than interpersonal MD. However, the increase of MD as a general entity is not statistically significant. As regards the increase of interpersonal MD in the edited texts, this is not statistically significant, but has an

importance which reveals itself in a qualitative analysis. With respect to the subcategories of interpersonal MD the original texts used fewer hedges and more emphatics than the edited ones. The study suggests that the existing small differences can be attributed to the culture differences.

The study thus provides evidence that the reading, by NES, of texts produced in English by Yemeni/Arab writers results in 'cross-cultural pragmatic dissonance' in terms of using MD.

4.7 Summary

This chapter has dealt with practical aspects and applications of the basic concepts of MD to text analysis. It started with a discussion of the general functions of MD, followed by the functions of the two major categories of textual MD and interpersonal MD. An important aspect of the application of a theory of MD is to establish typologies for the identification and classification of MD. This task has been fulfilled in this chapter. To this purpose several typologies have been reviewed. The review and the tables drawn from the two tests reflected overlap between the different subcategories within and across the typologies. They also showed a lack of harmony between the typologies and definitions. To minimise these things several steps have been taken. Firstly, a table was made to show the overlaps between the subcategories across the reviewed typologies. Secondly, the different subcategories were collapsed and classified as either TMD or IMD. Thirdly, a checking exercise was carried out to test the feasibility of the subcategories selected for the proposed typology. The results of this exercise demonstrated that all the subcategories chosen for the proposed typology exist at least once. The established typology is closely related to the working definition of metadiscourse provided in the previous chapter. Also this chapter reviewed the few

applied MD studies, which show the differences of language or genre reflect the MD usage and which it was hoped would be useful for the finding of the research gap in terms of two main contrastive approaches (language and genre).

Chapter 5: Methodology and Construction of the Corpora

5.1 Introduction

This chapter is devoted to the methodology and the construction of the corpora. It is organised into four main parts. The first part provides the purpose of the study by formulating the research questions. The second part explains the choice of research methodologies. It begins with an explanation of the methods adopted as a result of the pilot study and explains their main characteristics. Then, the procedures for collection and building corpora are described with the selection criteria for each group corpus. The final part explains the limitations of the methods especially the problems of classification of MD features.

The following research questions were addressed arising out of the discussion in Chapters 1-4.

5.2 Research Questions

1. Native versus Non-native writers: How differently and similarly is MD used in non-native English speaker (NNES) student writing and native English speaker (NES) student writing?
 - a. Is it the case that a particularly frequent type of metadiscourse in the NNES corpus is less frequent in the NES, or vice versa?
 - b. How differently are the features of MD used in the corpora?

2. Genres/expertness: How differently and similarly are MD used in the two genres: journal article and Master's assignment?
 - a. Are some types of metadiscourse more frequent in the student corpora (NES + NNES) than in the expert corpus, or vice versa?
 - b. How differently are the features of MD used in the corpora?
3. How do these factors (differences of language and genre) affect the pattern of metadiscourse in academic discourse?

Some researchers argue that the differences of MD usage are from native/non-native factors, others conclude that they are from genre/expertness factors in the applied MD studies. To build upon the applied MD studies, this study investigates how differently these factors (language or genre) influence the pattern of MD usage in academic writing in the same discipline.

For the purpose of this study, firstly, journals were nominated by specialist informants (five lecturers who taught the modules related to the topics in Cross Cultural Communication Studies and Applied Linguistics in the School of Education, Communication and Language Sciences at Newcastle University) in the leading journals in the relevant topics and then 30 papers were selected from recent issues as described below. Secondly, 55 assignments produced by the students, who had been taking modules in the same discipline, were collected, 25 by natives, 30 by non-natives. Thirdly, the following steps were taken sequentially. A concordance program (WordSmith 4.0) searched all eighty-five scripts (25 assignments produced by the postgraduate native English speakers, 30 assignments from the postgraduate non-native English speakers and 30 articles from journals) to find out textual MD (TMD) and interpersonal MD (IMD) features, which had been established as a proposed typology in chapter 4, to be explained in detail in the following sections.

5.3 Research Methodology

In this section, the methodological value of the methods adopted in the present study will be justified. Firstly, the pilot study will be illustrated and, secondly, the characteristics of the methods adopted in the main study will be discussed.

5.3.1 Pilot Study and the Methodologies

The pilot study was carried out to test how the proposed typology worked with the texts produced by the NES and NNES writers. The data from the two different groups (NES and NNES postgraduate students) were compared and contrasted to establish the significant characteristics of the MD usage.

The pilot study analysed four assignments from the NES and NNES postgraduate students in the topic of Cross Cultural Communication Studies for the following types of MD (Table 5-1) and Appendix 5-1 gives details of the texts analysed:

Table 5- 1: A List of MD Subcategories used in the Pilot Study

Textual elements	Interpersonal elements
Sequencers	Hedges
Topicalisers	Emphatics
Concluders	Evaluatives
Previews/reviews	Commentaries
Additives	Appeals
Concessives	Self references
Code Glosses	
Rationales	
Sources	

At this stage, the focus was not exclusively upon single-authored texts. However, the benefit of conducting the pilot is that the decision was made to collect only single-authored work, as discussed below. In addition, a corpus-based approach for the

frequency of MD items and semi-structured interviews (which will be discussed in the next section) were tried out.

As the research questions involve calculating the frequency and percentage of MD occurrences in each corpus, statistical tests were used to test differences between the two groups. Also, the semi-structured interview has been employed to support the above-mentioned methodologies, to gain insights into the data and discover something about the writers' own preferences and thoughts on academic writing in general and usage of MD in particular.

The corpus analysis was chosen for the present study because it can enable us to obtain a detailed description and frequency of the data in order to see how the factors (language and genre/expertness) influence the pattern of MD usage in the two different groups of writing. Also, by employing a semi-structured interview as a part of the qualitative methodology, a possible reason why NNES postgraduate students' writing is different from or similar to that of the NES postgraduate students, can be deducted.

As the pilot study went on, more examples of MD features were found although they had not been found from the initial stage of the framework, and added to the MD subcategory.

5.3.2 The Characteristics of the Methods adopted in the Main Study

The two main methods, corpus-based approach and semi-structured interview, were employed in order to investigate the effect of the difference of the factors in MD usage between the two groups' texts.

5.3.2.1 Characteristics of Corpus-Based Approaches (CBA)

Several advantages come from the use of CBA; for instance, the investigation of word usage, frequency, collocation and concordance (e.g., McEnery and Wilson 2001; Scott 2001; O’Keeffe and Farr, 2003). The fundamental features of corpus-based analysis include the following (Biber et al., 1998; Conrad, 1999):

- It is empirical, analyzing the actual patterns of use in natural texts.
- It utilises a large and principled collection of natural texts as the basis for analysis.
- It makes extensive use of computers for analysis, using both automatic and interactive techniques.
- It depends on both quantitative and qualitative analytical techniques, especially functional interpretations of language use which has been used in the present study for the analysis of the subcategories in TMD and IMD.

5.3.2.1.1 Corpus Studies to Investigate Academic Writing

Corpus studies are both quantitative and qualitative in nature. They involve gathering a representative database of language (which is ‘Lang-3’ from Cook’s (2006/2007) five meanings of the word ‘language’, called ‘a set of sentences’) and analysis of that corpus through computer programs (Biber et al. 1998; Conrad and Biber 2001). Corpus studies are valuable in linguistic research because, while in the past researchers have relied on intuition about the use of a linguistic item in a register, corpus studies rely on empirical evidence. This is supported by Biber et al. (1998, p3) who state that “Analyses cannot rely on intuitions or anecdotal evidence. In many cases, humans tend to notice unusual occurrences more than typical occurrences. Furthermore, we need to analyze a large amount of language collected from many speakers, to make sure that we

are not basing conclusions on a few speakers' idiosyncrasies". On the value of corpus linguistics over other methods of linguistic inquiry, Thomas and Short (1996) point out, "When language is actually used, it is for communicative purposes in social situations. The reality is that by starting with real texts, corpus linguistics has the potential to develop a new kind of linguistics with a much better theoretical foundation than has hitherto been the case" (p51).

The corpus studies method involves qualitative and quantitative methods, in that researchers not only simply look at frequencies of linguistic patterns but also make interpretations of those patterns, which were adopted in the present study.

5.3.2.1.2 Corpus Analysis and Reliability

To enhance the reliability of corpus analysis, ideally two or more raters would have worked on the data independently for both the quantitative and qualitative analysis (Markkanen et al. 1993). The weakness of any quantitative classification of MD is well-known (Hyland 2005b), which is the possibility of omitting and overlapping some MD elements because of their multifunctionality. Hence, two or more researchers may well have identified more multifunctionality in the data. For this reason, an independent reviewer (who is native speaker of English) was asked to recheck the examples of each category of the typology and to rescore the initial stage to obtain an agreement on the establishment of various types of MD elements and the initial analysis. Cases where there was disagreement were taken out from the examples of the typology (as described in the limitations of the method section). In addition to this, all the quantitative data were analysed twice to go some way towards raising the consistency and reliability of the analysis.

5.3.2.2 Semi-Structured Interviews

The method of semi-structured interviews is effective when the researcher collects data from particular individuals by meeting with them. During the semi-structured interviews, I tried to encourage the interviewees to freely express ideas to obtain significant unexpected information as well as the answers to the prepared interview questions.

As Potter (1996, p134) has noted, “interviews have been used extensively in discourse analysis, but they are constructed in a novel manner”. Traditional interviews used to aim at producing colourless interaction. However, in practice, interviews are as complex as any other social event, and responses are dependent on the structure of questions and atmosphere during the conversation. So, in order to conduct successful semi-structured interviews, I asked interviewees questions based on a prepared written list of questions and topics (see appendix 5-2), and tried to make interviewees feel comfortable about talking (Bryman 2001; Labov 1972).

7 NES and 7 NNES postgraduate students, who sent their assignments, were interviewed as a way of both gaining insights into the data and of discovering something about their own preferences and thoughts on academic writing. All of the 14 students were interviewed once. These interviews employed a semi-structured format to obtain writers’ beliefs about MDs and their motivations for using them in order to speculate the possible reasons of differences in the use of MD.

In the present study, I interviewed the actual writers of the text, unlike Hyland (2001) who did not interview the actual writer of the biology text he used. He was assuming that the biologist he was interviewing would be representative of attitudes in the biology community as a whole, and would be able to account for the ‘self mentions’

(self references in the present study) used by the biologist, writer of the text in the corpus (“The interviews were conducted with experienced researcher/writers from the target disciplines using a semi-structured format”, *ibid*, p178).

5.4 Data Collection Methods

Pedagogical applicability and details of each corpus, together with its characteristics and selection criteria, will be described in the following sections to explain the criteria of data collection and usefulness of each group’s data for valid comparisons.

5.4.1 Pedagogical Applicability

As it was envisaged that EAP textbook writers might use the findings of this study to inform their materials, it made sense to study one of the most popular disciplines with international students. With this in mind, statistics from H.E.S.A. (Higher Education Statistics Agency) were obtained of the most popular disciplines with international students at Master’s and doctoral level in the UK. The top 8 disciplines from the most recent set of figures available (2005) are given below.

Table 5- 2: The Top 8 Popular Disciplines with International Students

rank	discipline
1	Business and administrative studies 13.4%
2	Subjects allied to medicine 12.8%
3	Education 8.6%
4	Social studies 8.5%
5	Biological sciences 6.6%
6	Creative arts and design 6.2%
7	Computer science 6.1%
8	Languages 6.0%

Source: Higher Education Statistics Agency (HESA) (2005)

It seemed sensible to focus on one of those disciplines which were popular with international students. Hence it was decided to use texts from the nominated journals and students' assignments in the topics concerning education and social studies which were ranked 3rd and 4th in the above table. I chose the school of Education, Communication and Language Sciences in Newcastle University, which can be representative of many universities in the U.K., to answer the research questions. The considerations which influenced the eventual choice for data collection will be explained below.

A number of pieces of research have been done in the use of MD features in the NNES student and professional texts (Hyland 2002a; Hyland and Tse 2005; Hewings and Hewings 2002), and in NES student writing and professional academic writing (Harwood 2005). However, the study of MDs in the academic writing had not been analysed with two variables, language and genre (NES vs NNES and student vs PRO) at the same time to find out how the variables influence the pattern of MD in academic writing in order to use it in informing pedagogy. From this perspective, choosing the double contrastive analyses to find differences and similarities in the use of MD in the corpora would seem worthwhile.

5.4.2 The Expert Corpus

The selection criteria of the journals for the professional corpus can be summarised as follows:

- a. The topics of the professional texts are relevant to the modules for which the students have to produce assignments for their credits in the topics, C.C.C. and Applied Linguistics in the same department.

- b. Single-authored research articles were chosen because all of the student writings are also single-authored and this ensures that valid comparisons with the student corpus can be made.
- c. For the selection of journals, five lecturers who taught the modules in the topics of C.C.C and Applied Linguistics were asked to nominate the journals which represent the model of topics and writing style that the students concerned are commonly advised to aspire to in their assignments.
- d. Professional corpus size: approximately **226,597 words**

5.4.2.1 Selecting the Journals

A lecturer in C.C.C. (Cross Cultural Communications) and four lecturers in Applied Linguistics were asked to nominate the two most prestigious journals which represent the model of topics and writing style that the students can refer to in their assignments, and the two most popular nominations were chosen in the C.C.C. In the case of Applied Linguistics, the six highest nominations were selected.

As a result of this, *International Journal of Intercultural Relations* (four scripts), *Journal of Language and Social Psychology* (four scripts) were chosen by the lecturer in C.C.C., and *Second Language Research* (three scripts), *Applied Linguistics* (four scripts), *Annual Review of Applied Linguistics* (four scripts), *TESOL Quarterly* (four scripts), *Journal of English for Academic Purposes* (three scripts) and *Journal of Second Language Writing* (four scripts) were nominated by the four informants as prestigious journals in Applied Linguistics (see Appendix 5-3).

5.4.2.2 Selecting Single-Authored Research Articles (RAs) Only

Six years' (2001-2006) issues of the relevant journals were searched to find single authored RAs. Selection involved choosing texts that were single-authored as the student texts are necessarily by one student and choosing the topics is relevant to the modules that the students are taking, namely Discourse Analysis, Second Language Acquisition, Introduction to Applied Linguistics, TESOL Theory and Practice and C.C.C. as detailed in the students' titles in Appendix 5-4 and 5-5. This resulted in a corpus of 30 journal articles containing approximately 226,597 words (see Appendix 5-3).

5.4.2.3. Subdisciplinarity in the expert corpus

A recent study of the academy through an activity of graduate students and professors (e.g. Prior 1998) stresses its heterogeneity, even within what are supposedly single homogeneous disciplines. To a certain extent, no disciplinary subfield can be taken to be representative of the discipline as a whole, since different subdisciplinary communities have the potential to write in different ways. The issue of subdisciplinarity arises when considering the publications which were nominated as the leading journals in C.C.C. and Applied Linguistics. Then, in principle, the professionals writing in each of these subdisciplines could write in different ways – and presumably conform to different patterns of MD use. In the end, four scripts from *International Journal of Intercultural Relations* and four scripts from *Journal of Language and Social Psychology* were selected to be representative of the topic of C.C.C. represented in the expert corpus, and three scripts from *Second Language Research*, four scripts from *Applied Linguistics*, four scripts from *Annual Review of Applied Linguistics*, four scripts from *TESOL Quarterly*, three scripts from *Journal of English for Academic*

Purpose and four scripts from *Journal of Second Language Writing* were selected within Applied Linguistics where possible in order to ensure a spread of topics. This was because the student data were collected from the modules taught in the same school where the postgraduate students are allowed to take a choice of modules which are interchangeable for their MA programme (refer to Appendix 5-4 and 5-5).

5.4.3 The Student Corpus

There are important factors which dictated the choice of the student corpus:

- a. A practical point is that the length of student texts (NES + NNES) should be similar to the professionals' for valid comparisons
- b. the desire that the discipline chosen should be the same as the discipline in the expert corpus, so that a meaningful comparison in terms of similar topic between the two corpora would be possible; assignments produced by postgraduate students who are doing a Master's degree programme in the same department (each assignment is for 20 out of 120 credits for their Master's degree programme and about 5000 words in length from both the NES and the NNES)
- c. the student corpus included the similar amount of data from both postgraduate students that are native speakers of English (NES: 104,586 words) and non-native speakers of English (NNES: 118,858 words) respectively in the same discipline for valid comparisons.
- d. student corpus total size: approximately **223,444 words**

5.4.3.1 Considerations for Selection

There were three important factors which influenced the choice of what the student corpus should contain: (i) the requirement that the discipline chosen should be the same as the discipline in the expert corpus; (ii) the types of students' assignment as a requirement for their 120 credit in their MA courses should be similar to those in the professional corpus (both discursive and empirical studies); and (iii) that the student corpus included a similar amount of data from both postgraduate students that are native speakers of English (NES) and non-native speakers of English (NNES). However, there were a number of difficulties involved in obtaining the required number of assignments from NES postgraduate students, because there were only a few NES postgraduate students in the areas of C.C.C. and Applied Linguistics in 2004 and 2005 whereas there were many NNES students. In the end, it was only possible to obtain 25 assignments from NES postgraduate students which is fewer than the desired number, while 30 NNES assignments were collected.

5.4.3.2 Characteristics of Student corpora

The student corpora consist of a total of fifty-five assignments (223,444 words), twenty-five from native English speakers mainly from the U.K. and another thirty from non-native English speakers from a range of backgrounds, Arabic (5 scripts), Chinese (21 scripts), Japanese (3 scripts), and Korean (1 script). Their minimum English proficiency was 6.5 on the IELTS which is the admission requirement for the course. However their English proficiency varied from 6.5 to 8.5 in the IELTS. This needs to be borne in mind as an individual difference within a group.

Most of the NNES data was produced by students from China, Taiwan, Japan, Korea, and Arabic countries that are collectivist cultures (Hofstede 1991), and have high context (Hall 1983), and reader responsible (Hinds 1987) languages, as mentioned in the literature review chapter.

5.5 Building the Corpora

5.5.1 Expert Corpus

In the case of the professional data, the selected RAs were obtained directly from the electronic versions of the relevant journals (mostly PDF files) and converted to Text Only format. Then all the texts were checked by hand for errors to make sure that no errors occurred in the process of conversion. No abstracts, acknowledgements, footnotes, endnotes, and reference lists were included in the material discussed here to analyse contents only.

5.5.2 Student Corpus

The 55 student assignments were similarly collected and converted to Text Only format and edited as for the journal articles by cutting out abstracts, acknowledgements, footnotes, endnotes, and reference lists.

For ethical reasons it was necessary to have informed consent. Hence some assignments were obtained with the permission of the writers from the department and photocopied and scanned to produce an electronic corpus and the rest of the data was collected via e-mail which enabled the researcher to get an electronic version of the data from the students with their consent (see Appendix 5-11). Because data from the same writer might be biased, people were only allowed to send in one assignment so

there would be a wider range of assignments from different contributors, just like single-authored research articles.

As a result of these, approximately 104,586 words from the NESs' and 118,858 from the NNESs' texts were produced.

The relevant number of words in each group is given in the table below.

5.5.3 The Expert (RACORP) and Student (POSCORP) Corpora

The expert corpus is called here Research Article Corpus (RACORP), and it consists exclusively of expert texts. The postgraduate student assignment corpus is named Postgraduate Student Corpus (POSCORP). Key information about each of the corpora is summarized diagrammatically in the figure below.

Table 5- 3: Details of the RACORP and POSCORP

RACORP	POSCORP
A corpus of 30 research articles (RAs) containing single-authored papers in the discipline	A corpus of 55 assignments written by English native and non-native speakers in the discipline:
Total: 30 articles from eight journals	25 assignments from NES 30 assignments from NNES
CORPUS SIZE 226,597 words	CORPUS SIZE:
	NES 104,586 words
	NNES 118,858 words
	Total 223,444 words

A list of the 30 articles which comprise RACORP, and the 55 assignments which comprise POSTCORP can be found in the Appendices (see Appendix 5-3, 5-4 and 5-5).

In the corpus analysis which follows, a number of abbreviations are used for easy identification of the names of the teaching modules: SLA for Second Language Acquisition, TESOL for TESOL Theory and Practice, DA for Discourse Analysis, IAL for Introduction to Applied Linguistics and C.C.C. for Cross Cultural Communication. Appendices 5-3, 5- 4 and 5-5 show which script corresponds with which research articles and which assignment belongs to whom and from which modules, as they are interchangeable for the students. So, as can be seen from the appendices (Appendix 5-4 and 5-5), the students in C.C.C. took the modules of Applied Linguistics, Discourse Analysis, TESOL Theory and Practice etc. in Applied Linguistics and vice versa. Thus AL (Applied Linguistics) and CCC (Cross Cultural Communications) can be seen as one discipline in the sense since the students in both topics are allowed to take the modules interchangeably for their credits in the same school (refer to the MA programme and module name in the appendices 5-4 and 5-5).

In summary, the corpora consist of expert and student academic writing. The expert corpus consists of a total of thirty journal articles (226,597 words). The student corpus consists of a total of fifty-five assignments (223,444 words), twenty-five from native English speakers (104,586 words) and another thirty from non-native English speakers (118,858 words). All texts in both the expert and student corpora are in English produced by a single-author in the topics of Applied Linguistics and Cross Cultural Communication studies which concern 'language', 'culture' and 'communication'.

5.6. Searching the Corpora

Ideally the texts in the corpora would have been coded manually one by one and summed up as a group, in order to ensure that every example of Metadiscourse (MD) was identified and analysed. However, given the time constraints and the size of the

corpora, this was not feasible. The alternative was to compile a list of MDs which WordSmith Tools could search for. This list was compiled via two principal methods:

- (i) For the establishment of various types of metadiscourse elements, a sample of two texts from each group was initially coded manually to identify potential metadiscourse signals and to classify more delicate subcategories to establish the typology of MD. At this stage, the features which have textual and interpersonal functions were searched by the researcher and an independent reviewer to identify metadiscourse items according to the working definition (see Appendix 5-6).
- (ii) Then, more examples (which had been frequently used in other scripts) were found as the data analysis proceeded, and added to the MD subcategory, although they had not been found from the initial stage of framework. That means further modifications were made to the framework, resulting in the examples of MD elements presented in the following:

Metadiscourse features used in the texts were classified into the instances:

- i) textual MD, such as sequencers (e.g. *first/finally/lastly*),
topicalisers (e.g. *in regard to/based on/ in relation to*),
concluders (e.g. *in conclusion/in sum/to summarise*),
previews/reviews (e.g. *in chapter x/in this section/above/below*),
additives (e.g. *accordingly/furthermore/moreover*),
concessives (e.g. *but/by contrast/however/on the contrary*),

code glosses (e.g. *defined as/for example/i.e.*),

rationales (e.g. *the reason for this/as a result of/consequently*) and

sources (e.g. *according to x/x suggests*),

i) interpersonal MD, such as hedges (e.g. *presumably/approximately/uncertain*),

emphatics (e.g. *actually/certainly/in fact/doubtless/undeniably*),

evaluatives (e.g. *surprisingly/it is safe to say*),

commentaries (e.g. *you might add/let us*),

appeals (e.g. *what should I ignore at this time?/how should I respond?*)

self- references (e.g. *I/we/my/our*).

This reiterative process ended up with a list of potential MD subcategories; the elements for each category are listed in Appendix 5-7:

5.7. Data Analysis Methods

Like other research outlined in the previous chapter, the analysis showed that many MD elements are multifunctional, and can be assigned to more than one grouping in my taxonomy; I nevertheless follow other researchers like Hyland in assigning each element to a single category only. This means that my quantitative analysis of the functions of MDs in my corpora become less problematic (examples of how MD features are quantified into which subcategory can be found in section 5.8.4.3 and the analysis chapter).

5.7.1 Counting the MD Frequency

Once a framework had been established, the entire corpus for each group was searched electronically according to the lists of words in the each subcategory from the proposed typology for the usage of MD seen on pages 98 and 102, to get an idea of the environment of occurrences of each metadiscourse. Then every instance with metadiscourse features was scored as having either a propositional or a metadiscoursal function (see Appendix 5-8, Samples of Counting the MD Frequency).

WordSmith 4.0 (Scott 2004) was used for concordancing. This is a text analysis and concordance program, which provides an alphabetical list of every word in a text and its surrounding line. It is constructed as a one-item-per-search program (Concordance) with the option of searching for an extended sentence which contains the MD features. This function makes it possible to reduce ‘misinterpretation’ in the searched results. However the output of each search was controlled manually by the researcher in order to weed out irrelevant occurrences and ensure that they were functioning as MD from the extended sentences (examples of these can be found in the Appendix 5-8). So I had to manually remove ‘misinterpretation’ in the searched output. In addition, an independent reviewer was asked to recheck the analysis of the initial stage to ensure that the searched cases from each corpus were functioning as MD.

5.7.2 Statistical Analysis

After counting the features of MD in all the texts from each group, it may be asked which test should be used to answer the research questions before carrying out the comparisons, parametric tests or nonparametric tests. The reasons for choosing nonparametric tests in the present study, in particular the Mann-Whitney U test is

outlined in Appendix 5-9. Parametric tests were not employed because the data violated the following assumptions of the parametric tests that the distribution of the population is normal and that the variances in the different groups are equal or homogeneous (also see Salkind 2000). Furthermore, the present sample size (below 30 in each group) is not large enough to employ parametric tests and the size of the samples is unequal (25 texts for NES, 30 for NNES and 30 for PRO and different total words in three groups).

In fact, the majority of the frequencies of each subcategory of MD in both student and professional data were not normally distributed (see normality tests in Appendix 5-10), and a large number of texts did not contain the entire set of the features analysed in the study (see to Bryman and Cramer, 2001). Hence, the Mann-Whitney U Test was selected as a conservative measure of differences between the two groups of data. The Mann-Whitney U Test compares the two groups of data based on their ranks below and above the median (in this study, students' and professionals', and NESs' and NNESs' rates of frequencies for each individual subcategory of MD).

The reason that comparisons of averages (means in the parametric tests) cannot explain the difference of MD use in this study is that averages often obscure the distribution of frequencies in the sample. For example, if in one NES essay, the frequency of sequencers is 20 and in another only 2, then the average frequency of sequencers in the group NES would be 11, which clearly does not reflect the frequency distribution accurately since the present data was not normally distributed. Thus the mean ranks instead of means (averages) were used to identify the statistical difference in the study. The Mann-Whitney U Test ranks the frequencies in a combined array of NES and each group of NNES, one by one. In this way, it can be determined if, for example, the NES

or NNES frequencies of every feature are ranked in a cluster near the top or the bottom of the array, or if they are distributed more or less evenly.

It can be found in Appendix 5-10 that the distributions of each variable (sub-category of MD in each group) were not normally distributed from the Shapiro-Wilk test which is the normality test. Hence the non-parametric tests were performed to determine if the differences are meaningful between the two groups (refer to Appendix 5-9). In the present study, the Mann-Whitney test was employed in order to compare and determine whether the differences are statistically significant between the data produced by professional and student writers and the NES and the NNES students.

If the frequency of MD in a group comprises an extremely high number of MD frequency in one text, that does not represent the mean frequency of the whole group. For this reason, the statistical analysis was performed to identify if the differences are meaningful with statistical evidence rather than distorted results by chance, unlike most studies in MD which did not seem to use statistical analysis (Maurannan 1993a; Hyland 1999a; 2004a,b; Hyland and Tse 2004 and 2005; Crismore 1983; Hewings and Hewings 2002; Harwood 2005).

For the statistical analysis, an extremely short script (NNES 9) was excluded. The reason for leaving this out is that the extremely small number of words (under 1000 words compared to about 3000 words for the others) may skew the results for some subcategories, as this may yield a very high total number of one or a few subcategories of MD when the figures are converted to 5000 words for the direct comparisons of frequencies between groups.

5.8. Methodological Issues

5.8.1 Ethical Considerations

In order to have access to the written pieces of work (e.g. assignments), which were not published, unlike dissertations submitted by postgraduate students (NES and NNEs) to the department at Newcastle University, I initially asked for permission from individual postgraduate students for access to their work. After obtaining written permission from the students, most of the data were collected from them via e-mail with their informed consent (see Appendix 5-11) and the rest were collected from the school office by showing the school officers the contributors' written consent.

Having been granted access to their written work, I then made sure the participants fully understood the purpose of the research and whatever future use it might be put to. I provided the participants with the following information (see Appendix 5-12 for the British Association for Applied Linguistics list of ethics):

- a) My academic background and current study conditions
- b) The purpose of my research (I stated the purpose to be that I will examine certain discourse features of the writing in academic contexts.)
- c) Data collection procedure (after they knew that I needed to scan the paper version of their assignments if I cannot get the electronic version of the assignment, most of them e-mailed me their data.)
- d) The duration of data collection
- e) Data analysis procedure (e.g. concordance program and manual)
- f) The confidentiality of data (I informed them that the data collected will be treated as strictly confidential. The data will be reported only in aggregate

form. Individual participants will not be identified. I have no access to marks or other confidential information.)

- g) The participants' right to withdraw from the study before the data collection procedure

5.8.2 Validity

Bryman (2001, p30) defines 'internal validity' in terms of "the soundness, integrity, and credibility of findings". There is another type of validity, which Bryman (2001, p30) describes, 'external validity' which "is concerned with 'generalisability' (the extent to which the findings can be generalized beyond the specific research context)".

In Discourse Analysis (DA) research, the accomplishment of internal and external validity might be achieved by using corpus analysis which is a quantitative methodology for finding out the frequency of the occurrences.

However, it is well recognized that, unlike quantitative research, validity in qualitative research is very difficult to achieve, because replication as a way of testing the validity is not possible in social research in general and the present research in particular. As Bloor (1997) put it, qualitative research techniques may not validate the findings, but they can be regarded as relevant to the overall issue of validity in as far as they:

..... May yield new data that throws fresh light on the investigation and provide spur for deeper and richer analysis (Bloor 1997, p49)

The validity of the present research has been achieved in two main ways. Firstly, the contributors do not know the research topic, nor are they aware of its objectives, research questions or hypotheses; their writings are therefore natural and unbiased since the data was produced without knowing anything about the study. Secondly it was

validated by using a kind of ‘triangulation’ which refers to the combination of two or more different research strategies in terms of data collection by using corpus analysis, and semi-structured interviews with contributors (Denzin 1989).

5.8.3 Reliability

Reliability is “the question of whether the results of a study are repeatable or replicable” (Bryman 2001, p29). In other words, the question regarding reliability is whether other researchers, looking at the same data, will come to the same conclusion. To this end, an English native speaker who is a PhD student, studying the writing style of Ancient Greeks, was asked to check the elements found for the establishment of the typology and the initial stage of analysis regarding whether they function as MDs.

5.8.4 Limitations of the Method

5.8.4.1 From Corpus-based Study

Despite the care taken in their choice and design, some limitations resulting from the methodology adopted can be witnessed. One major weakness in this study was intrinsic to the corpus-based approach which I have adopted. Widdowson (2000) states these weaknesses:

..... Since what is revealed is contrary to intuition, then [corpus linguistics] cannot represent the reality of first person awareness. We get third person facts of what people do, but not the facts of what people know, nor what they think they do: they come from the perspective of the observer looking on, not the introspective of the insider. [.....] Furthermore, it can only be one aspect of what they do that is captured by such quantitative analysis. For, obviously enough, the computer can only cope with the material products of what people do when they use language. It can only analyse the textual traces of the processes whereby meaning is achieved: it cannot account for the complex interplay of linguistic and contextual factors whereby discourse is enacted. It cannot produce ethnographic descriptions of language use. [.....] [Corpus analysis] is necessarily only a partial account of real language. (pp6-7).

To these we can add other common objections to corpus methodologies outlined by Stubbs (2001) – although Stubbs is a corpus linguist himself, and subsequently goes on to address these objections. Briefly, Stubbs mentions the familiar complaint that corpora are by definition unrepresentative, since they ‘cannot represent a whole language’, and are ‘merely a collection of what it is convenient to collect’ (p.223). Another complaint is that corpora only provide positive data: ‘[a] corpus can reveal only what does occur and not what cannot occur’ (p.224).

Although the list of weaknesses Widdowson and Stubbs describe is formidable, one of Widdowson’s objections can be dismissed immediately, not only with regard to my own study, but also with regard to the vast majority of corpus studies being carried out today. What Widdowson effectively calls for is a methodology which provides a thicker description of why people write the way they do – in this case, why they use the MDs they do, what they know about MDs, and what they think they know about MDs. Widdowson is right to claim that the methodology I have adopted here cannot get behind the text in a way that will provide a sociological account of writers’ beliefs about MDs and their motivations for using them. In order to reduce the weaknesses, I have interviewed 7 postgraduate English non-native and native students respectively about their concepts and motivations of MD use in their academic writing.

I now focus on one of the complaints Stubbs details, that corpora cannot represent the whole language, which is similar to five meanings³ of language by Cook (2006/2007) as mentioned previously (i.e. corpora can only represent the ‘Lang 3’ which is ‘a set of

³ Cook (2006/2007) describes the issue of what language the L2 user knows depends on the meaning of word ‘language’ with five meanings: Lang 1- a representation system known by human beings, Lang 2 - an abstract entity – ‘the English language’, Lang 3 - a set of sentences – everything that has or could be said, Lang 4 - the possession of a community, and Lang 5 - knowledge in the mind of an individual.

sentences' from Cook's classification— but they cannot represent the all five meanings of language). As Stubbs himself goes on to argue, this objection to corpus linguistics misses the point, because any corpus which was truly representative of all of these disciplines would be enormous:

The population to be sampled ... is huge, and even corpora which seem very large by today's standards (thousand of millions of words) are hardly a drop in the ocean, when compared with the size of what is being sampled. (p.223)

It would clearly be beyond the scope of a doctoral project to construct a larger corpus. Also, there were few NES postgraduate students whereas there were many NNES postgraduate students in the school, so I was not able to collect enough data from NES postgraduate students, although anything over five million words is considered to be large enough for written corpora (Hunston 2002; Biber et al. 1998 and 1999; O'Keeffe et al. 2007).

However, two points must be addressed with respect to the issue of corpus size and representativeness. Firstly, as Leech (1991) argues, size is not all-important, mainly when texts are not easily available. In such cases, practical constraints override theoretical considerations. This was clearly the case in the research reported here. Secondly, the usefulness of small corpora is to serve as a sample of a specific kind of text for a particular kind of investigation. A small corpus such as the one used here cannot be taken as a representative whole sample of the MD usage of postgraduate students in the U.K. Its purpose is rather to help indicate possible characteristics of a part of the MD usage of the postgraduate student writers.

5.8.4.2 Different Genre

According to Swales (1990), genres are characterised by their 'communicative purposes' as well as by their patterns of 'structure, style, content and intended

audience' (p58). By this definition, the two types of writing I analyse in this thesis, journal articles and Master's assignments, are two different genres. So the students' assignments and journal articles are not identical in terms of style and readership. Thus, this needs to be borne in mind in making comparisons. However, in order to reduce the weakness I tried to get similar word lengths and topics for the professionals' articles (which is for similar patterns of content) from the relevant journals based on the nominations as the students' texts concerning the topics in C.C.C. and Applied Linguistics with about 5000 words, to persuade their reader (module leader) with their intention although they are of a different genre from research articles.

5.8.4.3 Problematic Classification

An important point in the analysis is the identification of the relevant metadiscoursal elements. This has implications for how features are counted. Williams (1981 p195) states that '....MD is discourse about discourse, words and phrases and clauses.... even sentences.....' and Mauranen (1993a p8) also argues that 'metatext can be realised through all kinds of linguistic units, ranging from affixes (like some Finnish clitics...) to whole clauses'. As an example, she provides the following sentence, which she maintains is metatextual in its entirety:

Let us now explore the implications of the above theorem.

Likewise, the present corpus contains such a sentence operating solely at a metadiscourse level, as in the following example too:

For the moment now, let us consider the previous exchange of utterances as a learning catalyst for the word chunk (NES 17).

As can be seen from the extract in NES 17, it was difficult to establish a potential MD element with the whole sentence without doing the analysis from each script manually.

Here are other examples by the different writers:

.....some leading specialists in L2 teaching and applied linguistics have maintained that the superiority of, for example, task-based instruction (PRO 10)

The work done in first language research has provided new perspectives for research dealing... (NNES 12)

While the '*some leading specialists.....*' in PRO 10 and '*The work done in first language research has provided*' in NNES 12 contain elements of the 'source' (Textual metadiscourse), it is difficult to identify them as metadiscoursal elements by using concordancing, unless the whole of each sentence has been added to the subcategory as an MD element.

Thus, if we want to quantify frequencies of various types of metadiscourse elements, each of the elements needs to be identified separately. For this reason, I decided to exclude potential MD elements consisting of whole sentences and clauses unless each script was analysed individually.

I also excluded a few instances with parentheses functioning as 'code glosses': (e.g. *also called genre based*) in sentence no 7 in the extract from the PRO 10 (see Appendix 5-6) that did not contain any elements that had been established in the proposed typology, as in:

...literacy based, community based, competency based, or standards based (and this is not a complete list by any measure) (PRO 10)

There are doubtless more MD elements that could have been classified as metadiscourse within the present corpora. But it is not feasible to check every instance of MD in each script for the size of the present corpora. So, for the purposes of counting of MD features in the present study, the analysis is restricted to metadiscourse elements which firstly had been found in the sample texts from each group and

secondly additional items which have been added to the subcategories as the analysis went on (see Appendix 5-7); i.e. the basic principle was to judge for the specific examples of MD listed in the typology on pages 98 and 102.

Some extracts from the corpora were difficult to fit into the MD functional taxonomy outlined in the previous chapter. These include multifunctional elements, which are the nature of MD, such as ‘e.g.’, ‘see’ and ‘for example’ followed by references which help the reader to elaborate the issue of the context, and could be classified as ‘code glosses’ or may alternatively, refer to the features that signal the origin of ideas as ‘sources’.

The extracts below convey both functions at once:

.....all of which have the goal of developing students' language proficiency and skills (for detailed discussion, *see, e.g., Ellis, 2003; Fotos, 2001, 2002; Snow, 2005*) (PRO 10 in appendix 5-6).

However, the particular issue of the effect of the learner's age on the acquisition of a second language (L2) is one that remains controversial. (*For example, Singleton, 1989, Flege, 1987 and Patkowski, 1990*) (NNES 12 in appendix 5-6).

There are other examples of multifunctional elements which can be assigned either to ‘code glosses’ (to elaborate the issues in a discourse) or to ‘pre/reviews’ (to guide the reader in connection with what has been going on other parts of the text):

An example is given in appendix E..... Students could be encouraged to keep records of subordinate relationships using diagrams like those shown in appendix E (NES 18)

However, for the purpose of the quantification of the frequency, I followed other researchers like Hyland and assigned the elements to the category that appears to describe their primary function, taken to be ‘code glosses’ in the present study. In fact during the analysis it rapidly became apparent that many MDs are multifunctional, and could be assigned to two, perhaps even three, of my categories.

Let us look at the case of *commonly*. In the following the italicised MD element seems to be fulfilling two functions:

Commonly accepted perspectives on language teaching and learning recognize that, in meaningful communication, people employ incremental language skills not in isolation but in tandem (PRO 10).

On the one hand a sentence with ‘*commonly*...’ can be classified as a hedge, since it can stress to the reader that the proposition is the writer’s opinion only, and that readers are free to disagree if they wish to. On the other hand it can also be seen as an emphatic; in the following case, emphasis is achieved by diverting attention away from the fact that it is the writer’s personal interpretation of the issue.

As some of the above examples imply, ellipsis is more *commonly* found in spoken than in written discourse. However, the feature is also found in written material.... (NES 10)

So those examples with *commonly* are classified as ‘emphatics’ in the present study.

The cases with *commonly* which could potentially be classified as ‘hedges’ were left out of the items search after the discussion with the independent reviewer since most of the cases with *commonly* were functioning as emphatics from the initial stage of the analysis. This is for assigning the elements to the primary function category although the precise classification of items is subjective.

Each putative MD had to be studied in context (Hyland 2005b): as Markkanen et al (1993) show, this contextual approach is the only way to ensure that all apparent instances of MD are in fact true MD. For instance, the following extract from my expert corpus contains many pronouns:

E5: *I* think *I* am more sympathetic. *I* would try to get sympathetic reviewers. *I* would get as many revisions as possible. *I* mean *I* still would uphold that *we* have to have high standards for the journal, but *I* would give as much support as *I* could to the NNS to publish in the journal. (PRO 17)

These pronouns cannot be classified as self-references (Interpersonal MD) because the writer is not using them to indicate the force of the discourse act but quoting interview data, where the informants are using self-references to interact with the interviewer:

As the data analysis proceeded with the elements of each category, there are overlaps of function for the elements for which the concordancer is searching, either propositional or metadiscoursal. The extract below illustrates the point, in that the element

furthermore fulfils both propositional (ideational) and metadiscoursal functions:

The single most pervasive outcome of this study is that English language policies and practices have been implemented, often at significant cost to other aspects of the curriculum, without a clearly articulated rationale and without a detailed consideration of the costs and benefits of such policies and practices on the countries in question. *Furthermore*, there is a widely articulated belief that, in public schools at least, these policies and practices are failing (PRO 28)

The conjunction *furthermore* functions to link a clause to extend the propositional content, but it also functions to signal the writer's intention to add information to the previous sentence which conveys metadiscoursal function.

So, distinguishing a propositional from a metadiscoursal role is problematic without consideration of the function of each element in the context. Therefore, it is necessary for a researcher to check and weed out irrelevant occurrences from each context to ensure that the searched items are functioning as MDs for the valid counting of MD frequency although the decision can be subjective.

5.9 Summary

The methodology used in this research was considered and procedures of data collection were introduced. Building the corpora with the characteristics of each corpus followed. Also the data analysis methods and methodological issues have been

discussed with acknowledgement of the limitations of the method. This study is qualitative and quantitative in nature since the corpus-based study and semi-structured interview were selected as a result of conducting the pilot study to answer the research questions.

Chapter 6: Results

6.1 Introduction

This chapter presents the results from the analysis of the text corpora used in the present work and from the semi-structured interviews with students. Using the MD typology established in Chapter 4, the TMD (textual metadiscourse) and IMD (interpersonal metadiscourse) characteristics were established for the NES (native-English-speaker) texts, the NNES (non-native-English-speaker) texts and the PRO (professional) texts.

After carrying out the main analyses, three other steps have been applied. Firstly, the NES and NNES texts have been compared to answer the research question 1 (namely, ‘what are the types and frequencies of MD used/ how differently are they used in the NNES and NES student writings?’). Secondly, the student texts (NES + NNES) have been compared with the journal article texts (PRO) to answer research question 2 (namely, ‘what are the types and frequencies of MD used/ how differently are they used in the journal articles and the MA assignments?’) Finally the results of the semi-structured interviews are reported to elucidate the reasons for the differences and similarities thrown up by the text analyses. Research question 3 will be answered in the discussion as it depends on results from research questions 1 & 2.

The presentation of results is in five parts. Section 6.2 deals with the quantity and density of MD in the whole corpus to show how MD frequency has been calculated for each corpus. Section 6.3 presents a general overview of MD frequency to answer the sub-research question 1 (i.e., ‘are there any differences in the frequency of MD in the

different group corpora?') in research questions 1 and 2. In Sections 6.4 and 6.5 the results for each MD subcategory are discussed with reference to the two variables 'language' and 'genre'. In this part, the similarities and differences of MD usage will answer the sub-research question 2 (namely, 'how differently are the features of MD used in the corpora?') in research questions 1 and 2. Section 6.6 is concerned with the students' beliefs and motivations regarding MD usage as shown in the interview to find out possible reasons for group preferences. In the final section 6.7, a summary of the findings concludes the chapter.

6.2 The Quantity and Density of MD in the Whole Corpus

As elaborated in the previous chapter, MD is a non-content aspect of discourse which represents the writer's point of view and attitude to the content of text in order to help the reader react to the content of the text and the writer's point of view/
/perspective/attitude.

The results discussed in the following sections deal first with the raw frequency of TMD and IMD, then with the mean frequency of total textual and interpersonal MD. Then the mean frequency of each MD subcategory per five thousand words will be illustrated for the direct comparison of frequencies between groups.

The concordance program WordSmith 4.0 was used to identify frequencies of analysis. All eighty-five texts were searched for examples of textual MD (TMD) and interpersonal MD (IMD) features (based on the proposed typology). The raw figures for the data are presented in the tables in the appendices 6-5 and 6-6 and should be referred to for detailed support for each section.

First, the total number of words in each script was counted. Then for quantifying the frequency of MD, all the features of each subcategory in the proposed typology (see appendix 5-7) were searched to identify the occurrence of MD. All the instances of searched cases were examined by the researcher according to the working definition and those that did not fit were excluded from the counting of MDs. Examples of this can be found in Appendix 5-8. Then the occurrences of each subcategory of TMD and IMD were calculated for each text. Secondly, the data were normalised by expressing them in terms of frequency per 5000 words (refer to Alkaff, 2000 who normalised the different size of his data per 1000 words for a direct comparison of MD frequency). This involved multiplying the total occurrence of MD by five thousand and then dividing it by the total number of words for each text. The reason for the counting of MD per five thousand words is that there are different numbers of words in each script and each corpus; the full details are given in table 6-1 and Appendices 5-3, 5-4, 5-5. Converting the sheer number of MD occurrences into the normalised five thousand words per text allows direct comparisons of the frequency, summed up as total TMD and total IMD for each group corpus. The following formula was used to calculate the MD frequency per 5000 words for each text in each group corpus:

$$\frac{\text{MD features per text} \times 5000}{\text{Words per text}}$$

Thirdly, for each corpus the percentages and mean frequencies of each TMD and IMD subcategory were calculated per five thousand words. After that, Mann-Whitney U tests were employed to find out the statistical difference in the frequency of MD use between groups.

The reasons for choosing nonparametric tests, in particular the Mann-Whitney U test, are outlined in the methodology chapter. Parametric tests were not employed here because the data violated the assumptions of parametric tests, as mentioned in Chapter 5. The Mann-Whitney U test was chosen from the other nonparametric tests to deal with raw frequencies of each subcategory in the TMD and IMD per 5000 words for each text.

The analysis begins by taking a general look at frequency of MDs (i.e. both TMD and IMD) in the three different group texts. Next, the group preference of MD choice in each student corpus (i.e. NES and NNES) and each genre (i.e. journal article and Master's assignment) will be described. Then, the results of text analyses of each individual subcategory will be discussed with a focus on the group preference of MD usage using the functional taxonomy described in Chapter 4 (see section 4.5) to find out how the factors affect patterns of MD use in academic writing.

6.3 General Overview:

6.3.1 MD Use in the Three Group Texts

Before discussing group differences, a general presentation of the two major categories of TMD and IMD in the whole corpus is needed, including the raw frequency of MD subcategories, mean frequency of TMD and IMD with percentages as well as the similarity or difference in the use of MD in three different group texts.

The total number of words in the NES texts is 104,586 with an average of 4,183 words per script and the total in the NNES texts is 118,858 words with an average of 3,961 words per script. The professional texts (PRO) contain 226,597 words with an average

of 7,553 words per text. In terms of the average number of words, the professional texts are the highest, NNES the lowest with NES corpus coming in between:

Table 6- 1: The Size of the Corpora

Groups	Total Words	Average Words
NES	104,586	4,183
NNES	118,858	3,961
PRO	226,597	7,553

For the counting of MD frequency, all the features of the functional typology were searched with WordSmith tools. The results were examined to find out if the target items were functioning as MDs, as described in Appendix 5-8. After classifying each instance, the MD frequency was obtained from each corpus as follows:

Table 6- 2: Raw Frequency of TMD in the Three Groups

TMD subcategories	NES (104,586 words)	NNES (118,858 words)	PRO (226,597 words)
Sequencers	244	514	501
Topicalisers	26	46	26
Concluders	26	20	106
Pre/reviews	54	77	190
Additives	365	452	756
Concessives	546	510	1295
Code Glosses	267	408	853
Rationales	125	169	349
Sources	155	125	319
TOTAL TMD	1808	2321	4395

(Figures in brackets are total number of words in each corpus: other figures refer to MDs as wholes – i.e. they may consist of more than one word)

Table 6- 3: Raw Frequency of IMD in the Three Groups

IMD subcategories	NES (104,586 words)	NNES (118,858 words)	PRO (226,597 words)
Hedges	195	127	476
Emphatics	267	315	540
Evaluatives	81	73	127
Commentaries	33	45	50
Appeals	11	9	21
Self references	474	345	946
TOTAL IMD	1061	914	2160

(Figures in brackets are total number of words in each corpus: other figures refer to MDs as wholes – i.e. they may consist of more than one word)

As the tables show, the total number of TMD is 1,808 in the NES texts, 2,321 in the NNES and 4,395 in the PRO texts; the total number of IMD in the NES is 1,061, 914 in the NNES and 2,160 in the PRO texts. However a direct comparison cannot be made between these figures because the total numbers of words in each corpus are different, as elaborated above (104,586 in the NES, 118,858 in the NNES and 226,597 in the PRO texts).

Table 6- 4: Mean Frequency for TMD per 5000 words in the Three Groups

TMD Subcategories	NES	NNES	PRO
Sequencers	11.67	21.62	10.41
Topicalisers	1.24	1.94	0.64
Concluders	1.24	0.84	2.05
Pre/reviews	2.58	3.24	3.95
Additives	17.45	19.01	16.88
Concessives	26.10	21.45	28.57
Code Glosses	12.76	17.16	19.76
Rationales	5.98	7.11	7.50
Sources	7.41	5.26	7.04
TOTAL TMD	86.43	97.63	96.8

Accordingly, table 6-4 presents the normalised data for the corpora according to the formula given above, which will be used as a basis for most of the following account.

The mean frequency of total TMD per 5000 words is 86.43 in the NES, 97.63 in the NNES and 96.80 in the PRO texts. So, the NNES writers employed TMD the most, followed by the PRO writers; the NES group used TMD the least.

The texts from the three groups have the top four TMDs in common (sequencers, additives, concessives, code glosses), as seen in Table 6-4. Their use of TMD is therefore similar, although the frequency varies slightly in the three group corpora.

Table 6- 5: Mean Frequency for IMD per 5000 words in the Three Groups

IMD subcategories	NES	NNES	PRO
Hedges	9.32	5.34	10.50
Emphatics	12.76	13.25	11.92
Evaluatives	3.87	3.07	2.80
Commentaries	1.58	1.89	1.10
Appeals	0.53	0.38	0.46
Self references	22.66	14.51	20.87
TOTAL IMD	50.72	38.45	47.66

Table 6-5 shows that the mean frequency of total IMD is 50.72 in the NES, 38.45 in the NNES and 47.66 in the PRO texts. The NES writers employ IMD the most, followed by the PRO writers; the group that used IMD the least were the NNES. These are summarised by the frequency of main categories of MD in the following (see Tables 6-6). They show similar choice of IMD subcategory with three most frequent subcategories (hedges, emphatics, self-references). However, some of the IMD subcategories (hedges, self references) which are used less frequently by the NNES compared to their counterparts show the differences.

Table 6- 6: Mean Frequency of Total MD, TMD and IMD per 5000 words in the Three Groups

Sample Group	Total TMD	Total IMD	Total MD
NES	86.4	50.7	137.1
NNES	97.6	38.4	136
PRO	96.8	47.6	144.46

The table shows that the mean frequency of total TMD per 5000 words in the NES texts is 86.4 times and 50.7 times for the IMD while the NNES texts contain 97.6 for the total TMD and 38.4 for the total IMD, and the PRO texts have 96.8 for TMD and 47.6 for IMD.

This total number of TMD is therefore more than the total IMD for all three groups. In line with Hyland and Tse (2004), TMD is more frequent than IMD in the academic

writing. The NNES texts have the highest proportions of TMD whereas the NES the lowest.

A general overview of the corpora shows broad similarities and differences. The MD use in each student corpus (NES and NNES) and each genre (journal articles and Master's assignments) will be compared in more detail in the following section to answer the first sub-questions in the research questions 1 and 2 respectively (i.e. 1. 'Is it the case that a particularly frequent type of MD in the NNES corpus is less frequent in the NES, or vice versa?' and 2 'Are some types of MD more frequent in the student corpora (NES +NNES) than in the PRO corpus, or vice versa?').

6.3.2 MD Use in the NES and NNES texts

Before discussing the frequency of each subcategory, a general presentation of the two major TMD and IMD categories is in order. This will include the mean frequencies and percentages of the two major categories of MD per 5000 words as well as the significance of the difference between the two groups. Then the mean frequencies with percentages of individual subcategories of TMD and IMD from the two corpora, NES and the NNES, will be compared. In addition the mean ranks from the statistical analyses will be presented to test their significance.

The total number of words in the NES texts is 104,586 with an average of 4,183 words per script while the NNES texts contain 118,858 words with an average of 3,961 words per script (see Table 6-7). The average number of words in the NES texts is 222 more than the NNES ones.

Table 6- 7: The Size of Data in the NES and NNES

Groups	Total Words	Average Words
NES	104,586	4,183
NNES	118,858	3,961

6.3.2.1 The Use of TMD in the NES and NNES texts

Table 6- 8: Mean Frequencies of total TMD and TMD Percentages out of MDs per 5000 words in the NES and NNES texts

Sample Group	TMD	Total MDs	Percentage of TMDs in MDs
NES	86.4	137.1	63%
NNES	97.6	136	72%

The data was filtered for the counting of MDs as described above (see Appendix 5-8).

The table shows that the mean frequency of total TMD in the NES texts is 86.4 times per 5000 words, with a density of 63% out of total MD occurrence. In the NNES texts, the frequency of TMD is 97.6, with a density of 72%. The difference between the NES and NNES texts in TMD frequency is 11.2 occurrences per 5000 words. However, the comparisons of averages (means) cannot explain the statistical difference of MD use since the majority of the frequencies of each subcategory of MD in the corpora were not normally distributed, as already mentioned in the methodology chapter. So the mean ranks, which list the average of the frequency ranks for each group, are as follows:

Table 6- 9: Mean Ranks for TMD per 5000 words in the NES and NNES

Sample Group	N	Mean Rank
TMD NES	25	21.20
NNES	29	32.93
Total	54	

As can be seen from Table 6-9, the NNES writers employed more textual MDs (11.73) than their NES counterparts. The difference between the two groups in TMD use is

highly significant (Mann-Whitney U test, $z = -2.732$, $p < 0.006$). The NNEs students therefore employ significantly more TMD than their NES counterparts.

The table below provides details of the TMD use in each sub-category for each corpus.

Table 6- 10: Mean Frequencies of TMD subcategories per 5000 words in the NES and NNEs texts

Sample group	Sequencers	Topicalisers	Concluders	Pre/reviews	Additives	Concessives	Code Glosses	Rationales	Sources	Total TMD
NES	11.7	1.2	1.2	2.6	17.4	26.1	12.8	6.0	7.4	86.4
NNEs	21.6	1.9	0.8	3.2	19.0	21.5	17.2	7.1	5.3	97.6

Table 6- 11: Percentages of TMD Subcategories per 5000 words out of total TMDs in the NES and NNEs texts

Rank Order	NES	NNEs
1	Concessives (30)	Sequencers (22)
2	Additives (20)	Concessives (22)
3	Code Glosses (14)	Additives (20)
4	Sequencers (13)	Code Glosses (17)
5	Sources (9)	Rationales (7)
6	Rationales (8)	Sources (6)
7	Pre/reviews (3)	Pre/reviews (3)
8	Topicalisers (1)	Topicalisers (2)
9	Concluders (1)	Concluders (1)
Total TMD	100%	100%

(Percentages out of Total TMD are given in brackets)

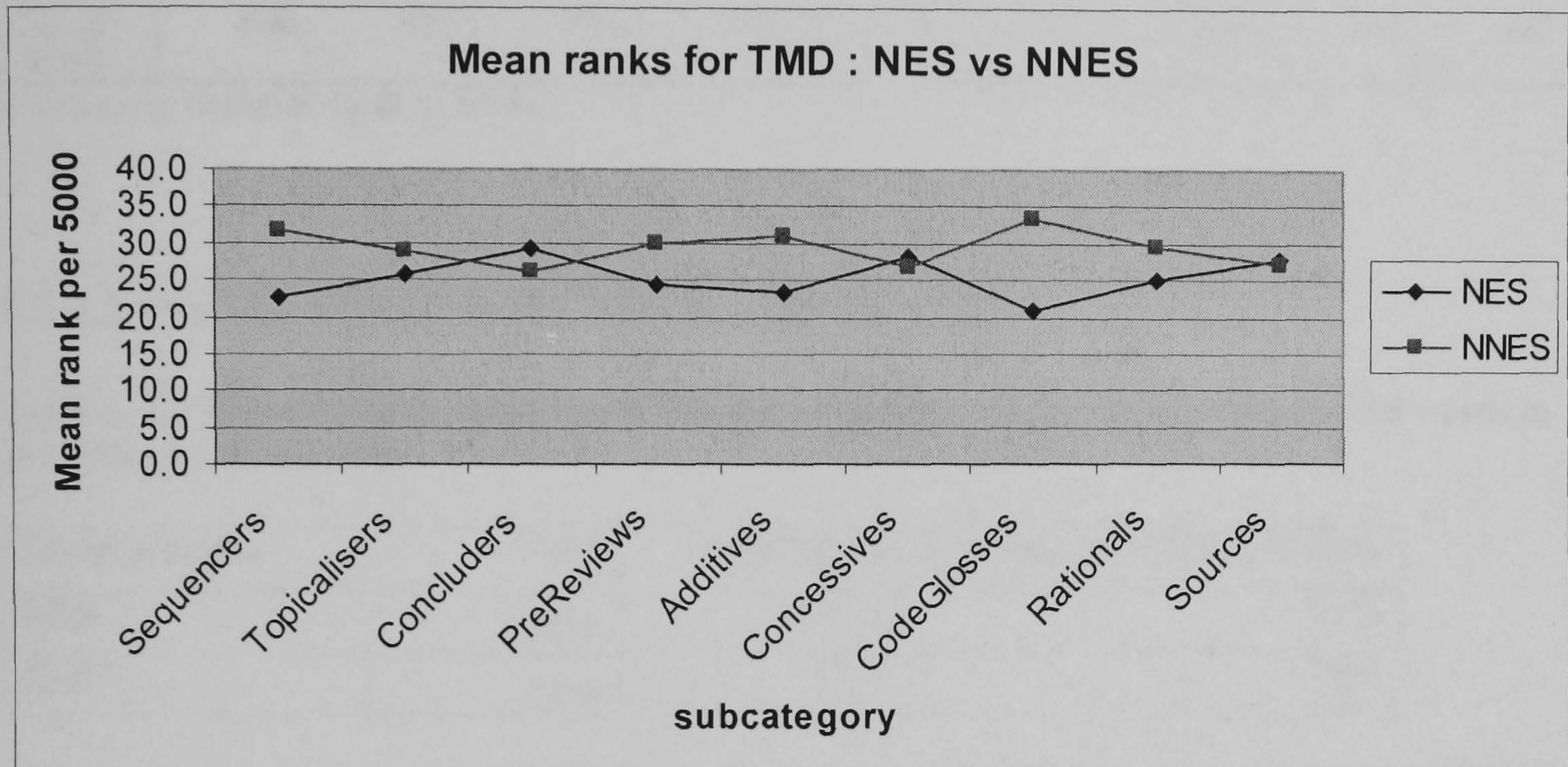
Tables (6-10 and 6-11) show that, in the textual MD, both groups used concessives, additives, sequencers, and code glosses more frequently than the other subcategories.

One of the most frequent sub-categories in both groups is concessive (e.g., *however, whereas, on the other hand* etc, 30% in the NES and 22% in the NNEs). Additives (*therefore, moreover, likewise*) are also far more common in both texts. The use of additives in both texts shows lack of differences with the same percentage of usage (20% in both groups). However there are group differences in the use of TMD subcategories with high usage of sequencers (*first, second, finally, lastly*) and code glosses (*i.e., in other words, for example*) in the NNEs texts. The group preferences in the use of TMD are as follows:

Table 6- 12: Mean Ranks for the Frequency of TMD Subcategories per 5000 words in the NES and NNES texts

Sample group	Sequencers	Topicalisers	Concluders	Pre/reviews	Additives	Concessives	Code Glosses	Rationales	Sources
NES	22.7	25.9	29.1	24.5	23.4	28.3	20.7	25.2	27.9
NNES	31.6	28.8	26.1	30.0	31.0	26.7	33.3	29.4	27.0

Figure 6- 1: Mean Ranks for the Frequency of TMD subcategories per 5000 words in NES and NNES texts



As can be seen from Table 6-12 and Figure 6-1, the NNES used most of the TMD subcategories more often than the NES. However, ‘concluders’ (*in summary, to summarise, to conclude*) and ‘concessives’ (*however, on the other hand, whereas*) were relatively more employed by the NES writers. Although there are apparent differences in the frequency of most of the TMD subcategories between the two groups, there was no statistically significant difference in their use apart from ‘sequencers’ and ‘code glosses’ (sequencers: $z = -2.073$, $p < 0.038$; code glosses: $z = -2.940$, $p < 0.003$) which have been more frequently used by the NNES writers (details are summarised in the table below).

Table 6- 13: MWU Tests for TMD Sub-category in the NES and NNES

	Sequencer s	Topicali- sers	Conclu- ders	Pre/ reviews	Additives	Conce- ssives	Code Glosses	Rationales	Sources
Mann-Whitney U	243.00	324.00	322.50	289.00	261.00	342.00	193.00	305.00	350.50
Z	-2.073	-.688	-.712	-1.292	-1.761	-.356	-2.940	-.998	-.208
Asymp. Sig. (2-tailed)	.038	.491	.476	.196	.078	.722	.003	.318	.835

a Grouping Variable: NES vs NNES

6.3.2.2 The Use of IMD in the NES and NNES texts

Table 6- 14: Mean Frequencies of total IMD and IMD Percentages out of MDs per 5000 words in the NES and NNES texts

Sample Group	IMD	Total MDs	Percentage of IMDs in MDs
NES	50.7	137.1	37%
NNES	38.4	136	28%

Table 6-14 shows that the average frequency of total IMD per 5000 words in the NES texts is 50.7, in the NNES texts, 38.4. In terms of density, the NES texts have 37% out of total MD use whereas their counterparts, NNES texts, have 28%. In general, the NES texts included more occurrences of IMD than the NNES ones. However the difference between the two groups in the use of IMD is not significant (MWU test $z = -1.466$, $p < 0.143$).

Table 6- 15: Mean Frequencies of IMD Subcategories per 5000 words in the NES and NNES texts

Sample group	Hedges	Emphatics	Evaluatives	Commen- taries	Appeals	Self references	Total IMD
NES	9.3	12.6	3.8	1.5	0.5	23.0	50.7
NNES	5.3	13.2	3.1	1.9	0.4	14.5	38.4

Table 6- 16: Percentages of IMD Subcategories per 5000 words out of total IMDs in the NES and NNES texts

Rank Order	NES	NNES
1	Self references (45)	Self references (38)
2	Emphatics (25)	Emphatics (34)
3	Hedges (18)	Hedges (14)
4	Evaluatives (8)	Evaluatives (8)
5	Commentaries (3)	Commentaries (5)
6	Appeals (1)	Appeals (1)
Total IMD	100%	100%

(Percentages are given in brackets: out of total IMD)

From Tables 6-15 and 6-16, it can be seen that there are similarities in the choice of IMD subcategories with the three most frequent subcategories (hedges, emphatics and self references) in both group texts. The use of ‘emphatics’ (e.g. *certainly, clearly, surely*) in both groups is high (25% in NES and 34% in NNES). The most frequent subcategory in both groups, however, is ‘self references’ (*I, we, my, our*) (45% in the NES and 38% in the NNES). Hedges (e.g. *perhaps, possibly, probably, tend to*) are also common in both texts.

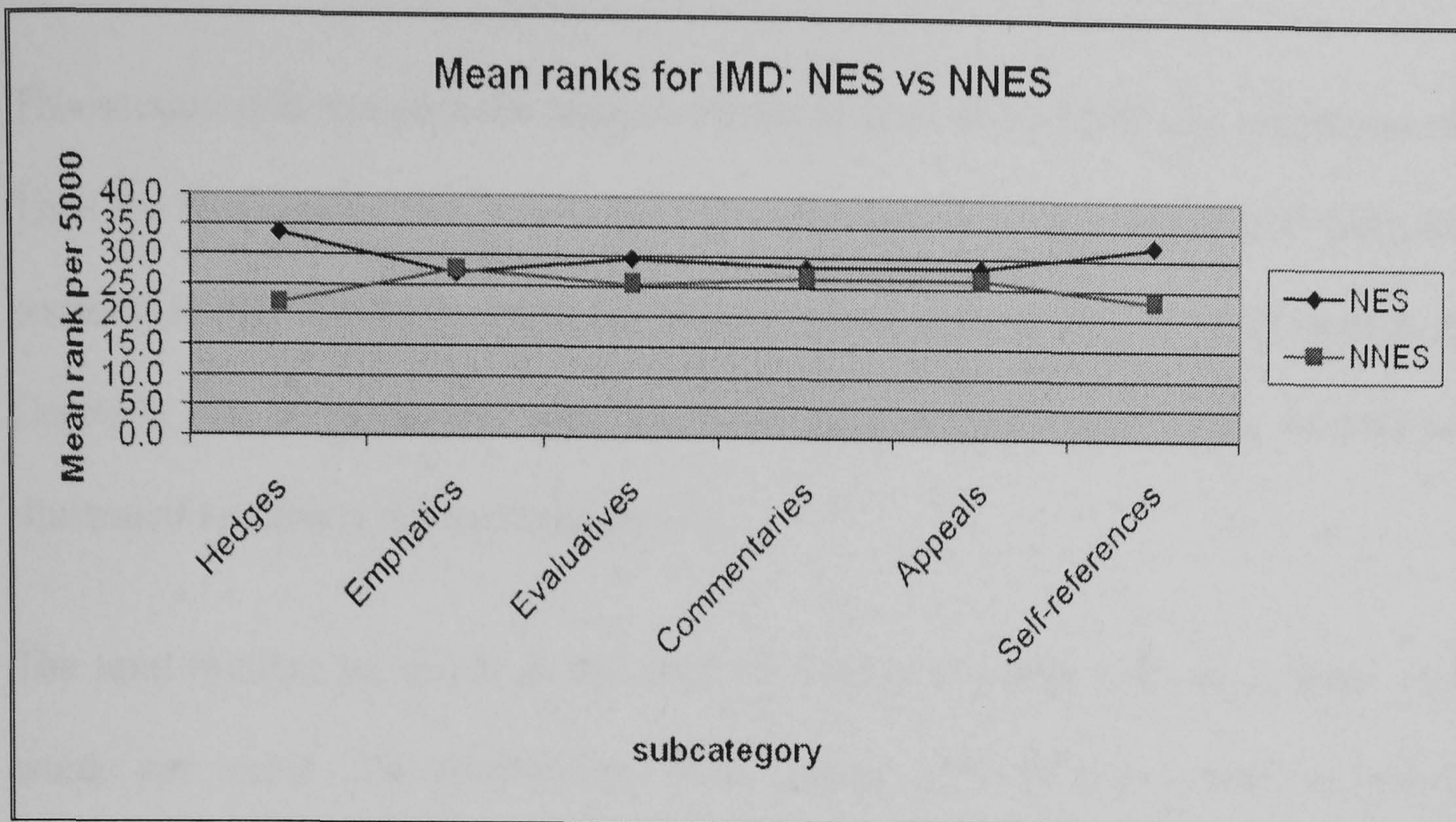
The NES writers used hedges and self references more often while the NNES writers made more use of emphatics and commentary markers (*you, your, imagine, allow me*).

The followings are mean ranks per 5000 words for IMD subcategories between the two group corpora analysed with the Mann-Whitney U test.

Table 6- 17: Mean Ranks for IMD Subcategories per 5000 words in the NES and NNES texts

Sample group	Hedges	Emphatics	Evaluatives	Commentaries	Appeals	Self references
NES	33.7	27.1	29.7	28.4	28.6	32.4
NNES	22.1	27.7	25.6	26.7	26.5	23.2

Figure 6- 2: Mean Ranks for IMD Subcategories per 5000 words in the NES and NNES texts



As can be seen from Table 6-17 and Figure 6-2 above, the NES writers used more IMD subcategories overall than their NNES counterparts, especially hedges and self-references, for which the differences are statistically significant (see Table 6-18: hedges: MWU test $z = -2.706$, $p < 0.007$; Self references: $z = -2.125$, $p < 0.034$).

Table 6- 18: MWU Tests for IMD Sub-category in the NES and NNES

	Hedges	Emphatics	Evaluatives	Commentaries	Appeals	Self references
Mann-Whitney U	206.500	354.000	307.500	339.500	335.000	240.000
Z	-2.706	-.147	-.955	-.423	-.678	-2.126
Asymp. Sig. (2-tailed)	.007	.883	.340	.672	.498	.034

a Grouping Variable: NES vs NNES

In summary, the NNES writers make more frequent use of TMDs than NES writers; the NES writers employed more IMDs than their NNES counterparts. In more detail, the NNES writers made greater use of sequencers, topicalisers, pre/reviews, code glosses and rationales while the NES writers employed more hedges, evaluatives and self-references.

6.3.3 MD Use in the PRO and Student texts

This section will compare the mean numbers of total textual MD and interpersonal MD. Then the frequencies and percentages of individual subcategories of MD from the two corpora, POSTCORP (Postgraduate Students Corpus) and RACORP (Research Article Corpus), will be presented. Also the mean rank of each subcategory of MD will be illustrated to answer research question 2.

The total number of words in the student texts is 223,444, with an average of 4,062 words per script. The professional texts contain 226,597 words with an average of 7,553 words per text. The average number of words in the professional texts is thus 3,492 words more than that in the student ones (see Table 6-19).

Table 6- 19: The Size of Data in the Student and PRO Texts

Groups	Total Words	Average Words
Student	223,444	4,062
PRO	226,597	7,553

Before discussing the frequency of each subcategory, a general presentation of the two major categories (TMD and IMD) in the student and PRO corpora is needed.

6.3.3.1 The Use of TMD in the Student and PRO texts

Table 6- 20: Mean Frequencies of total TMD and TMD Percentages out of MDs per 5000 words in the Student and PRO texts

Sample group	TMD	Total MDs	Percentage of TMDs in MDs
Student	92.30	136.8	68%
PRO	96.80	144.46	67%

The Table shows that the average frequency of total TMD in the student texts is 92.30 per 5000 words. In the PRO texts, it is 96.80. This use of TMD in both texts is similar to each other with similar proportion of TMD out of total MD (68% in POSTCORP and

67% in RACORP). There are no significant differences in the use of total TMD between the groups (MWU test, $z = -.541$, $p < 0.588$).

Table 6- 21: Mean Frequencies of TMD Subcategories per 5000 words in the Student and PRO texts

Sample group	Sequencers	Topicalisers	Concluders	Pre/reviews	Additives	Concessives	Code Glosses	Rationales	Sources	Total TMD
Student	16.9	1.5	1.0	2.9	18.3	23.6	15.2	6.6	6.3	92.3
PRO	10.4	0.6	2.0	4.0	16.9	28.6	19.8	7.5	7.0	96.8

Table 6- 22: Percentages of TMD Subcategories per 5000 words out of total TMDs in the Student and PRO texts

Rank Order	STUDENT	PRO
1	Concessives (26)	Concessives (29)
2	Additives (20)	Code Glosses (21)
3	Sequencers (18)	Additives (17)
4	Code Glosses (16)	Sequencers (11)
5	Rationales (7)	Rationales (8)
6	Sources (7)	Sources (7)
7	Pre/reviews (3)	Pre/reviews (4)
8	Topicalisers (2)	Concluders (2)
9	Concluders (1)	Topicalisers (1)
Total TMD	100%	100%

(Percentages are given in brackets: out of total TMD)

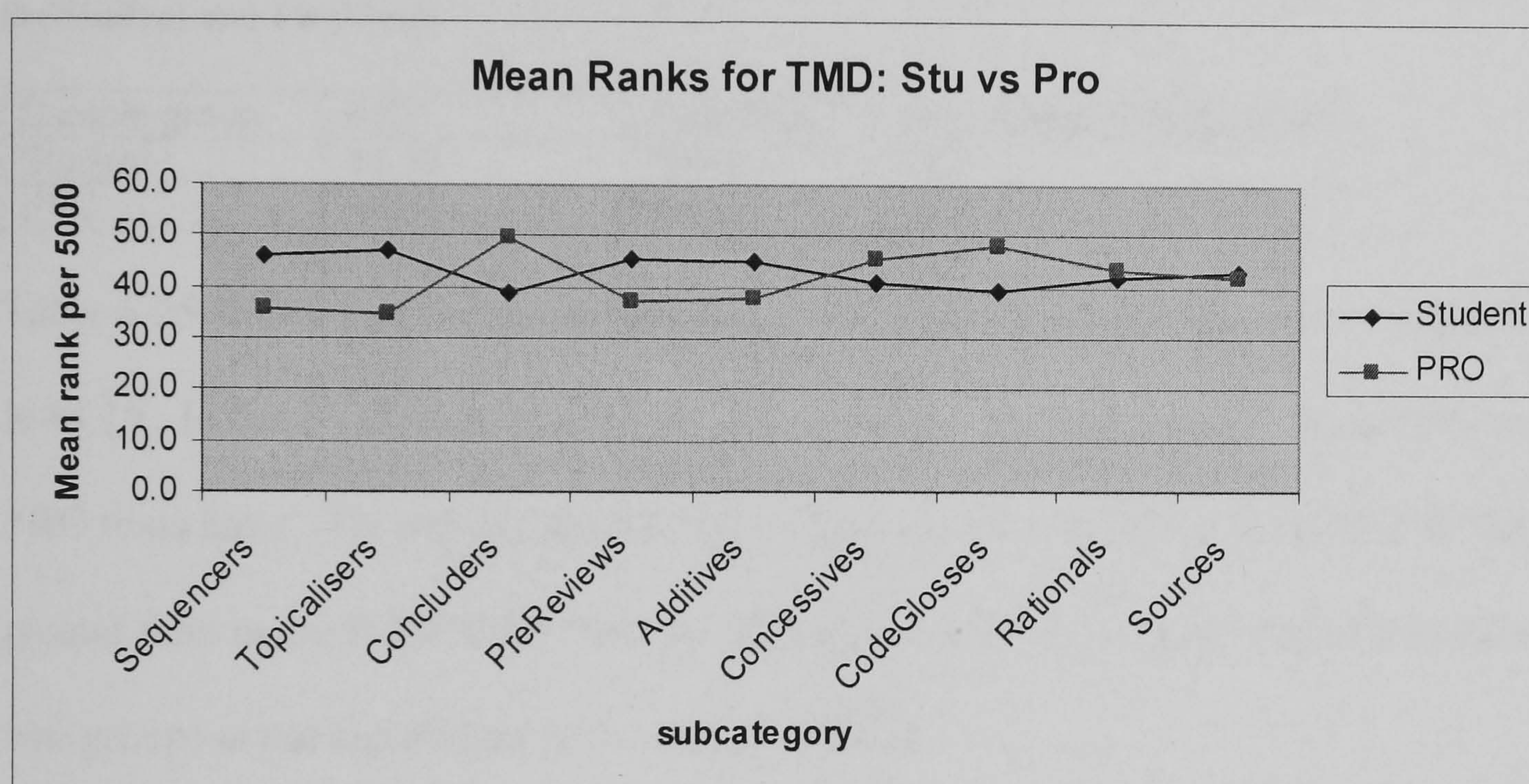
Tables 6-21 and 6-22 show that, in the textual MD, both group writers used concessives, additives, sequencers and code glosses more often than the other subcategories. The most frequent sub-category in both groups is concessives (29% in the RACORP and 26% in the POSCORP in Table 6-22). Additives (e.g. *furthermore*, *in addition*, *moreover*) are also more common in both texts (20% in the student texts and 17% in the journal article texts). Concessives (e.g. *however*, *on the other hand*, *by contrast*, *whereas*), code glosses (e.g. *defined as*, *for example for instance*, *i.e. in other words*, *known as*, *such as*) and concluders (e.g. *in conclusion*, *in sum*, *in summary*, *so far*, *to conclude*, *to summarise*) have been more used by the journal article writers while sequencers (e.g. *first*, *finally*, *lastly*, *next*), topicalisers (e.g. *in regard to*, *as for*, *in relation to*) and additives (e.g. *in addition to*, *in the same way*, *moreover*, *therefore*, *furthermore*) were more used by the student writers (see Table 6-21).

The Mann-Whitney U test was employed and the mean ranks from the test are as follows:

Table 6- 23: Mean Ranks for TMD Subcategories per 5000 words between the Student and PRO texts

Sample group	Sequencers	Topicalisers	Concluders	Pre/reviews	Additives	Concessives	Code Glosses	Rationales	Sources
Student	46.2	46.9	38.6	45.3	45.1	40.7	39.2	41.9	42.7
PRO	35.7	34.5	49.5	37.4	37.7	45.6	48.3	43.5	42.0

Figure 6- 3: Mean Ranks for TMD Subcategories per 5000 words between the Student and PRO texts



From Table 6-23 and Figure 6-3 we see that the student writers used more sequencers, topicalisers, pre/reviews and additives than the PRO writers while the PRO writers employed more concluders and concessives, code glosses and rationales. However, there was no statistically significant difference in these subcategories (see Table 6-24) apart from topicalisers which were used more by the student writers (Mann Whitney, $z = -2.292$, $p < 0.022$) and concluders which were more frequently used by the PRO writers (Mann Whitney, $z = -1.981$, $p < 0.048$).

Table 6- 24: MWU Tests for TMD Sub-category in the Student and PRO

	Sequence rs	Topical- sers	Conclu- ders	Pre/ reviews	Additives	Conce- ssives	Code Glosses	Rationales	Sources
Mann- Whitney U	607.00	572.00	600.00	658.50	666.00	714.50	634.50	779.50	795.00
Z	-1.895	-2.292	-1.981	-1.436	-1.344	-.892	-1.638	-.285	-.140
Asymp. Sig. (2- tailed)	.058	.022	.048	.151	.179	.373	.101	.776	.889

a Grouping Variable: Student vs PRO

6.3.3.2 The Use of IMD in the Student and PRO Texts:

Table 6- 25: Mean Frequencies of total IMD and IMD Percentages out of MDs per 5000 words in the Student and PRO texts

Sample group	IMD	Total MDs	Percentage of IMDs in MDs
Student	44.19	136.8	32%
PRO	47.66	144.46	33%

Table 6-25 shows that the mean frequency of total IMD per 5000 words in student texts is 44.19. In the PRO texts, it is 47.66. Density-wise, the student texts have 32% and PRO texts have 33% out of total MD use. Thus, the use of IMD in RACORP is slightly greater than in the POSCORP. But the difference in the use of total IMD between the two groups is not significant ($z = -.322, p < 0.747$).

However the frequency of IMD subcategories is not consistent across the sub-categories. While the student writers make greater use of emphatics (e.g. *always, certainly, clearly, definitely, obviously*), evaluatives (e.g. *appropriately, disappointingly, more importantly, fortunately*) and commentaries (e.g. *you might be missing, you might add, imagine, let us*), the journal article writers make more use of hedges (e.g. *perhaps, possibly, presumably, probably, seems, tend to*) and self-references (e.g. *I, my we, our*), as seen in the table below.

Table 6- 26: Mean Frequencies of IMD Subcategories per 5000 words in the Student and PRO texts

Sample Group	Hedges	Emphatics	Evaluatives	Commentaries	Appeals	Self references	Total IMD
STU	7.2	13.0	3.4	1.7	0.5	18.3	44.10
PRO	10.5	11.9	2.8	1.1	0.5	20.8	47.60

Table 6- 27: Percentages of IMD Subcategories out of Total IMDs in the Student and PRO texts

Rank Order	Student	PRO
1	Self references (42)	Self references (43)
2	Emphatics (29)	Emphatics (25)
3	Hedges (16)	Hedges (22)
4	Evaluatives (8)	Evaluatives (7)
5	Commentaries (4)	Commentaries (2)
6	Appeals (1)	Appeals (1)
Total IMD	100%	100%

(Percentages are given in brackets: out of total IMD)

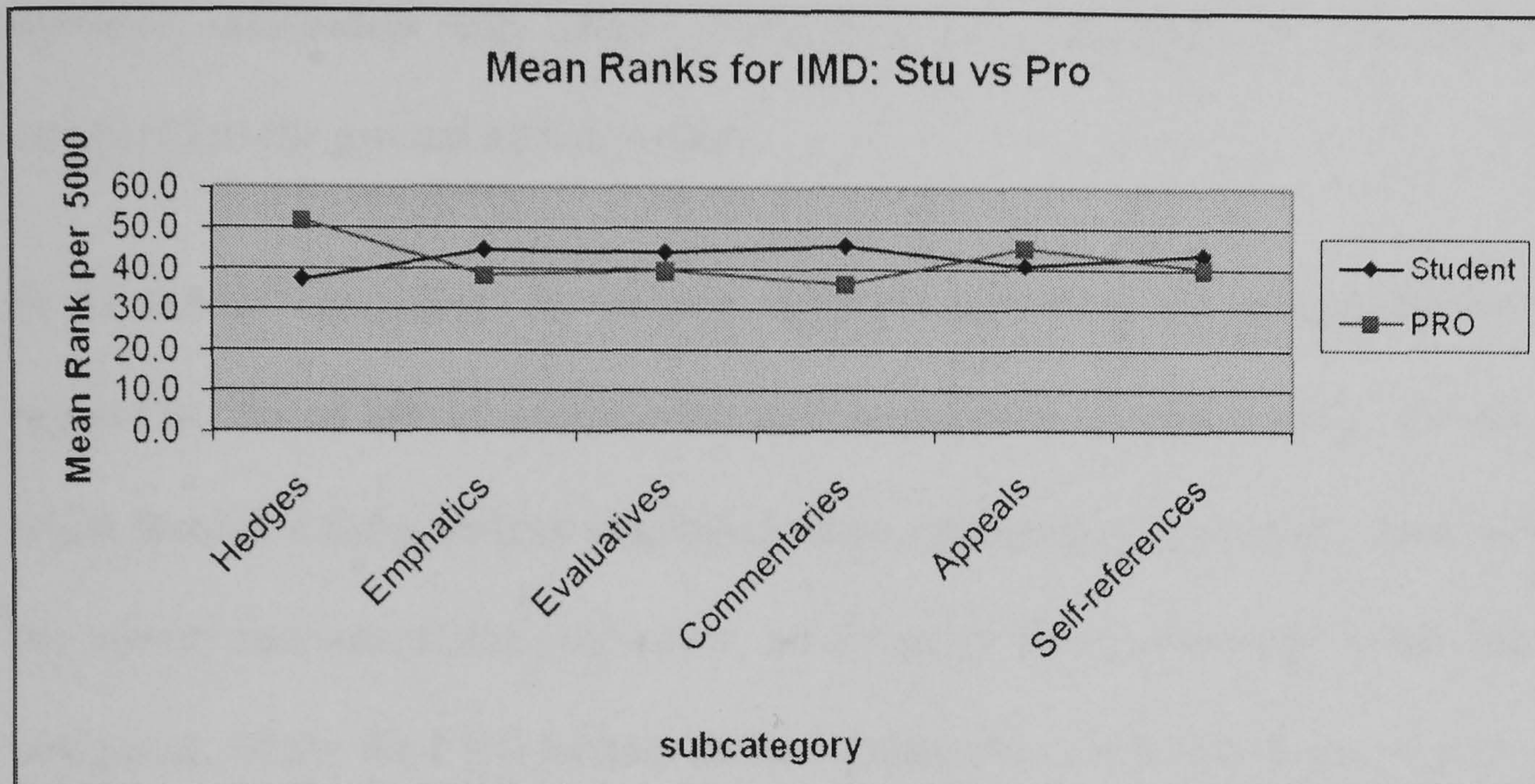
There are similarities between the three most frequent subcategories (hedges, emphatics and self-references) in both group texts. The most frequent sub-category in both groups is self-references (42% in the POSCORP and 43% in the RACORP), and the second most is emphatics. Hedges are also far more common in both texts with more frequent use in PRO texts (16% in the student and 22% in the PRO in Table 6-27).

Some of the IMD subcategories (emphatics, evaluatives, and commentaries) were used more frequently by the student writers while hedges were employed more by the journal article writers. These differences between the two groups can be revealed by the mean ranks from the MWU tests as follows:

Table 6- 28: Mean Ranks for IMD Subcategories per 5000 words in the Student and PRO Texts

Sample group	Hedges	Emphatics	Evaluatives	Commentaries	Appeals	Self references
Student	37.3	44.7	44.1	45.9	40.8	43.7
PRO	51.8	38.3	39.4	36.3	45.4	40.2

Figure 6- 4: Mean Ranks for IMD Subcategories per 5000 words in the Student and PRO texts



As can be seen from the mean ranks in Table 6-28 and Figure 6-4, the student writers used emphatics, evaluatives, commentaries and self-references more frequently than the PRO writers while hedges and appeals were employed more by the PRO writers. The differences are not statistically significant apart from hedges which were more frequently used by the journal article writers (MWU test, $z = -2.605$, $p < 0.009$, see Table 6-29).

Table 6- 29: MWU Tests for IMD Sub-category in the Student and PRO

	Hedges	Emphatics	Evaluatives	Commentaries	Appeals	Self references
Mann-Whitney U	531.000	686.500	718.500	625.000	720.500	742.500
Z	-2.605	-1.153	-.855	-1.902	-1.099	-.630
Asymp. Sig. (2-tailed)	.009	.249	.393	.057	.272	.529

a Grouping Variable: Student vs PRO

The overall percentages of the main categories in student and journal article texts are similar (68% for TMD and 32% for IMD in RACORP and 67% for TMD and 33% for IMD in POSCORP from tables 6-20 and 6-25). However, the frequency of subcategories in TMD and IMD is not consistent across the sub-categories. For example, concluders and hedges were used less frequently by the student writers compared to the journal article writers.

In the textual categories, the student writers made more use of topicalisers (e.g. *with regard to, based on, in relation to*) and sequencers (*firstly, finally, list of numbers*) while journal article writers employed more concluders (*to sum up, in conclusion, on the whole*) and concessives (*however, on the other hand, whereas*). In the interpersonal categories, while the PRO writers made significantly (MWU test, $z = -2.605$, $p < 0.009$) more use of hedges (*might, perhaps, seems, probably*) the student writers more frequently used emphatics (*always, clearly, especially*).

Hence the general overview of the corpus analysis reveals broad commonalities in the main categories of MD (e.g. TMD and IMD) choice is concerned in the comparisons of the student texts (the NES and the NNES) and student and journal article writing. However, considerable differences in the frequency of individual MD subcategories between groups in the same discipline were found.

In the following sections, the results from individual subcategories will be used to explore the differences and similarities between the two group texts in more detail in the frequency of MD in general and when and how they use MD in particular (the NES vs NNES and the Student vs PRO) to answer research question 1: What are the types and frequencies of MD used/ how differently are they used in non-native English speaker (NNES) student writings and native English speaker (NES) student writings?,

and research question 2: What are the types and frequencies of MD used/ how differently are they used in each genre (i.e. journal article/Master's assignment)?

6.4 Results from Individual Subcategories in TMD

Talking about general figures in the last section does not give a clear picture of the differences and similarities in the use of MD. So individual subcategories of the main categories (TMD and IMD) will now be presented focusing on when and how the MD features are used in the different group texts.

The tables below show the mean frequencies of TMD sub-categories including the indications of statistical difference for the comparisons of two groups (the NES vs NNES and the Student vs PRO). Also the percentages of each TMD sub-category are provided to compare TMD frequency between groups.

Two clear results can be observed in Tables 6-30 and 6-31. The first relates to 'language differences' (i.e. native versus non-native) in the comparison between the NES (native English speaker) and NNES (non-native English speaker). As shown in Table 6-30, the NES group uses significantly less TMD than does the NNES (MWU test, $z = -2.732$, $p < 0.006$).

Table 6- 30: Mean Frequencies of TMD Subcategories per 5000 words with the indication of the Statistical Difference in the NES and NNES texts

Subcategories of TMD	NES	NNES	Statistical Differences
Sequencers	11.67	21.62	✓ (NES<NNES) (p<0.038)
Topicalisers	1.24	1.94	Non sig.
Concluders	1.24	0.84	Non sig.
Pre/reviews	2.58	3.24	Non sig.
Additives	17.45	19.01	Non sig.
Concessives	26.10	21.45	Non sig.
Code Glosses	12.76	17.16	✓ (NES<NNES) (p<0.003)
Rationales	5.98	7.11	Non sig.
Sources	7.41	5.26	Non sig.
TOTAL TMD	86.43	97.63	✓ (NES<NNES) (p<0.006)

Table 6- 31: Percentages of TMD Subcategories out of total TMDs in the NES and NNES texts

Rank Order	NES	NNES
1	Concessives (30.2)	Sequencers (22.1)
2	Additives (20.2)	Concessives (22.0)
3	Code Glosses (14.8)	Additives (19.5)
4	Sequencers (13.5)	Code Glosses (17.6)
5	Sources (8.6)	Rationales (7.3)
6	Rationales (6.9)	Sources (5.4)
7	Pre/reviews (3.0)	Pre/reviews (3.3)
8	Topicalisers (1.4)	Topicalisers (2.0)
9	Concluders (1.4)	Concluders (0.9)
Total TMD	100%	100%

(Percentages are given in brackets: out of total TMD)

The second (see Table 6-32) relates to ‘genre differences’ (i.e. journal articles versus student assignments); we see that the journal article texts display a higher frequency in total TMD than do student texts (MWU test, $z = -.541$, $p < 0.588$).

Table 6- 32: Mean Frequencies of TMD Subcategories per 5000 words with the indication of the Statistical Difference in the Student and PRO texts

Subcategories of TMD	Student	PRO	Statistical Differences
Sequencers	16.90	10.41	Non sig.
Topicalisers	1.49	0.64	✓ (Stu>PRO) ($p < 0.022$)
Concluders	1.03	2.05	✓ (Stu<PRO) ($p < 0.048$)
Pre/reviews	2.93	3.95	Non sig.
Additives	18.28	16.88	Non sig.
Concessives	23.63	28.57	Non sig.
Code Glosses	15.22	19.76	Non sig.
Rationales	6.58	7.50	Non sig.
Sources	6.27	7.04	Non sig.
TOTAL TMD	92.39	96.8	Non sig.

Table 6- 33: Percentages of TMD Subcategories out of total TMDs in the Student and PRO texts

Rank Order	Student	PRO
1	Concessives (25.6)	Concessives (29.4)
2	Additives (19.8)	Code Glosses (20.8)
3	Sequencers (18.4)	Additives (17.4)
4	Code Glosses (16.5)	Sequencers (10.8)
5	Rationales (7.1)	Rationales (7.6)
6	Sources (6.8)	Sources (7.2)
7	Pre/reviews (3.0)	Pre/reviews (4.0)
8	Topicalisers (1.7)	Concluders (2.1)
9	Concluders (1.1)	Topicalisers (0.7)
Total TMD	100%	100%

(Percentages are given in brackets: out of total TMD)

The results from the individual subcategories will now be presented to find out how the major factors govern the MD usage in their academic writing. Full details of the figures for each sub-category of MD are presented in the tables in appendices 6-5 and 6-6, which can be referred to by the reader throughout the following sections.

6.4.1 Sequencers

Sequencers, as the name suggests, refer to the MD features that signal how propositions or pieces of information are ordered (e.g. *first, second, lastly, finally*).

6.4.1.1 NES vs NNES

In the NES texts, this subcategory is ranked fourth while the NNES writers used it the most. In terms of occurrence, Tables 6-30 and 6-31 show that in the NES texts, sequencers are used at the rate of 11.67 per 5000 words with a density of 13.5%. In the NNES texts they are used 21.62 per 5000 words with a density 22.1%. There is therefore a tendency among non-native speakers to use more sequencers than their NES counterparts with a statistical difference (Mann Whitney U, $z=-2.073$, $p<0.038$).

The result shows that the vast majority of sequencers in the typology used here are found in both types of texts; the full figures are given in appendix 6-5-1. Examples of these are: *first, second, firstly, secondly, finally, first of all, listing (a,b,c and 1,2,3),subsequently*. In addition to these common types, the NNES writers also used sequencers such as *thirdly* and *lastly*, as in (sequencers italicised by researcher):

Firstly, it is easy and convenient for researchers to control and present the questions. *Secondly*, Questionnaire can be used in either small scale, such as in office, home, or in large scale, in big class, in army and public. *Thirdly*, there is no strict limit for collecting data. Questionnaires can be collected all at once as they are completed in classroom, for example; or respondents can also choose suitable time to return questionnaires (NNES 3)

The first section of the paper illustrates the theoretical background of this study. After a brief review of the discourse analysis approach to language teaching, the model used for this study is presented. *In the second section*, the data collected for this study is presented and analysed following Sinclair and Coulthard's coding scheme. The focus of the analysis is on the teacher-initiated three-part exchange sequence of Initiation-Response-Follow-up (IRF) (Sinclair and Coulthard 1975; 1992). *Lastly*, the paper discusses the implication of the findings for second/foreign language teaching. (NNES 17)

The mapping function is achieved in the extracts above, which contain sequencers (*firstly, secondly, the first section, in the second section*) to orient the reader and give them a hint of what is to come next. Also *thirdly* and *lastly* are used in the NNES texts to describe what will finally follow the previous propositional contents whereas these were not found in the NES corpus.

At times this mapping function (e.g. listing with numbers *2, 3, 4, and 5*) in the NES texts combining with self-references (*I* and *my* which will be discussed later) results in a more self-promotional function. Note how the following passages promote subsequent parts of the assignment by using sequencers (*section 2, section 3, section 4 and section 5*) to the readership and encourage the reader to read on:

.... I set out the theoretical background to my research questions in *section 2*; detail my research methodology in *section 3*, provide my analysis in *section 4* and evaluate my findings, with a conclusion, in *section 5* (NES 23).

However, the context of the discourse suggests that certain features which would be classified as additives can also be considered sequencers, as in the following extract:

The natural content facilitates language acquisition by provision of an abundance of authentic interaction. *In addition*, Cook maintains that such exposure to authentic material can help students to fill the language gap when put in a similar situation as that portrayed. *Moreover*, this particular extract involves L2 users in authentic communicative interaction which is rare in EFL material (NES 13).

The writers used *in addition* and *moreover* as parts of a process of listing and these can be classified as sequencers. They also exemplify the multifunctional nature of MD as already mentioned in the section 'characteristics of MD' in Chapter 3. However, for the purpose of the quantification of the frequency, I followed other researchers, as mentioned in chapter 5, and classified them as additives for the counting of MD frequencies because their primary function appears adding and reinforcing information. Thus these items were not counted as sequencers for the frequency of MD.

While the sequencers were only the fourth most common TMD subcategories in the NES corpus, they were the most common in the NNES texts.

6.4.1.2 PRO vs Students

The subcategory of sequencers is ranked fourth in the journal article texts, among all the TMD subcategories, while it comes third in the student texts (Table 6-33). Tables 6-32 and 6-33 show that sequencers are used 16.90 times per 5000 words in the student texts with a density of 18.4%. In the PRO texts they are used 10.41 per 5000 words with 10.8% density. The sequencers were used more frequently in the student texts.

The student texts have almost identical patterns of sequencers for the presentation of the structure of their assignment, found at the end of the introduction sections in the student texts:

The purpose of this essay is to explore the means by which focus and ground are evident in modern Turkish in terms of Vallduvian information packaging and to consider areas that need to be developed. *Firstly*, a brief background into this research will clarify terminology and delimit the scope of the assignment. *In the second section*, an attempt will be made to show how Turkish utilises syntax and phonology to show focus and ground. *In the third section*, a brief look at how some of the monoclausal syntactic data from Turkish might be accounted for using an HPSG framework will be looked at. *Finally*, the conclusion will highlight some areas for which further research is required (NES 24).

In its detailed justification of the use of video camera in the language classroom, this paper focuses on the following points; 1) stimulus: activities – video project, in-class recording and role-play; 2) evaluation and self evaluation strategies – language analysis, communicative approach analysis, and feedback; 3) strategies awareness (NNES 12)

This analysis, predictably, will be split into two parts; *the first part* will analyse the authentic text, and *the second part* will analyse the simplified text taken from an intermediate level E.L.T. coursebook. (NES 7)

Firstly, this essay will talk about the definition and the history of sustainable development, the SD is thirty year old and is a concept about how to use the ecological recourse reasonable and correctly, and then will discuss its three aspects: Economic aspect, Social aspect, Economic aspect. *The next main part*

will talk about that the SD will perfect itself to meet the rule of human society to become success. (NNES 12)

The student writers in the above extracts describe the order of the texts by employing a similar pattern of sequencers (*firstly, in the second section, in the third section, finally, lists of no, the first part, the second part, the next main part*) to explain how the writing is sequenced, whereas these features are much less commonly present in the journal articles. The following extract shows an alternative method of using sequencers in the journal article through additives (*in addition*) and rationales (*because of this*):

Each of these studies suggests various roles that language play may have in facilitating SLL. Taking this prior research as a base, for the present study I examined the playful interaction of three adult L2 speakers of English interacting outside of the classroom with NSs of English and asked whether and how this play may have furthered their SLL. Using Tarone's (2000) suggestions for ways in which language play may aid SLL as a starting point, I examined my data for potentially facilitative patterns of interaction that occurred in the humorous language play of the three participants. In this case, the learners are much more advanced than those of the studies mentioned above. *In addition*, whereas Davies' participants interacted with sympathetic NSs, the NNSs in this study recorded their conversations with a wide variety of NSs, from family members to strangers, and thus were not always certain to encounter a NS willing to collaborate with them. Thus, the examples I present of humour are not always co-constructed, or even apparently humorous to both parties. *Because of this*, I begin with a definition of humour/language play. (PRO 1)

This extract also demonstrates the multifunctionality of MDs, in that the writer employed features of additives and rationales (*in addition*, and *because of this*) to describe the sequence of the text structure. These features were used to show the reader the writers' intention within a running discourse in order to imply what will come next with reasons and logic of the paragraph by using the features of rationales and additives for the functions of sequencers.

Another passage by a PRO writer uses *both* to imply what propositions of information are included within a chapter ('how new technologies facilitate acquisition of L2

literacies and what L2 literacies are needed for learners to participate in an increasingly digital world') instead of using overt devices of sequencers (*first, second, last*), as in:

The title is also ambiguous in that, on first glance, it seems to refer only to how technologies can support L2 literacy development; however, there is the additional reading of what L2 literacies are needed to use the technologies, an equally important focus for L2 education because knowing the literacies required to use the new technologies is often critical for learners to meet their social, personal, and educational needs (Goodwin-Jones, 2000). This chapter will therefore discuss *both* how new technologies facilitate acquisition of L2 literacies and what L2 literacies are needed for learners to participate in an increasingly digital world (PRO 5)

6.4.2 Topicalisers

As classified in this work topicalisers cover the MD features that guide the reader to what is happening at certain points in the text like topic shift, main topic of sections, paragraphs, or sentences. They include expressions such as *in regard to, move on, turn to, with regard to, as for, based on, bearing in mind* which were used for the quantification of the frequency as opposed to the phrases which have implicit meanings for the functions of topicalisers, as in the PRO texts, and will be described in the next section.

6.4.2.1 NES vs NNES

NES writers used 1.24 items per 5000 words while the NNES corpus contains 1.94 items per 5000 words. This means the NNES writers employed 0.7 items more than their NES counterparts per 5000 words (56% more). This difference is not statistically significant ($z = -.148, p < 0.883$) due to the uneven distribution of data (see Appendix 5-10, normality test) but nonetheless seems important. As for the density, it is 1.4% and 2% in the NES and NNES texts respectively.

In terms of the frequency of topicalisers, the NES writers and their NNES counterparts show a different choice of phrases in this category. For example, *move on*, *in relation to*, *in respect to* were used in the NNES texts while these features were not found in the NES corpus (for fuller information the reader is referred to Appendix 6-5-2).

NES: *in regard to* (8), *with regard to* (5), *turn to* and *as for* (4 respectively), *based on* (2), *move on* (0), *bearing in mind* (0), *in relation/respect to* (0)

NNES: *based on* (14), *with regard to* and *as for* (10 respectively), *in regard to* and *in relation/respect to* (4 respectively), *turn to* (2), *move on* (1)

One similarity is that both groups of writers used the items which concern the topic of sections or paragraphs (e.g. *in regard to*, *as for*, *based on*) more frequently than topic shift (*turn to*, *move on*):

In regard to the research questions, they were clearly stated and were closely associated with the underlying hypothesis (NNES 1)

Nevertheless, just *as for* morphosyntax, the interface between universals and transfer remains unclear for phonological issues such as acquisition of final consonants, especially consonant clusters... (NES 23)

Both writers used *in regard to* and *as for* to indicate the main topics of sentences, to be dealt with later.

In the next extracts, *based on*, which was most frequently used in the NNES corpus, was employed to guide the reader to the main topic of sentences as in:

Based on Flander's and Moskowitz's descriptive system, Sinclair.... (NNES 10)

Based on this definition, four functions are found among..... (NNES 13)

The NNES students employ topicalisers more often than their NES counterparts, using most of the topicaliser devices such as *based on*, *with regard to*, *as for*, *in regard to*, *in relation to*, and *in respect to*. However both groups of students used the items which

concern the topic of sections (e.g. *in regard to, as for, based on*) more frequently than topic shift (*turn to, move on*), suggesting a similarity in their use of topicalisers.

6.4.2.2 PRO vs Students

The table shows that in the student texts topicalisers are used 1.49 times per 5000 words with a density of 1.7%. In the journal article texts they are used 0.64 times per 5000 words with 0.7%. The topicalisers were employed more frequently by the student writers, such as topic shift and main topic of sentences, with significant difference ($z = -2.292, p < 0.022$).

The student writers and their counterparts show similarities in this subcategory, with the two most frequent phrases in the two corpora being *with regard to* and *based on*.

With regard to population, the participants in this research study were students at school of professional and Continuing Education at the University of Hong Kong. (NNES 1)

With regard to accuracy rates, their judgements of wh-structures may still derive from their knowledge of UG principles (PRO 3).

Based on research carried out by Riseborough (1981-cited in Kellerman, 1992) it appears that parallel processing occurs of information coming from auditory and visual channels (NES 13).

Based on recent research on the role of cognition in L2 learning, L2 pedagogy,.... (PRO 10)

However, as can be seen from Appendix 6-5-2 (raw frequency of topicalisers), topicalisers are far more common in the student texts with a limited range of devices (i.e. some of the phrases were not found in the student sub-corpus, NES) to express topicality whereas the professional writers employed a broad range of topicalisers.

As for the students, it seems clear that L2 under graduates in the US do not all fit the proverbial descriptions of international students as gushing with respect (PRO 5)

Except in rare instances and brief glimpses, the teachers of L2 undergraduates have usually explained their expectations, plans, opinions *in relation to* L2 students as a group, not as individuals (PRO 30).

In regard to the research questions, they were clearly stated and were closely associated with the underlying hypothesis (PRO 1).

As for the student, *in relation to* and *in regard to* can act as topicalisers in the above extracts to show the topic of sentences and to spell out what will follow.

At times the PRO writers use implicit meanings for the functions of topicalisers as in:

Developing the ideas presented in Cook (1997), Cook (2000) argues for the importance of language play for adult language learning, as well as for child language acquisition (PRO 1)

Let me refer to studies like these as ‘POS-oriented’ accounts of how UG constrains grammatical representations in SLA, to distinguish them from the ‘difference-oriented’ accounts of access to UG described earlier (PRO 3).

The writers started sentences with ‘developing.....’ and ‘let me refer to’ to signal the topics of sentences instead of using overt topicalisers which have been defined in the proposed typology and used for counting the frequency. Hence the journal article writers use the broad range of phrases, which were not defined in the proposed typology, to imply the function of topicalisers instead of only using overt features of topicalisers which were defined in the proposed typology and were used to count the frequency.

The students employ topicalisers significantly (MWU test, $z = -2.292$, $p < 0.022$) more than their PRO counterparts. However the features of this sub-category used by journal article writers are broader (starting the sentences with varied features, such as ‘developing.....’ and ‘let me refer to’) while the student writers mainly use the overt devices of topicalisers (*with regard to* and *based on*).

6.4.3 Concluders

With reference to tables 6-30 and 6-32, it can be observed that concluders is one of the least used subcategories in TMD from all groups.

6.4.3.1 NES vs NNES

Concluders were used only 1.24 times per 5000 words in the NES corpus, and only 0.84 occurrences in the NNES texts. Their use is ranked 9th, the least frequent subcategory in both corpora. The density in the NES texts is 1.4% and in the NNES it is 0.9%, both extremely low.

Examples of the ‘concluders’ that were used are: *in conclusion, overall, so far, thus far, on the whole, to conclude* and *to sum up*. Of these two group texts, *overall* was the most commonly used item as in:

His results (see Appendix B) show the linguistic distinction between definiteness and indefiniteness, and the pragmatic distinctions between echo and non-echo contexts. *Overall*, accuracy rates are higher in all contexts that require the definite article. And, in echo contexts the accuracy rate is considerably lower, which may stem from the hearer accommodating their own speech to that of the speaker (NES 11).

Overall, there are 7 paragraphs with five characters (the first, second and third little pigs, the big bad wolf and mother in this story. Like other children stories, the repetitive element like purely (NNES 29)

Both writers used *overall* to summarise the contents that have been mentioned up to now.

In the following extract, *in conclusion* was employed with a similar function to *overall* in the NES text:

In conclusion after closely examining leadership through the literature, "there is nothing more practical than good theory. It can help the practitioner to unify and

focus his views on an organisation, on his role and relationships within the organisation and on the elusive phenomena of leadership..... (NES 1)

The writer employed the phrase *in conclusion* to introduce a summary of what has been said in the previous section. This phrase was found three times in the NES corpus while it was not found in the NNES corpus.

As in the NNES text, the concluding phrase *to sum up* was used to establish a similar purpose to *in conclusion* in the above extract, as in:

To sum up, children before school ages have acknowledge the discourse procession and acquired various devices and strategies to initiate and reinitiate to get attention and response (NNES 13).

In addition to the preferred phrases mentioned above, there are three occurrences of *thus far* in the NES corpus as in:

We should recognise as teachers' that L1 can assist learning and not be afraid to include activities where it may take place. By doing this we are limiting the spheres in which learning and acquisition can take place. L1 Use in Teaching Resources *thus far*, I have argued that there is a role for use of the L1 by both teachers and students. In addition we must look at the tools which assist learning, the issue of teaching resources (NES 15).

The purpose of this review is to discuss what has been written *thus far* about two concepts which are central to this assignment, namely (NES 7)

Thus far indicates that a topic which has been mentioned in a previous section will be dealt with at a later point. In other words, the authors try to summarise what they have been writing. This phrase was employed three times in the NES texts but did not occur in the NNES corpus.

A comparison of the NES student writing with the NNES ones shows that the two sets of writings are similar in terms of the rank order among the TMD subcategories.

However there are some differences in the preferred devices in the subcategory of

concluders, such as *in conclusion* and *thus far* which were found in the NES corpus whereas no instance of these items was found in the NNES texts.

6.4.3.2 PRO vs Students

From tables 6-32 and 6-33, it can be seen that the student writers employed 1.03 instances per 5000 words with a density of 1.1%. In the PRO texts, on the other hand, they are used 2.05 per 5000 words with 2.1% density. The concluders were significantly more used (99% more) by the journal article writers (MWU test, $z = -1.981$, $p < 0.048$).

As far as variety is concerned, the PRO writers used a broad range of concluding phrases whereas the student writers employed limited devices of concluders (see appendix 6-5-3). To compare the frequency of phrases for concluders in each group, their frequency of occurrence is listed as follows:

POSCORP: overall (20), so far (7), in conclusion (3), thus far (3), to conclude (3), to sum up (3), on the whole (2),

RACORP: overall (72), in sum (11), on the whole (6), so far (5), thus far (4), in summary (3), to conclude (2), in conclusion (1), to sum up (1), to summarise (1)

The student writers and their counterparts are similar in their choice of phrases in this subcategory, with the most frequent phrase being *overall*, although its frequency in the RACORP is much higher than in POSCORP.

There are similarities in the use of concluder devices with similar purposes, such as *overall*, *in conclusion*, *in sum*, *thus far*; they are used to summarise what has been done up to now.

For example, *thus far* was employed 4 times in the RACORP while it was used three times in the POSCORP, although it does not occur in the NNES sub-corpus. In RACORP, one of the four occurrences of the concluder phrase *thus far* was used to show that a topic which had been mentioned in a previous section might bring up some questions which the author anticipates. By throwing these questions to the reader, the writer can get more attention and encourage the reader to read on:

The lawful pattern of data observed *thus far* leads one to ask, how do people "know" what latency values are most (un)acceptable for a given evaluative dimension, and what source of information are they relying upon? From a cognitive perspective, it seems unlikely that people have mathematical values stored in long-term memory for each individual speech act (PRO 18).

The other three instances of *thus far* in the RACORP were used for similar purposes to what the NES students did in order to summarise what has been done in the previous section.

The PRO writers make significantly ($p < 0.048$) more use of concluders with a broader range of features compared to their student counterparts.

6.4.4 Pre/reviews

The subcategory of pre/reviews is one of the least used TMD subcategories.

6.4.4.1 NES vs NNES

There are only 2.58 occurrences of pre/reviews per 5000 words in the NES texts compared with 3.24 in the NNES ones, making the density 3.0 in the NES and 3.3 in the NNES texts respectively. So the pre/reviews used by the NNES writers are slightly more than by the NES ones.

Appendix 6-5-4 (raw frequency of pre/reviews) shows a group preference in the use of items in this subcategory. While *previously/earlier* were employed more in the NES texts, *above* was used more frequently in the NNES texts (see appendix) to direct the reader in connection with what has been going on in the previous parts of the text before moving on:

As was *previously* explained CLT was the best method for this group. Following the five basic principles contributed to the overall success of the activity (NES 21).

In the example given *earlier*, where the teacher reverted to using the L1 following comments from students, the students' lack of confidence prompted them to ask for use of the L1. (NES 15)

In the examples *above*, advertisers did not take a cross-cultural analysis into account before advertising abroad. This shows us again that standardization can fail in a lot of cases ... (NNES 2)

Below was used (10 and 11 times in the NES and NNES texts respectively) to alert the reader to what will come next. As can be observed from Appendix 5-8, this feature was used as MDs with the frequent use of the describing verbs (e.g. listed, presented, summarised, outlined, elaborated, provided, illustrated). In addition, they can act as sequencers, to signal to the reader how pieces of information are ordered in the following text. These extracts are therefore good examples of how the items can behave multifunctionally, as already discussed (in section 6.4.1.1):

Listed below are the eight roles with a brief description of their traits (NES 20)

Two tasks of such exercises are *provided below*. Task 1: Word Families Please ... (NNES 7)

This can be usefully represented as a series of cycles as illustrated in the figure *below*. (NNES 25)

The two sets of writings are similar in terms of the rank order among TMD subcategories although there are some differences in the preferred devices in the

subcategory of pre/reviews, such as *previously/earlier* which were employed more frequently in the NES texts while *above* was more frequently used in the NNES texts.

6.4.4.2 PRO vs Students

There are only 2.93 occurrences of pre/reviews per 5000 words in the student texts compared with 3.95 in the PRO ones, making the density 3.0 in the student texts and 4.0 in the journal article texts. The pre/reviews used by the student writers is thus 23% less than the PRO one. Despite the apparently higher number of pre/reviews in the PRO texts, the difference is not significant (MWU test; $z = -1.324$, $p < 0.186$).

Nevertheless the student writers and their counterparts differ in their choice of items. For instance, items which have a guiding function when the writer refers to diagrams, tables, or figures away from the main body of the text (e.g. *table x* and *figure x*), were more frequently used by journal article writers to clarify what the writer is saying to the reader as in:

As can be seen from *Table 1*, some differences among the groups were apparent as measured by a GLM ANOVA. One way to make the groups equal in proficiency would have been to eliminate many participants (PRO 7)

Figure 1 proposes that cognitive abilities are influential on the early input-based stages of language learning. (PRO 14)

On the other hand the pre/review elements *previously/earlier* and *above* were more frequently employed in the student texts to refer back to ground already covered, in order to guide the reader in connection with what has been going on in the previous parts of the text.

As has been mentioned *earlier*, preparation for participation in another culture requires more than simply learning its language, so teaching behaviours is of course beneficial to students (NES 25)

In the example given *previously*, where the teacher reverted to using the L1 following comments from students, the students' lack of confidence prompted them to ask for use of the L1. To be a good language learner you need to be a risk taker (NES 15).

It can be seen from Appendix 6-5 that there are differences in their choice of items in this subcategory. They are *table x* which has been used more by PRO writers (72 times in PRO texts and 5 times in the student texts), followed by *figure x*, used only twice in the student corpus but 16 times in the PRO texts.

In summary, a comparison of the frequency between the student and journal article writing shows no statistical significance. However there are some differences in the preferred devices in the subcategory of pre/reviews: while the PRO writers more frequently used features *table x* and *figure x* for providing guidance to the reader with tables or figures, the student writers more frequently employed features, *previously/earlier* and *above*, to alert the reader concerning what has been going on in the previous parts.

6.4.5 Additives

Additives include MD features such as *moreover*, *furthermore*, *in addition*, which writers use to add or reinforce information.

6.4.5.1 NES vs NNES

Among all the TMD subcategories, additives are ranked second and third in the NES and NNES corpus respectively. Additives are used 17.45 times per 5000 words in the NES texts and 19.01 in the NNES ones. The density is 20.2% in the NES texts and 19.5% in the NNES texts. The difference between the two groups is not significant (MWU test, $z = -1.761$, $p < 0.078$).

Regarding the usage of the features in this subcategory, *also* was used the most in both texts (219 instances for the NES and 296 for the NNES). Other features were also used in both types of text and can be presented in the following descending order of their occurrence: *therefore* 72, *furthermore* 16, *moreover* 10, *in addition* seven *in the same way*, and *accordingly* each of which was used six times, and *likewise* three in the NES texts, while *therefore* was used 111 times, *moreover* 29, *furthermore* 28, *accordingly* 18 times, and *likewise* once in the NNES texts.

There is a similarity in the frequency of this subcategory with the predominant use of *also* and *therefore* in the two corpora as in:

However, unless the native speakers have been trained, they will be unable to explain the grammatical rules whilst a proficient Non-native speaker or a L2 learner will be able to explain as they would have learnt these rules. *Also*, in vocabulary, some non-native speakers maybe far superior to some native speakers in the vocabulary of specialist registers for example. (Boyle, 1997) (NES 10).

Skimming and scanning are two essential reading techniques for quick and efficient reading. *Therefore*, teachers should provide their students with the necessary information about these two techniques and do sufficient practice (NNES25).

However, some of the features preferred by the NES writers were used infrequently in the NNES texts:

These are adjectives which forms part of a scale of values or qualities that can be compared to a norm or average. They can be used as intensifiers and comparatives. Eg: huge, very big, quite big, medium sized, quite small, small, tiny. *In the same way* scales for hot / cold, love / hate, good / bad, interesting, boring can be used (NES 10).

The writer used *in the same way* to add another example of ‘a scale of values’ to enhance the reader-friendliness of the text (6 times in the NES whereas no instance was found in the NNES corpus from the raw frequency of additives).

Likewise was used in the NES texts to extend the previous information in a neutral way instead of a more overtly reinforcement of the information (3 in the NES and 1 in the NNES);

This may lead to an increased sensitivity to save face amidst the more individualistic and less class/gender distinctive Israeli Jews. *Likewise*, the collectivist nature of Arabic culture is countered by individualist tendencies in Israel's majority and can be observed in the Jewish and Arab communication behaviours (NES 19).

Furthermore and *moreover* were more frequently used in the NNES texts to reinforce the information which had been stated (*furthermore* 16 times in the NES corpus and 28 in the NNES ones, and *moreover* 10 times for the NES and 29 for the NNES):

Therefore, techniques must be designed and sequenced in a way that is consistent with these steps. *Furthermore*, they must be designed to ensure that "repetition, retrieval, generation and thoughtful processing occur" (NNES 22).

Firstly they believed that some teachers never well prepared lessons before class. *Moreover*, in terms of some basic linguistic questions, such as grammar, students could not find a satisfactory answer from teachers (NNES 6)

In both extracts, the writers used the features *furthermore* and *moreover* to give support to the previous idea by providing a more important reason in the following statement.

A comparison of the use of additives between the NES and NNES writing shows no significance. However some features preferred by the NES writers were infrequently used in the NNES texts. For instance, *in the same way* and *likewise* were more frequently used by the NES writers while *furthermore* and *moreover* were employed more frequently in the NNES texts.

6.4.5.2 PRO vs Students

In the student texts, this subcategory is ranked second while the PRO writers used additives the third most among all the TMD subcategories. Tables 6-32 and 6-33 show

that additives are used 18.28 times per 5000 words in the student texts with density of 19.8%. In the journal article texts they are used 16.88 times per 5000 words with 17.4% density. Student writers tend to use more additives to explicitly add information in their academic writing. Despite the apparent high number of additives in the student texts, the difference is not significant (MWU test, $z = -1.324$, $p < 0.186$).

The student writers and their counterparts show similarities in their choice of features in this subcategory, with the three most frequent features in the two corpora (e.g. *also*, *therefore* and *in addition*). However, as can be seen from the list below and appendix 6-5-5 (raw frequency of additives), additives are far more common in the student texts with a relatively high frequency of some features (e.g. *therefore*, *moreover*, and *furthermore*) as mentioned in the previous section (comparison between the NES and the NNES) to reinforce the previous information in the following sentence, while the PRO writers employed relatively more items, *also*, *in addition*, and *in the same way* to extend the information as a neutral way.

To compare the frequency of features for additives in each group, their frequency of occurrence is listed as follows:

POSCORP: also (515), therefore (183), in addition (64), furthermore (44), moreover (39), accordingly (24), in the same way (6), likewise (4)

RACORP: also (542), therefore (85), in addition (80), furthermore (17), moreover (15), in the same way (8), likewise (5), accordingly (4)

The two sets of writings show similarities in the choice of features in this subcategory, with the three most frequent features in the two corpora, such as *also*, *therefore* and *in addition*, although the student writers use more additives than their counterparts.

6.4.6 Concessives

Concessives mark a change of direction in the running discourse or a modification of preceding propositions. Examples of concessives are *but*, *by contrast*, *however*, *in spite of*, *on the contrary*, *on the other hand*, *whereas* and etc.

6.4.6.1 NES vs NNES

Among all the TMD subcategories, concessives are ranked first and second in the NES and NNES texts respectively. Concerning the total number of occurrences, they are used 26.10 times in the NES texts and 21.45 times in the NNES ones per 5000 words.

The density of concessives is 30.2% in the NES texts and 22.1 % in the NNES ones.

The above figures indicate that the NES writers make more use of concessives than their counterparts (21% more). However, the result of the statistical test is $z = -0.356$, $p < 0.722$, showing no indication of any significance.

Out of the features used in this subcategory, the most widely used concessive is *but* which scored 261 and 295 times in the NES and in the NNES texts respectively. It is followed by *however* (163 in the NES and 183 times in the NNES texts) and *although* which was used 68 times in the NES and 63 times in the NNES. These are the three most frequent features used in both group texts and can be seen as a similarity of their choice of features in this subcategory. However, some of the phrases were used more frequently by the NNES writers but infrequently in the NES texts (e.g. *on the other hand* and *whereas*):

Reading short stories, for instance, has been proven to be effective because these attract the reader and help them to interact with the theme of the story till the end. *On the other hand*, extensive reading plays a fundamental role in improving reading skills in particular and language learning in general (NNES 25).

The writer used *on the other hand* to modify the propositional information from the importance of ‘reading short stories’ to ‘extensive reading’.

The first paper, “Differences in the metacognitive awareness of reading strategies among native and non-native readers”, is quantitative in nature, *whereas* the second paper, “Reading and the non-academic learner: a mystery solved”, is apparently a piece of qualitative research. In this critique, the comparison and contrast of the two papers to highlight some prominent features in terms of similarities as well as differences of the two research articles will be undertaken as follows (NNES 1).

The same effect is achieved in the extract above, in which the writer tries to change the topic by using the concessive device, *whereas*, from ‘quantitative method’ to ‘qualitative method’ in a running discourse.

As in NES texts, the concessive feature *yet* was used to achieve these purposes:

Through using repetition, the host is looking for clarification of understanding. *Yet*, the facial expression of the contestant is defiant as she does not realise why her answer causes amusement (NES 13).

It can be seen from the frequency of raw occurrence (see appendix 6-5) that this feature was used more frequently in the NES texts (39 times) than the NNES texts (14 times).

The three most frequent devices features chosen in this subcategory in the NES and NNES writing are: *but*, *however*, *although*. However, some of the phrases that were used more frequently by the NNES writers (*on the other hand* and *whereas*) were used infrequently in the NES texts, showing how the groups differ in their use of concessives.

6.4.6.2 PRO vs Students

Among all the TMD subcategories, concessives are ranked first in both group texts. Tables 6-32 and 6-33 show that in the student texts concessives are used 23.63 times per 5000 words with density of 25.6%. In the journal article texts they are used 28.57 times per 5000 words with 29.4% density. So the journal article writers make more use of concessives in their texts (20% more) although the result of the MWU test ($z = -1.344$, $p < 0.179$) shows that the difference is not statistically significant.

The student writers and their counterparts show similarities in their choice of features in this subcategory, with the three most frequent items (*although*, *but*, *however*) as in:

Since Lenneberg's (1967) revolutionary work, it has been assumed that the earlier the brain injury is sustained, the greater the chance of full recovery from aphasic symptoms, with most adults being left with long-term traces. *However*, I will argue that things are not so clear-cut and that, with evidence from more up to date research than Lenneberg's, such a simple proposal is not appropriate for the complexity for what we now know about the human brain (NES 2).

The contrastive *however* was employed to show the reader what is surely the main focus of the writer's argument. We are told that the assumption concerning brain injury is vague and that in fact the researcher's work is noteworthy.

Another example of the PRO writer using *although* to modify the preceding ideas (bottom-up and top-down skills are important in the learning of reading):

According to Birch, *although* both bottom-up and top-down processing skills are necessary to learn to read in an L2, the reading fundamentals must be in place before top-down instruction can benefit learners (PRO10).

Like the previous extract, the main focus of the author's argument is 'the reading fundamentals must be in place before the learning skills of the reading'.

However, the journal article writers show that there are group preferences in the choice of features in this subcategory, such as *whereas* and *on the other hand* (*whereas*: 83 occurrences in the PRO and 28 in the student texts, *on the other hand*: 39 in the PRO and 29 in the student corpus), as in:

...., there might be an effect for word order: for instance, Spanish speakers may pattern with Chinese speakers (mainly SVO), *whereas* Japanese speakers (SOV) might pattern differently. This latter prediction is based on Fender (2003), who reports that Japanese-speaking learners have particular problems with integration when reading simple sentences when compared to Arabic speaking learners (PRO 7).

As we have noted, the earlier chain-building and weighted network experiments found few systematic differences in either path length or path strength for learners and native speakers. The alternative methodology reported here, *on the other hand*, appeared to suggest that the number of connections per word in both L1 and L2 networks was far greater than other work in the field had led us to expect (PRO 13).

Whereas and *on the other hand* were used to mark a change of direction in the running discourse by the journal article writers.

Although there is no statistical difference in the comparison of the frequency between the student and journal article writing, it is noticeable that the PRO writers employed concessives more frequently. However both groups of writers show similarities in their choice of features in this subcategory, with regard to the three most frequent items (*although, but, however*).

6.4.7 Code Glosses

Code glosses are used to guide the reader towards the definition, meaning of terms, concepts or to elaborate or illustrate propositions. Examples of code glosses are *defined as, e.g., for example, for instance, in other words*.

6.4.7.1 NES vs NNES

The subcategory of code glosses appeared 12.76 times in the NES texts and 17.16 in the NNES ones, with a density of 14.8% and 17.6% respectively. The rank order is 3 in the NES texts and 4 in the NNES ones. The use of code glosses differs between the groups: the NNES used more code glosses than their NES counterparts (18% more) with highly significant differences (MWU test, $z = -2.940$, $p < 0.003$).

Regarding the usage of the features in this subcategory, the NNES writers made great use of most of the features in general and some of the devices in particular (*e.g. for example, for instance, i.e.*) as in:

..... definite article and location adverbs (*e.g. here, now, there, then*) and Comparative reference including adjectives in comparative degree and certain adjectives and adverbs (*e.g. same, identical, other, different*). (NNES 14)

In the extract, the teacher gives the learners feedback every time they answer the question. A follow-up move is made up of either a single act or a combination of two or more acts. *For example*, a single evaluation act, “That’s right” (Turn 3), constitutes one follow-up move, and the follow-up in Turn 7 is made up of two acts (“Very good” = evaluation; “Ta ye hen mang” = repetition of the learner’s reply = accept) (NNES 17).

Decision must be made after consulting the superior officers, and especially in the strict hierarchy divided culture, one can only contact with his immediate superior, therefore, it take too long time for a report to be reached to the top ranked officer. While in a low power distance culture, freedom is under control of the individual negotiators, and quick decisions could be made without consulting the top boss. *For instance*, in America, it is quite common for contract to be signed during the first business meeting. (NNES 11)

Additionally, the author failed to give details of how the interviews were undertaken, *i.e.* what kind of interview was applied, *e. g.* structured, semi-structured or unstructured interviews, and what methods were used to record the interview data. (NNES 1)

In the above extracts, the writers use code glosses to explain what the previous propositions mean by providing some examples to the reader.

As in the NES texts, the code gloss phrases *defined as* and *for example* can be used to clarify the term, as in ‘acquired aphasia’ in the NES text:

....Acquired Aphasia is *defined as* impairment of language abilities following acquired brain injury (Bates and Roe, 2001:2), *for example* as the result of a stroke or a car accident. The symptoms of childhood and adult acquired aphasia are.... (NES 2)

So, code glosses in both NES and NNES texts were used to explain definitional or propositional terms.

The similarities between the groups mainly concern explanation of definitional terms. However there are differences in the frequency of the features; the NNES made more use of most of the features in general and some of the devices in particular (*e.g. for example, for instance, i.e.*) with significant difference (MWU test, $z = -2.940$, $p < 0.003$).

6.4.7.2 PRO vs Students

In the student texts, this subcategory is ranked fourth while the PRO writers used code glosses the second-most among all the TMD subcategories. Tables 6-32 and 6-33 show that in the student texts code glosses are used 15.22 times per 5000 words with density of 16.5%. In the journal article texts they are used 19.76 times per 5000 words with 20.8%. Thus the code glosses were more frequently used by the journal article writers (29% more). Despite the apparent difference in the mean frequency scores between the two group texts, the difference is not significant (MWU test, $z = -1.638$, $p < 0.101$).

The student writers and their published counterparts show similarities in their choice of features in this subcategory, with the three most frequent items in the two corpora (*e.g., for example, such as*). These are the three most frequent features used in both group texts and can be seen as an indication of the similarity of their choice of features in this subcategory. However, differences in their choices can be found from Appendix 6-5;

some of the features were used more frequently by the journal article writers (*e.g., for example, i.e., I mean*) as in:

Such fundamental factors as who given L2 learners are, why and where these individuals undertake to learn an L2 and what their available resources are (*e.g., time, cognitive, financial*) should and often do determine how particular L2 skills are taught and learned (*e.g., Breen, 2001; Breen & Littlejohn, 2000*). (PRO 10)

Therefore, we need to create mechanisms through which we can help students communicate with us when we are not face-to-face, and specifically through which students can inform us about what they are trying to accomplish. *For example*, we can create assignments that allow students to let us know their intentions. Students can complete cover sheets and attach them to their papers when they turn them in or directly annotate their texts (Leki, 1990). (PRO 12)

Street and his colleagues (*e.g., Street, 1982, 1984; Street & Brady, 1982; Street, Brady, & Putman, 1983*) found similar results when investigating the effects of speech rate and response latency on perceived competence (*i.e., intelligence, status, expertise*)-another dimension related to self-image and abilities. (PRO 18)

....because there are some so-called NNSs who are far more knowledgeable. And I don't just mean grammatical knowledge. *I mean* awareness of cross-cultural pragmatics and all kinds of other things that NSs are just not aware of (PRO 17).

In the extract above (PRO 17), the writer tries to clarify what he has in mind when he uses the term, 'knowledge' by using the feature of code glosses, *I mean*, which was more frequently used by the PRO writers to help to clear up anticipated queries regarding the propositional content ('NNSs who are far more knowledgeable').

Other types of code glosses (*defined as, in other words, known as*) are similarly concerned with things for the reader, and were more frequently employed by the student writers:

Aphasia is *defined as* impairment of language abilities following acquired brain injury (Bates and Roe, 2001:2), for example as the result of a stroke or a car accident. (NES 2)

....students should become conscious of generalizations, stereotypes and prejudices, and become increasingly open and positive towards different cultures. *In other words*, they hope that this awareness and understanding will result in the respect, acceptance and tolerance of cultural differences (NES 25).

And Jandt (2004) introduces that ‘many academic disciplines refer to the Sapir-Whorf hypothesis (also *known as* the Whorfian thesis) when accounting for the differences in languages across cultures. (NNES 23)

In general, there are differences between the two groups; the journal article writers employ code glosses more frequently in the comparison of quantity (15.22 per 5000 words in the POSCORP and 19.76 per 5000 words in the RACORP). While the student writers employ code glosses for expounding their propositional content to enhance the reader’s understanding, as in the extracts above (from the NES2, NES25 and NNES 23), the PRO writers use them for the same reason as the students do and also for answering anticipated queries regarding their propositional content, as in the extract from the PRO 17 (*I mean*).

6.4.8 Rationales

The features included under this subcategory refer to markers of cause and effect relations, justifications, results e.g. *the reason for this, as a result of, because, as a consequence, etc.*

6.4.8.1 NES vs NNES

Tables 6-30 and 6-31 show that in the NES texts rationales occurred 5.98 times compared with 7.11 times in the NNES ones. The density is 6.9% in the NES and 7.3% in the NNES texts. So the NNES employed more rationales than their NES counterparts although this is not significant (MWU test, $z = -0.998$, $p < 0.318$).

Only three types of rationales have been used repeatedly in relatively large numbers in the two corpora. They are *as a result of, because* and *as a consequence/consequently* which have been used 14, 107 and 18 times in the NES texts and 30, 171 and 13 times in the NNES texts respectively. They are followed by *thereby* twice in the NES

whereas no instance of this item was found in the NNES texts and the *reason for this* occurred once in each corpus. In both texts, 'because' is the most frequent feature in this subcategory as in:

Israel's legal system also reflects the essential role of Judaism, *because* religious tribunal courts hold authority over citizens in certain issues such as marriage. Partially similar to the Western judicial system, Israeli courts have three levels: Magistrate, Appellate, and Supreme (NES 19).

According to questionnaires, nearly 90% of Chinese students choose red colour as their preference of gift giving. Though red is a fearful colour which means danger in western countries, red is the most preferable colour of gift giving in China *because* Chinese people believe that red is the symbol of good luck. (NNES 15)

The rational feature, *because* helps the writer spell out why 'Israel's legal system reflects the essential role of Judaism' in the extract above. This extract again demonstrates the multifunctionality of MDs, in that the fact that the writer feels compelled to explain the reasons which also can constitute a hedge (which will be discussed later) to avoid criticism by using the feature of rationales (*because*). But this item was counted as rationales since the primary function of this feature appears to be a rational item.

The feature, *as a result of* is relatively frequent in the NNES corpus. The NNES writers used this phrase 30 times while this was found 14 times in the NES texts.

As a result of modern demands, in-class methods and rates of emphasis on language skills are changing. Schools are coming to focus on the speaking and listening skills more and a listening component has been introduced into the University Entrance Examination (NNES 12).

Sifleet claims that sometimes an intermediary should be involved in a negotiation *as a result of* cultural or style differences. Effective negotiation needs to be peer-to-peer. Involving an intermediary can help to bridge the differences and find the common ground or 'win-win' outcome. Relationship in a collectivistic oriented culture is so important that with a contact or intermediary, one can make an appointment and quickly establish a harmonious atmosphere (NNES 11).

The writers justify the reasons for their statements by using the phrase, *as a result of*.

The NNES employed rationales more often than their NES counterparts although they share the three most frequent features in the use of this subcategory (*as a result of*, *because* and *as a consequence/consequently*).

6.4.8.2 PRO vs Students

The subcategory of rationales appeared 6.58 times per 5000 words in the student texts and 7.50 in the PRO ones, with a density of 7.1% and 7.6% respectively. The rank order is 5 in both group texts. There are differences in the use of rationales: the PRO used more rationales than their student counterparts (9% more). However, the difference between the two groups is not significant. (MWU test, $z = -0.285$, $p < 0.776$).

Concerning the features used in this subcategory, *because* is the most frequent feature in both group texts. However, differences in their choices can be found from Appendix 6-5; while the student writers employed the features, *the reason for this*, *as a result of*, *as a consequence* more often, the journal article writers made more use of *because* and *thereby*. In fact, the PRO writers used *thereby* 19 times while the NES writers employed it only twice and no instance of this item was found in the NNES sub-corpus:

The effect of invoking some executive process, then, is to augment the activation level of the coalitions to which it is applied, *thereby* increasing the likelihood that those action features will become consciously available (PRO 26)

Writing allows learners access to observable units of text and *thereby* induces attention to form (PRO 23)

The writers in the above extracts used *thereby* which marks the cause and effect relationship here.

Similarly, as in the student corpus, the rational phrase *as a consequence* was used to achieve these purposes:

She believed that this accusation stemmed from the students' lack of confidence in their own language levels. *As a consequence*, in subsequent lessons she began to use more L1 in her lesson in order to encourage her students (NES 15)

These features (*as a consequence/consequently*) were more frequently used in the student texts (31 times) than in the journal article texts (13 times).

The PRO used more rationales than their student counterparts and there are some features (*because* and *thereby*) more frequently used by the journal article writers whereas they have been used infrequently in the POSCORP. However there are similarities in the purposes of the use of rationales and they share the most frequent feature used, *because*, in both group texts.

6.4.9 Sources

Sources refer to the TMD features that signal the origin of ideas, concepts or content in general. Whereas other studies have used 'attributors' and 'narrators' as two separate subcategories (Williams 1981; Vande Kopple 1985), here both of them are put under the heading of 'sources'.

6.4.9.1 NES vs NNES

Within the nine TMD subcategories 'sources' is rank ordered 5th in the NES texts and 6th in the NNES ones (see table 6-31). Its occurrences are 7.41 and 5.26 per 5000 words in the NES and NNES texts with 8.6% and 5.4% respectively. The NES writers use more sources than their NNES counterparts (40% more). Despite the apparent difference between the two group texts, the MWU test shows that the difference is not significant ($z = 0.043$, $p < 0.835$) because of the uneven distribution of the data (see

Shapiro-Wilk tests in Appendix 5-10).

Writers in both groups show similarities in their choice of features in this subcategory, with the two most frequent features in the two corpora (e.g. *according to x*, *x argues*).

However, as can be seen from the raw frequency of sources (appendix 6-5-9), there are group preferences in the choice of features: in the NNES student texts *according to x* was more frequently used compared to the NES texts while the items, *x suggests*, *x argues*, *x notes* and *x states* were employed relatively more by the NES writers.

To compare the frequency, the frequency of occurrence of features in the subcategory in each group is as follows:

NES: *according to x* (59), *x argue* (33), *x suggest* (21), *x state* (17), *x propose* (15),
x note (10)

NNES: *according to x* (94), *x argue(s)* (23), *x propose(s)* (17), *x suggest(s)* (13),
x state(s) (4), *x note(s)* (1)

(listed in descending order: spelling variants disregarded)

Although there are differences in frequency, both the NES and the NNES writers employ them for similar purposes, to show the origin of ideas, concepts or content. The next extracts are typical:

According to Gestalt theory in psychology, the sum is more than the combination of its parts. Using this idea, perception can be divided into two parts: the figure and the ground. This idea, coupled with the realisation that syntax cannot cover all aspects of meaning, was related to sentence organisation, the figure being the emphasised part, and the ground being the given part of the sentence. JOHN is flying to Paris tomorrow. (Who is flying to Paris tomorrow?) (NES 24)

According to Sinclair and Coulthard (1992, p.25), the two moves (framing and focusing) often occur together within a boundary exchange, but not always. Here is an example from my data of a focusing move that occurred within a boundary exchange (NNES 17).

The extracts above help to connect the writers' work with earlier works which have already been accepted in the community.

In general, the NES writers use more 'sources' than their NNES counterparts although there are similarities in the purposes of the usage (i.e. showing the origin of ideas).

6.4.9.2 PRO vs Students

The subcategory of sources appeared 6.27 times per 5000 words in the student texts and 7.04 in the PRO ones, with a density of 6.8% and 7.2% respectively. The rank order is 6 in both group texts (see table 6-33). The journal article writers use more sources than their student counterparts (7% more) with no significant difference (MWU test; $z = -0.140$, $p < 0.889$).

There are nevertheless differences in the preferred choice of the features. For example, the journal article writers employed *x suggests* more than their student counterparts while the student writers used *according to x* more frequently than their PRO counterparts.

Examples are given in the extracts below:

According to Krashen (1982), for language acquisition to occur there has to be a considerably large input of authentic language which is presented in context and at a level which is linguistically challenging to the learner. This should be carried out in an atmosphere with low anxiety levels to enable students to achieve language acquisition. The use of video-taped material in the language classroom can fulfil these prerequisites with careful planning (NES 13).

According to Markus & Kitayama (1991, cited by Kumar, 1999, 65), the characteristic of collectivist societies is the "interdependent self construal". Truly, this is a reality for China, given the collectivistic character of the negotiation team and process. Group orientation describes the management style in China (NES 16)

In the student corpus, they use the features of sources (*according to x*) to link their work with established work and agree with it without any evaluation. In the extracts above, the writers used *according to x* to show that their arguments are backed up by the established work done by other researchers.

In the following extracts, however, it is demonstrated how the writers from the PRO corpus can show not only other researchers' work as an origin of their work, but also make it better suit their own contexts:

The findings for a positive relationship between phonetic sensitivity, memory ability and learning from recasts in Robinson and Yamaguchi study, and phonological working memory capacity and noticing of recast information in Mackey et al. (2002) *suggest* that these abilities are positively implicated in aptitude for learning from the recasting technique. *However*, as with the finding for incidental learning in Robinson (1997a), reported earlier, in Robinson and Yamaguchi (1999), there were nonsignificant correlations of learning of relative clause during task-based interaction (supplemented by targeted recasts) and the grammatical sensitivity aptitude subtest. These findings therefore allow an inference across contexts (laboratory studies of incidental learning, and classroom studies of focus on form during task-based learning) about the noninfluence of individual differences in grammatical sensitivity on incidental learning during processing for meaning (PRO 14).

The OLP groups produced more accurate output in terms of both speaking and writing. In addition, the analyses of the spoken discourse revealed that the online planners spoke more slowly, and used more self-repairs suggesting a greater level of monitoring of their language. Yuan and Ellis (2003) *suggest* that the availability of online planning time allows learners to access their explicit knowledge about grammar, *whereas* in pretask planning learners tend to advantage content over form. Results from questionnaire and interview data for the writing study supported this view with participants reporting a greater focus on language (i.e., accuracy) with online planning (i.e., during the execution of the task) for both the PTP and OLP groups, but less for the NP(no-planning) group who presumably focus more on content planning during the limited time they have to write (PRO 23).

The above extracts from the RACORP achieve a similar effect to what the students' writers do, linking the writer's work with established work by using the feature of sources (*suggest*); however, other MD devices are used in the PRO texts to evaluate the earlier literature which the writer is building upon by using concessive devices

(*however* and *whereas*). In other words, employing the devices of sources (*suggest*) shows that the writer's work is from earlier research or from other researchers' ideas and methods; however using concessive devices (*however* and *whereas*) not only shows the writers offering up their own conclusion or methodology as an alternative to other researchers' results or methodologies but also signals that they do not follow the previous work carried out by the other researchers and make it better to suit their own contexts.

The frequency of features in each group is as follows:

POSCORP: according to x (153), x argue(s) (56), x suggest(s) (34), propose(s) (32), x state(s) (21), x note(s) (11)

RACORP: x suggest(s) (104), x argue(s) (87), according to x (73), x propose(s) (37), x note(s) (12), x state(s) (6),
(listed in descending order: spelling variants disregarded)

A comparison of the student writing with the PRO writing shows that the two sets are different in terms of frequency and purpose of the use of sources. For the difference in frequency, the journal article writers used more sources than their student counterparts. In addition to this, there are differences in the purposes of their preferred choice of the devices. For instance, *x suggested* was employed more frequently by the journal article writers to signal that they will evaluate the previous work while *according to x* was more frequently used in the POSCORP to connect their work with established work rather than evaluating the previous work.

6.5 Results from Individual Subcategories in IMD

The IMD results for the different corpora are presented in below. The figures given are the mean frequencies and percentages of each IMD sub-category out of total IMDs per

5000 words in the NES and NNES corpora, and the student and PRO corpora respectively.

Table 6- 34: Mean Frequencies of IMD Subcategories per 5000 words with the Indication of the Statistical Difference in the NES and NNES texts

IMD subcategories	NES	NNES	Statistical Difference
Hedges	9.32	5.34	✓ NES>NNES (p<0.007)
Emphatics	12.76	13.25	Non sig.
Evaluatives	3.87	3.07	Non sig.
Commentaries	1.58	1.89	Non sig.
Appeals	0.53	0.38	Non sig.
Self references	22.66	14.51	✓ NES>NNES (p<0.034)
TOTAL IMD	50.72	38.45	Non sig.

Table 6- 35: Percentages of IMD Subcategories out of total IMDs in the NES and NNES texts

Rank Order	NES	NNES
1	Self references (44.7)	Self references (37.7)
2	Emphatics (25.2)	Emphatics (34.5)
3	Hedges (18.4)	Hedges (13.9)
4	Evaluatives (7.6)	Evaluatives (8.0)
5	Commentaries (3.1)	Commentaries (4.9)
6	Appeals (1.0)	Appeals (1.0)
Total IMD	100%	100%

(Percentages are given in brackets: out of total IMD)

Table 6- 36: Mean Frequencies of IMD Subcategories per 5000 words with the Indication of the Statistical Difference in the Student and PRO texts

IMD subcategories	Student	PRO	Statistical Difference
Hedges	7.21	10.50	✓ Stu<PRO (p<0.009)
Emphatics	13.02	11.92	Non sig.
Evaluatives	3.45	2.80	Non sig.
Commentaries	1.75	1.10	Non sig.
Appeals	0.45	0.46	Non sig.
Self references	18.33	20.87	Non sig.
TOTAL IMD	44.19	47.66	Non sig.

Table 6- 37: Percentages of IMD Subcategories out of total IMDs in the Student and PRO texts

Rank Order	Student	PRO
1	Self references (41.5)	Self references (43.8)
2	Emphatics (29.5)	Emphatics (25.0)
3	Hedges (16.3)	Hedges (22.0)
4	Evaluatives (7.8)	Evaluatives (5.9)
5	Commentaries (3.9)	Commentaries (2.3)
6	Appeals (1.0)	Appeals (1.0)
Total IMD	100%	100%

(Percentages are given in brackets: out of total IMD)

The NES thus make more use of IMD than their NNES counterparts (see Table 6-34) and Table 6-36 shows that the PRO writers use more IMD than their student counterparts.

The results of individual subcategories in the following sections will shed more light on the issue of differences in the purpose and preference of IMD use between groups.

6.5.1 Hedges

Writers use hedges to express uncertainty or doubt towards the content. They can be said to reflect a mild or soft attitude as opposed to the strong one expressed by ‘emphatics’ which will be discussed in the next section.

6.5.1.1 NES vs NNES

Tables 6-34 and 6-35 show that in the NES texts, the mean number of hedges is 9.32 per 5000 words, that is 18.4% density, whereas in the NNES ones they are 5.34 and 13.9%. This means that hedges were more used in the NES texts by 74%. The apparent difference is statistically significant from the mean rank, MWU test ($z = -2.706$, $p < 0.007$). As far as rank order is concerned, this subcategory was rank-ordered third among all the IMD subcategories in both types of text.

Although the NES writers employed more hedges overall, there are similarities with the three most frequent features in the two corpora (e.g. *might*, *seems*, *tend to*) in their choice of features in this subcategory. However, as can be seen from the raw frequency of hedges in Appendix 6-6, most of the features were more frequently used by the NES writers while only the feature, *might* was more frequently employed in the NNES texts, as in:

This activity *might* help some students who have no idea what the teacher or the next activity run.....(NNES 27)

Perhaps was used very often in the NES texts whereas the NNES writers used it seldom (41 times in the NES texts and 3 times in the NNES texts)

Perhaps one of the more obvious advantages of the visual channel, then, is that it communicates the shapes, movements or relative positions of objects rather than the audio channel or the written word (NES 18).

Both the writers of the above extracts used hedges (*might* and *perhaps*) to convey their own opinions, which they cannot assume will be accepted by everyone; therefore they try to acknowledge their uncertainty about their readers' agreement.

Small cuts were made in an area of the monkey's brain that controlled movement and, compared to the more mature monkeys, Kennard found the younger monkeys were able to walk again sooner (Savage, D, and Urbanczyk, 2002). The

problem with this ‘Kennard principle’ is that it is in danger of oversimplifying the phenomenon of brain plasticity. Whilst in general it is *probably* safe to assume that the earlier the brain damage is sustained, the better the chance of recovery, the human brain is intricately complex and such a simple proposal does not suffice for the misfit cases, for example, those showing collateral sprouting in mature brains (Lynch and Gall, 1979). In a study of brain-damaged children, Teuber and Rudel (1962) showed that there is a complex interaction between age at injury and function (NES 2).

At times, the writer feels that the previous study is limited and is open to criticism. In the extract above, the hedge device *probably* is used to express doubt towards the others’ study.

There are statistical differences between the NES and NNES texts in terms of frequency. The NES make more use of hedges with highly significant difference (MWU test $z = -2.706$, $p < 0.007$) although both groups use the features of hedges with similar purposes; expressing their uncertainty towards the content.

6.5.1.2 PRO vs Students

Hedges were used 7.21 times per 5000 words in the student corpus, with 10.50 occurrences in the PRO texts. They are ranked third in both of them. The density in the student texts is 16.3% and in the PRO it is 22.0%. The hedges were more frequently used in the journal article texts (45 % more). The difference was highly significant (MWU test, $z = -2.605$, $p < 0.009$).

The student writers and their counterparts show similarities in their choice of features in this subcategory, with the four most frequent features in the two corpora being *might*, *perhaps*, *seems* and *tend to*. However, as can be seen from the list below, hedges are far more common in the journal articles.

The raw frequencies in this subcategory in the two corpora are given below:

POSCORP: might (95), seems (57), tend to (46), perhaps (44), probably (19), possibly (18), a assumed (14), approximately (10), broadly (6), presumably (6), unclear (3), uncertain (2),

RACORP: might (181), perhaps (72), seems (52), tend to (39), assumed (28), probably (27), approximately (15), broadly (14), possibly (8), unclear (8), presumably (6), uncertain (3)

RACORP writers used hedging features as protection from anticipated attacks, mainly in connection with their methodological choices and to justify the limitations of their research:

At first sight it *might seem* that these concrete measurements of L1 and L2 network density are somewhat difficult to interpret since they are in some sense decontextualized measures (PRO 13).

Participants were presented with a set of vocalized adjectives followed by one of delay times, and then a true/false response that *presumably* indicated whether this adjective applied to the speaker (PRO 18).

In addition to this, when the journal article writers fear the readership may disagree with the writer's interpretation of results, hedges were used to protect the writer from attack:

These analyses of real-life language in use have delved into, for example, the frequencies and patterns of syntactic, morphological, lexical, pragmatic, or discoursal features that *tend to* occur in particular types of text (PRO 10)

Not surprisingly, students *seemed to* write more when their writing is not being marked for any kind of surface error than when it is (PRO 21)

This is because the violation of social communication rules could *possibly* lead to a loss of face. For example, in strong UA cultures, people pay particular attention to fashion in order to feel protected when facing a threatening world (PRO 30)

However, hedges were used in the POSCORP when the student writer evaluated earlier research, and created a research space because she does not support the previous claim from her experience. In this case, the writer also hedges her disagreement with a hedging device (*as far as to say that*):

Quoted in Cook (2001), Sridhar (1986) goes *as far as* to say that ‘the goal of SLA is bilingualism’. However, this assumption is not universal in some EFL countries. I had the unique personal experience in which despite what the theoretical and empirical SLA findings show, the principal of the school I was teaching in requested that I hide my bilingual identity (NES 10).

Although the student writers and their counterparts show similarities in their choice of features in this subcategory with the four most frequent features (e.g. *might*, *perhaps*, *seems* and *tend to*), as can be confirmed by the statistical analysis, hedges are used significantly more (MWU test, $z = -2.605$, $p < 0.009$) in the journal articles. The purpose of using the hedges in the student texts, as can be observed from the extracts in the comparison between the NES and NNES in the previous section, is to acknowledge uncertainty about the readers’ agreement while they are employed in the RACORP not only to acknowledge uncertainty like the students’, but to protect them from readers’ attack in their methodological choices and interpretation of results.

6.5.2 Emphatics

Emphatics are a type of IMD that writers use to express certainty and indicate what they really believe to be true (e.g. *certainly*, *undoubtedly*, *surely*). Besides, they reflect a sense of strength of their claim.

6.5.2.1 NES vs NNES

In the NES texts the number of occurrences of emphatics is 12.76, while in the NNES ones it is 13.25 per 5000 words, and the density is 25.2% and 34.5% respectively. This shows that emphatics in the NNES texts are relatively frequent (3% more). This subcategory is rank ordered second in both groups’ texts.

With respect to the frequency of the features in this category, *clearly*, *always*, and *especially* are the three most frequent features in both NES and NNES texts, as in:

A deputy head with whom I worked recently and with whom I discussed a bullying problem in my year group, told me that bullying was “inevitable” and that school bullying can only be “managed” and “contained.” *Clearly* this is not the case. Definitions and Causes: Definitions of bullying range from “having no one to talk to” (Smith and Shu, 2000 p.194), to “aggressive and anti-social behaviour” (Salmon et al, 2000, p.566), to a cause of “post-traumatic stress” (NES 3)

It is *always* useful to approach an empirical study, such as this, in a cynical manner, remembering to question the writers’ approach to hypotheses, methodology and ultimately drawing anthropological conclusions from the results (NES 9)

Especially to those senior teachers, they have to spend some time to adapt this new concept (NNES 20).

The writers in the above extracts used emphatic devices (*clearly, always, especially*) in order to attract maximum attention to the writer’s opinion. This apparent desire to express a dramatic impact can reflect a sense of strength of their claim in the use of emphatics.

At times, the features of emphatics in both student texts were used to build on what had already been covered, in order to summarise the writer’s argument or findings before moving on:

As already mentioned, a similar approach could have been taken with regard to the second text, with the students putting themselves in the position of the parents writing the problems or of their children. Conclusion An issue that appeared to permeate this section of the lesson was the problem of teaching a fairly receptive skill through the extremely active process of developing cultural awareness (NNES 6).

The use of specific anaphoric references, as *already* mentioned, is more commonplace; “he” is quite a popular choice; the term is used in lines 7, 19,28,35 (NES 7)

Although there are similarities in their choice of features in this subcategory, with the three most frequent items in the two corpora, the NNES students used more emphatics than their NES counterparts.

6.5.2.2 PRO vs Students

Tables 6-36 and 6-37 show that the student writers employed emphatics 13.02 times per 5000 words with a density of 29.5%. In the PRO texts, they were used 11.92 times per 5000 words at 25.0% density. The emphatics were more employed (9% more) by the student writers.

As far as features of emphatics are concerned, both student and PRO writers made more frequent use of emphatic devices, such as *always*, *clearly*, *especially*, *in fact*, *really*, *simply* to express their certainty about their statement as in:

It was *always* the subject matter that counted, not personal relationships with individual faculty. (PRO30).

Clearly no real experimental study is going to be dealing with entirely homogeneous groups of participants with lexicons of uniform density. We may therefore expect that the real data will show a greater degree of variation within participant groups (PRO 13).

Many theorists have criticised this aspect of Brown and Levinson's model of politeness, focusing on both the overextension and the limitation of use of the term 'face' in Brown and Levinson's use. It seems that Brown and Levinson's model is almost unable to analyse politeness beyond the level of the sentence. This is something that seems especially pertinent in this study, *especially* when the reader might consider the idiom that "actions speak louder than words", indicating that holding a door open for someone might more accurately represent tacit civility than spoken forms of politeness. The reader will, however, notice some shortcomings in Moser and Corroyer's methodology (NES 9)

The emphatic *especially* was applied by the NES student to draw attention to the example of the difficulties in 'Brown and Levinson's model'. The use of emphatics helped the writer to emphasize the idea that there is a limitation of the use of the term 'face'.

Another passage by the PRO writer which uses *especially* for a similar purpose, to emphasize an example of the previous sentence, is:

At a more fundamental level, a significant proportion of teachers, *especially* in the rural areas, do not have sufficient command of the English language to conduct their classes with confidence (PRO 28)

Emphatics in the journal article texts at times combine with the first person pronoun to construct the image of a conscientious researcher. Note the use of *I* and *always* in the following extracts: *always* allows the writer to emphasise the limitation of methodology and shows that he/she is concerned about this. Thus the author is trying to minimise the limitation by starting with a definition of humour first:

.....in this study recorded their conversations with a wide variety of NSs, from family members to strangers, and thus were not always certain to encounter a NS willing to collaborate with them. Thus, the examples I present of humour are not *always* co-constructed, or even apparently humorous to both parties. Because of this, I begin with a definition of humour/language play (PRO 1).

The student corpus, on the other hand, did not have a single case of emphatic features, which shows that the writers are meticulous about their methodology.

The limitations of the author's study in the passages below are partly conveyed through the use of emphatics (*clearly*) which are used to demonstrate that the writers hold that there are some of the anticipated criticisms of their work in the RACORP:

Clearly there are potential problems with this sort of endeavour... (PRO 15)

These suggestions show that the notion of ways lead us in the right direction and should alert us to the dangers of accepting too superficial an analysis of the popular metaphor of the lexical network rather than opting for a more formal approach. *Clearly*, the model we have used is a very simple (PRO 13)

However, in the POSCORP, *clearly* is used to help to emphasize the propositional contents as in:

This statement *clearly* reflects the view that the L1, in this case English, should not be used in the class (NES 15)

The author described *clearly* how the study was undertaken particularly how the teaching of the experimental groups was rendered... (NNES 1)

A comparison of the student writing with the expert writing shows that the student writers employed more emphatics overall to express certainty and create a dramatic

impact about what writers believe to be true, apart from a few features; *in fact* (68 in RACORP and 43 in POSCORP), *certainly* (33 in RACORP and 15 in POSCORP), *simply* (67 in RACORP and 46 in POSCORP), *of course* (40 in RACORP and 25 in POSCORP) which were used more frequently by the journal article writers:

Those portions of texts for which differences could not be resolved were not used in the subsequent analysis. *In fact*, these represented a very small portion of the data (PRO 13)

Whether such larger batteries add any incremental validity to the predictive power of earlier, more parsimonious tests is an empirical issue. *Certainly*, if early and later stages of language learning draw on different abilities, or combinations of abilities, then larger batteries will be necessary to inform selection, and also diagnosis and pedagogic use of aptitude profiles (PRO 14).

What are the social and educational consequences that their instructional practices may have for the intellectual and social lives of the students they teach? Moreover, *simply* legitimizing teachers' ways of knowing will not automatically lead to praxis (PRO 20).

They reported very few difficulties in understanding the words used in the questionnaire. *Of course*, it is possible that a few participants were either under-reporting the gaps in their knowledge or had misunderstood some vocabulary, but this seems unlikely (PRO 13)

They were used to extend the propositional content with emphasis in the RACORP.

Although the student writers employ more emphatics than their PRO counterparts they mainly use them to emphasise the propositional content with a limited purpose (to emphasise the propositional contents in order to create a dramatic impact) whereas the PRO writers employ them for variety of purposes, such as emphasis of the previous sentences, demonstrating that the writers are concerned about the limitations of the methodology and their anticipated criticism of their work.

6.5.3 Evaluatives

Evaluatives are a feature of IMD that refer to the writer's attitude and assessment of propositional content, ideas, facts, things or people.

6.5.3.1 NES vs NNES

In the present corpus, evaluative occurrences in the NES texts are 3.87 per 5000 words with a density of 7.6%. The quantity of evaluatives in the NNES texts is 3.07 with 8.0%. The evaluatives were used more frequently in the NES texts (26 % more). From a statistical point of view there is no significant difference between the two groups ($z = -0.955$, $p < 0.340$).

Regarding the usage of the features in this subcategory, *appropriately* and *meaningful* were used the most in both texts (*appropriately*: 8 in the NES and 10 in the NNES; *meaningful*: 8 in the NES and 19 in the NNES). In addition to the most frequent features in both texts, *correctly* was more frequently employed in the NES texts while *unfortunately* was more frequently used in the NNES texts.

Although there are differences in the frequency of evaluatives between the NES and NNES texts, there are some similarities in the purpose of evaluatives in both texts, in that both the NES and NNES writers employed evaluative features for similar purposes:

They had three consecutive hour of class per week. In short, the population and sample of this study were described *appropriately* giving full details in an easily readable fashion (NNES 1).

Grammar is introduced in situations and context where it is *meaningful* to the students (NES 12)

They are used to signal the writers' points of view on the propositional content.

Similarly, *unfortunately* and *correctly* were used to express the authors' attitude:

Listening is also an important skill which also requires sufficient attention. *Unfortunately*, most of the Libyan teachers are not aware of this or they complain that time is not sufficient to conduct the listening activities (NNES 22).

We know long before line 42 however that the events are happening in a court; the writer *correctly* assumes that words like charges, bail, magistrate, bench, sentencing..... (NES 7)

Although the NES writers employ them more frequently, both groups of writers use them with a similar purpose (expressing their attitude and assessment of propositional contents), with the two most frequent features, *appropriately* and *meaningful*.

6.5.3.2 PRO vs Students

The subcategory of evaluatives appeared 3.45 times per 5000 words in the student texts and 2.80 in the PRO ones, with a density of 7.8% and 5.9% respectively. There are differences in the use of evaluatives, the students using more evaluatives than their PRO counterparts (23% more). However, the difference between the two groups is not of any significance (MWU test, $z = -0.855$, $p < 0.393$).

Correctly and *meaningful* are the most frequent features in this subcategory in both group texts. While the student writers employed the features, *disappointingly*, *fortunately*, *desirably*, *hopefully* more they did not occur in the RACORP:

Unfortunately, from my research, there seems to be, in fact, a complete absence of agreement among researchers as to the particular mechanism that underlies social contagion. This is arguably why Moser and Corroyer, *disappointingly* so, do not commit to any conclusions regarding its role in this particular experiment. (NES 9)

Fortunately, those partial understandings and the underplayed status of preference organization were noticed and carefully examined by Boyle (2000).... (NNES 30)

..... and accept simple and inadequate word choice or, more *desirably*, increase a student's awareness of the form of words and encourage them to check (NES 22)

The second task is more difficult. Each sentence in the three texts has had the theme section removed. *Hopefully*, the students will notice that in text three the same concept (NES 7)

The evaluatives (*desirably* and *hopefully* in the NES 22 and NES 7) in the passages above underline the writers' hope that the methods used in their study encourage the learner to get the point.

The journal article writers employed *surprisingly* (5 in POSCORP and 13 in RACORP), and *essentially* (9 in POSCORP and 13 in RACORP) relatively more often to convey the authors' attitude:

Most studies of the WC have focused on tutor-writer interaction and writing center theory. *Surprisingly*, very little of the research reports on the effects of tutoring on the subsequent written products (PRO 22).

In addition to responding to what students have written, we can also consider responding to student processes for writing. *Essentially* what we read at any one time is a product, and it is quite important to understand how that product came to be (PRO 12).

It was noted above that it is often hard to distinguish between emphatics and evaluatives. Indeed, one may argue that writers deliberately exploit the fuzziness that exists between emphatic and evaluative functions. While *surprisingly* and *essentially* can be read as evaluatives, meaning that it is the author's attitude to the following facts, an alternative interpretation is that the authors have emphasised in the facts that 'there is very little research on the effects of tutoring' and 'it is important to understand how the product came to be'. The effect is to maximize their emphasis on the propositional content.

These extracts are therefore good examples of how MD features can behave multifunctionally, already discussed above in sections 6.4.1.1, 6.4.1.2, 6.4.4.1 and 6.4.8.1.

Both student and journal article writers employ evaluative features for similar purposes although students employ them more frequently.

6.5.4 Commentaries

This subcategory refers to addressing the audience directly by using either second person pronouns such as *you*, *your*, or by other forms such as *let us*, *imagine*, *allow me*.

6.5.4.1 NES vs NNES

Commentaries are one of the least used subcategories among IMD. Tables 6-34 and 6-35 show that they occurred 1.58 and 1.89 times in the NES and NNES texts respectively, with density of 3.1% and 4.9% respectively. This means that the commentaries were more frequently used in the NNES texts (19 % more). But the difference is not statistically significant (MWU test, $z = -0.423$, $p < 0.672$).

With respect to variety, 8 different features are used in the corpora, some of which are used repeatedly while others appear only once. The features that are used in the NES texts include: *you* 18 times, *imagine* 4, *(the) reader's* 4, *let us* 3, *your* 3, *allow me* 1. In the NNES corpus; *you* 36 times, *your* 3, *you will* 3; the rest have been used only once and include: *you might be call*, *imagine*, *let us*.

Both writers mostly use the pronouns *you* to address their audience as in:

In building an effective team *you* must consider which employees hold which attributes. *You* need to establish where these individual strengths lie and if further training is necessary (NES 20)

... When a language learner says, 'He go to school everyday', corrective feedback can be explicit, for example, 'no, you should say goes, not go', or implicit, 'yes he goes to school every day', and *you may* or may not include metalinguistic information, for example, 'Don't forget to make the verb agree with the subject'" (Lightbown and Spada 1999, p 171-172). In the case of first language acquisition, it is widely believed, namely by the nativists, that negative evidence plays no significant role in the process of learning a mother tongue (NNES 21)

Where the writers need to talk to the readership in the above extracts, commentaries like *you* and *you may* serve to enhance the reader-friendliness of the text, or to manipulate the audience's thinking by addressing them directly.

Both NES and NNES student writers employ commentary features for similar purposes (i.e. manipulation of the audience's thinking by addressing the reader directly), with the two most frequent features *you* and *your*, although the NNES writers employ them more frequently.

6.5.4.2 PRO vs Students

Commentaries in both student and PRO texts are one of the least used subcategories among IMD (see table 6-36). In the student texts, commentaries occurred 1.75 times per 5000 words with a low density of 3.9%. In the journal article texts, on the other hand, there are 1.10 occurrences, density 2.3%. This means that the student writers employ more commentaries (by 59%) than their PRO counterparts although the difference is not statistically significant (MWU test, $z = -1.902$, $p < 0.057$).

Concerning the features used in this subcategory, the student writers relied on a more limited range of devices, mainly second-person pronouns (*you*, *your*) to address the reader directly, while the PRO writers employed a broad range of devices of commentaries (*imagine*, *let us*, *you might consider*, *you*, *your*). To compare the

frequency of features for commentaries in each group, their frequency of occurrence is listed as follows:

POSCORP: you (54), your (6), imagine (5), (the) reader (4), let us (4), you will (3), you might be call (1), allow me (1)

RACORP: your (16), let us (12), imagine (9), you (7), you might be missing (2), consider (2), you might add (1), you will (1),

In the extract below, the PRO writer employed *let us* to talk to readers about the detail of density measurement step by step in order to guide the reader to understand the complicated figures in the text:

Let us now apply this method of density measurement to the figures for the experiment here. For the native speaker group the hit rate recorded in the study led us to assume an average of 45 links per word. In a network of 1000 items, this would imply a total of 45 000 links. In graph theoretical terms, then, the sum of the degrees of the graph is 45 000, and the actual number of lines (l) in that graph is half that figure: 22 500. Thus, the density of the net of the net is formula range from 0 (a completely unconnected set of points) to 1(PRO 13)

The writer in the extract below is acknowledging the downside of lexical items used in his study to encourage the reader to evaluate the drawbacks of his method in order to justify the methodology later. On this occasion the addressing of the audience is achieved by using the commentary feature *we might imagine*, to describe an alternative lexical item.

In associative terms, then, *we might imagine* that some lexical items are more important than others, either in that they have more connections than other points, or in that they hold positions of strategic significance in the overall structure of the association network by acting as link points between different clusters of associations within the network (PRO 13)

Students employ commentaries more frequently, but rely on a more limited range of devices, mainly second-person pronouns to address the reader (*you, your*). In contrast, the journal article writers have a tendency to use them as a broad range of purposes to guide the reader and justify their methodology as mentioned above.

6.5.5 Appeals

The items classified under this subcategory involve the questions asked by the writer either to address a certain person or institution directly or to attract the readers' attention and involve them in the issues under discussion.

6.5.5.1 NES vs NNES

Appeals are the least used subcategory among IMD (see table 6-34). In the NES texts, it occurred 0.53 per 5000 words only with a very low density of 1.0%. In the NNES texts, there are 0.38 occurrences, density 1.0%. This means that the NES writers use more appeals than their NNES counterparts (39%). However, the difference is not statistically significant (MWU test, $z = -0.678$, $p < 0.498$).

The apparent reason for the very limited quantity of appeals is that in the present work the definition restricts them to questions asked by the writers.

In this subcategory, there are very few occurrences as in:

Teachers of languages are often advised to follow a communicative, task-based approach to teaching. *Do we as teachers follow the more traditional route of limiting the use of L1 in the classroom or should we recognise that the L1 can help our students learn?* When we reflect on the fact that the communicative and task based approaches were devised for use in multicultural classes or for use by ex-patriate teachers who were not proficient in the language of their students, then we must begin to judge how rigidly we follow the guidelines for these methods. Use of L1 by teachers of multilingual classes it is true to say that it is unlikely the teacher of a multilingual class will have knowledge of all of their students' languages. Therefore, "One reason for the lack of reliance on the L1 has undoubtedly been convenience for the teacher", (Cook, 1999:201). Consequently, needs necessitate that the L2 should be used within the classroom (NES 15)

....What circumstances is it made available to the L2 learner in the first place? From Pica, Young, and Doughty (1987, and Pica et al. henceforth), one can extract answers to these questions. Providing contexts in which different kinds, or types, of input are made available in a linguistic environment for L2 learning. The first, being that which is characterized by premodifying the input in some way or another, all taking place before it actually reaches (NNES 21)

The writers here attract the reader's attention by formulating questions which they answer immediately. Thus, the questions in the extracts above were used to construct reader/writer interaction, so that their attention is attracted.

Both NES and NNES student writers employ appeals for similar purposes, to get the reader's attention and involvement in the issues, although the NES writers employ them more frequently.

6.5.5.2 PRO vs Students

Tables 6-36 and 6-37 show that appeals occurred 0.45 and 0.46 times per 5000 words in the student and journal article texts respectively, with a density of 1.0% in both texts. So the PRO writers employ them slightly more than their student counterparts (2% more) without any statistically significant difference. Appeals are the least used subcategory among IMD in both types of text.

In line with the results in Hyland (2001, 2002b), some of the questions in the journal corpus were not answered immediately by the writer:

The immediate questions for educators (including assessors) are the following: *When is it appropriate to give standardized content assessments to ELLs? That is, when are the inferences made about the performance of ELLs on standardized content assessments valid?* (PRO 4)

The central problems of ideational dynamics are, first, *what is the nature of these ideations as cognitive entities*, and second, *how do message-relevant ideations arise?* A third question emerges as a natural extension of these problems: *How do message-relevant ideations come to impact overt behavior?* In other words, if thought drives talk, then *how does it do so?* In the remainder of this article, I explore some of the implications of AAT2 (Greene, 1997, 2000), one theory for addressing these and related issues..... (PRO 26)

By asking a couple of questions which the writers will answer later on in the extracts above, the readers' interest is maintained.

Similarly, the extract below by the student writer which uses questions in her arguments would seem to function for similar purposes, to attract the reader's attention and involve them in the issues under discussion.

...But the question is, *which native speaker do we model and why do we need to use a native speaker model?* Higgins (2003) warns that the NS-NNS dichotomy is 'more of a social construction than a linguistically based parameter'. Advocates of successful L2 user model in language teaching (Cook, 1999; Kwon, 1994) claims that native speaker model is no longer relevant and unrealistic (.....) as it sets unachievable goals for L2 users and can demotivate them (NES 10)

Although there are frequency differences in the use of appeals between the student and journal article writers, there are also similarities in that both groups of writers employ them for similar purposes (to get the reader's attention and encourage the readers to be involved in the issues of discussion).

6.5.6 Self-References

These are pronouns or terms of reference for the writer in the text, e.g. *I, we, me, my, our, us*, etc.

6.5.6.1 NES vs NNES

The frequency of self references is 22.66 per 5000 words in the NES texts and 14.51 in the NNES texts with 44.7% and 37.7% density respectively. The NES students thus employ them significantly more than their NNES counterparts with statistical evidence (56% more, from MWU tests; $z = -2.126$, $p < 0.034$). Self-references are the most frequent subcategory among IMD in both texts (see tables 6-34 and 6-35).

There is a neat native-nonnative split in the choice of self references in this subcategory.

The NES prefer *I* with 165 occurrences while the NNES writers use *I* 94 times.

However, the use of *we* does not differ much between two groups: *we* is preferred to *I*

by the NNES writers (146 in the NES and 142 in the NNES) despite the fact that all the scripts in POSCORP are single-authored. The NNES writers use *we* about as often as the NES writers, while they use *I* much less than their NES counterparts mentioned above. The features used in the NES texts include: *I* 165 times, *we* 146, *my* 73, *our* 51, *us* 24, *me* 15, *researcher* 1 time; in the NNES corpus, *I* (94), *we* (142), *my* (40), *our* (36), *us* (26) and *me* (7).

A comparison of the NES writing with the NNES writing shows the preference for the use of self-references as a collectivist way of elaborating an argument from a Japanese student (e.g. preference for *we* over *I* which will be shown in the following extracts):

....we Japanese tend to apologize more often than U.S. American even though under the situations where Japanese do not need to be accused and they were aware of them. I dare say this is a kind of stereotype. We Japanese express many different feeling (meanings) by using the structures of apology. (NNES 23)

We is used to suggest that the writer and at least one other person were grouped together to represent the Japanese community in order to bond the reader to the same community as the writer. Thus the extract above uses ‘we’ to indicate the force of the discourse act performed not only by the writer but also the extended community of Japanese people together.

However, the more individualistic *I* is used to convey the writer’s own opinion, by which she assumes that the opinion she agrees with will be useful for raising awareness and encouraging the L2 learners in the NES text:

Hence, from a teaching perspective, *I* agree with Cook (1999) who argues that ‘language teaching should adopt the successful L2 user than the native speaker as a model for the L2 learner. This idea will be useful for raising awareness and perhaps encouraging the L2 learners (NES 10)

Despite the fact that there are group preferences of devices in the use of self-references, their purpose in the student texts is for explaining the methodology and procedure of their study:

The transcripts were analysed using the so-called Birmingham Model (Sinclair and Coulthard 1975; 1992), the coding scheme based on discourse analysis. *I* chose their model for this study because, as McCarthy (1991, p.12) and many other researchers point out, it is a “relatively simple and powerful model” which enables researchers to capture not only the details of turn-to-turn exchanges but also the larger structures of the discourse (NNES 17)

The two texts *I* have chosen for analysis are both newspaper reports of men appearing in court to answer criminal charges. The first text is taken from the Guardian of January... (NES 7)

At times, self-references can be linked with procedural detail and research findings throughout the student text to show the discourse act is performed by the writers. Note how *I* is linked to procedural details as well as findings in the following:

In the reduplication of Robertson’s (2000) study, *I* gave a brief explanation of the test, although for control purposes the participants were not informed of the specific nature of the study, and what aspect of English *I* was focusing upon. After the test was completed *I* then informed them what the purpose of the study was (NES 11)

Although *I* agree that predominant L2 use is beneficial to learning, *I* also believe there is a role for the L1 in the classroom. The L1 is already present when teaching the L2. It should be used as a positive tool in the classroom to motivate students, manage tasks, to ensure understanding and negotiating syllabuses (NES 15)

The frequency analysis in this subcategory shows similarities as well as differences between the two groups of texts. The similarities are suggested by the fact that self-references are the most frequent subcategory among IMD in both sets of texts. The differences are suggested by the statistical difference in the frequency of self-references and group preference of *I* over *we* by the NES and vice versa in the NNES texts. However, they use them for similar purposes, mainly describing the procedure of their methods and findings in their academic writing.

6.5.6.2 PRO vs Students

Self references in both student and PRO texts are the most frequent subcategory in IMD (see tables 6-36 and 6-37). In the student texts, they occurred 18.33 times per 5000 words with a density of 41.5%. In the journal article texts, there are 20.87 occurrences, density 43.8%. This means that the PRO writers employ self-references more frequently (13%) than their student counterparts although they are not statistically significant (MWU test, $z = -0.630$, $p < 0.529$).

Interestingly, the student writers and their published counterparts show similarities in their preference of devices in this subcategory, with *we* preferred to *I* in the two corpora, although the PRO writers make more frequent use of *we* than their student counterparts. To compare the frequency of features for self-references in each group, their frequency of occurrence is listed as follows:

POSCORP: I (259), we (288), my (113), our (87), us (50), me (32), researcher (1).

RACORP: I (276), we (345), my (107), our (114), us (70), me (34), researcher (7), author (2).

The following extracts from the RACORP can be read as promotional devices, to advertise the value of their work:

We will need to consider what resources might be harnessed to ameliorate these conditions. The suggestions in this paper are guidelines, ones that teachers may need sometimes to bend and shape to..... (PRO 12)

The interactions between UG and other modules of mind remain constant (so that SLA research will not necessarily contribute anything to the understanding of interface relations). It then becomes of interest to ask whether the changes can be linked to identifiable components of UG. If this turned out to be the case, it would not only produce testable proposals about how SLA differs from FLA, but also about what might count as a discrete module in UG. Since this is a line of research that *I* believe offers considerable promise, *I* will illustrate in slightly more detail with two examples (PRO 3)

However the extracts below from the student corpus focus on what subsequent sections of the paper will do:

We shall begin with a look at the language teaching potential of video generally; while the practical aspect is not ignored, the emphasis in this first section (part A1) will be on theoretical concerns... (NES 17)

I will look briefly at these items and then look in depth at Hofstede's model which the authors base a large part of their work on. Cross Cultural Communication Competence Model One of the two main foci of the study is measurement of Cross Cultural Communication Competence (NES 5)

I will present a short classroom activity aiming at developing learners' awareness of the discourse markers (NNES 19).

In addition to this, the professional writers used *author* to talk about themselves:

In this study, all the interview analyses were performed by the author. An independent rater also coded responses to questions in four of the interviews... (PRO 7)

Similarly, sometimes *researcher* was used instead of more overtly self references (e.g. *I, we*) in journal article texts:

This enabled the researcher to take into account non-verbal means of communicating as well (Hulstijn, J. and Vreeswijk, G. A. W., 2003). Sacks et al have defined a period of speech as a "turn-constructive unit", which may be regarded as the "turn" of one speaker... (PRO 4)

....to support pedagogical practices, suggested practices are also grounded in the experiences of the researcher and of others from.... (PRO 13).

Most often, however, *researcher* and *author* are used in the student texts to refer to other researchers as opposed to the assignment writers themselves:

The author has provided many useful expressions/phrases in each unit to assist learners opening up or ending a conversation (NNES 26).

Annamma Joy, the researcher for the article about Hong Kong interviewed students to retrieve information on their thoughts on gift-giving, (NES 14)

The most frequent purpose of self-references in the student writing is a describing methodology and procedure without evaluation of their methodology:

I then recorded two tests to reconfirm my research questions: 1 Is there evidence of systematic variability? 2 Is it possible to identify a principled account of variability from an interface perspective? 3.4 Recorded Tests for morpheme production 3.4.1 Test 4A – cued oral task I first tested participants with an oral cued sentence completion task about a life-changing situation, using pictures and phrases which required grammatical and contextual adaptation to make a complete accurate paragraph. (NES 23)

In this section, I will use Halliday and Hasan (1976) organised the cohesive resources.... (NNES16)

However, in the RACORP, the effect of self-references is to construct the writer as an expert of the discipline, as someone who is knowledgeable and careful in the methodology, and to underline that the writer tries to solve the potential limitation of the methodology:

To illuminate my retrospective reading, I also pay careful attention to the cover sheet or the annotations students have made detailing their revisions (see previous discussion on communication). Thus, as I plan my response, here too I consider not only the product... (PRO 12)

I added verb voice (active versus passive) in addition to verb tense, word division in addition to spelling, and sentence structure in addition to run-on sentences and fragments. I also added categories of idiom, awkward (not grammatically incorrect but quite infelicitous stylistically), subject verb agreement, repetition or redundancy, pronoun, and need for new paragraph in order to cover all the errors these students made even though most of them were not frequent.... (PRO 21).

A comparison of the student writing with the expert writing shows similarities as well as differences between the two groups' texts. The similarities can be found from the fact that self-references are the most frequent subcategory among IMD in both texts. The differences can be observed from the frequency and functions of the self-references between groups as mentioned above.

6.6 Results from the Interviews

In this section the opinions related to students' preferences and motivations regarding the use of MD, together with their own thoughts and beliefs about good academic writing, derived from the interviews of 7 NES and 7 NNES postgraduate students will be presented, illustrated with quotations of some outstanding viewpoints in order to find out possible reasons for the differences in the use of MD. All respondents are asked to describe what makes academic writing better in order to discover their perceptions of academic writing in general and MD use in particular.

The interview questions were divided into mainly two parts (refer to Appendix 5-2). The first part (parts A.B.C in Appendix 5-2) seeks information about their general perception about their assignments (e.g. audience of their writing, language support, concept of good academic writing, their worries about their academic writing etc.) in order to find out their motivation of the MD usage for research question 1. In the second part (parts D.E in Appendix 5-2) they were asked some questions about the use of the features in academic writing (e.g. the use of signposts, self-references, code glosses etc.) in order to find out their preference regarding the use of MD. The details of the questions are given in Appendix 5-2 and their answers will be compared where appropriate. The followings are the questions asked in the first part:

What do you think makes a good assignment?

Most students think that a good assignment should have relevant content with coherence to convey their argument clearly:

“I think that a good assignment has clear arguments and a clear line of thought. So, the person reading the assignment can understand what the writer is talking about even if they have no prior knowledge of the subject”.

Another respondent added,

“I think the most important thing is relevant contents/ideas with coherence to convey the content clearly”.

They also believe that the way they present the content is important, saying:

“it should be easy to follow and understand. A good assignment needs to have a clear structure and each section must be clearly linked to the next. I think clear presentation and structure are very important”. “I think that a good organisation and good writing style can make a good assignment”.

The postgraduate students think the factors that make good assignments are not only relevant ideas with coherent arguments (the propositional content), but also clear presentation and structure with good links (the characteristics of MD). In other words, they are aware of the importance of using relevant arguments with good presentation skills for a good assignment.

Who do you think is going to read your assignment?/Do you think about the reader when you are writing? Do you think this influences your writing? In what ways? (Do you change anything?):

Even though all of the NNES students think that their reader is mainly the module leader, they try to assume that their reader knows nothing about the topic. Thus they try to explain to the reader the details of their propositional contents with guidance towards

the definition and meaning of terms and concepts. This idea is supported by a non-native speaker who points out the influence of academic writing course,

“Even if my reader is the module leader who is knowledgeable, I was taught in the academic writing course that I should treat the reader as if it is being introduced for the first time to person who doesn’t know anything, so I try to explain meaning of terms which have been introduced in the first time”.

The NES student also mentions that when she writes an assignment she always thinks about the reader and imagines that her reader does not know anything about the topic, thus she tries to write an essay which anyone can understand easily (this can be considered as ‘writer-responsible’ writing, as mentioned in the literature review chapter). And the writer’s understanding of the need to treat their reader as their centre of interest can be a common factor of MD (refer Nash 1992):

“I try to write in a way which would enable anyone to understand the essay, not just people who already have a knowledge in the area”.

“I always think about the reader whenever I write an essay, so I try to explain the details of the contents as possible as I could in order to help my reader understand easily”.

Do you have any worries about writing up your assignment? What worries you most? Regarding language supports, they were asked; Do you think the standard of English is important in an assignment?/Are you worried about your own standard of English when writing an assignment?

Most of the NNES students were worried about the standard of their English when writing an assignment. Therefore 4 out of 7 NNES students use an English native proof-reader for their assignment in order to get a good mark saying,

“Yes, I am worried about my standard of English, because if my module leader doesn't understand my English, then I will get a poor result. That's why I ask my proof reader to change my essays before final submission in order to make it understandable and to get a good mark”, “I started to have a proof reader before submitting a paper and I found that the marks got higher”

whereas one of the seven NES students asked her parents to proof-read for typing errors or stylistic issues.

Most NNES postgraduate students (4 out of 7) have an English native proof-reader for improving their academic writing style in terms of reader-oriented writing style (which is different from their own style) instead of sticking to their own writing style, which they were taught in their own country before they came to England to pursue their degree, because they are worried that the reader will not fully understand what they are writing about:

“I am not used to the English writing style which I suppose a bit different from the way we do in my country. So I ask my proof reader to read my assignment and correct my writing in terms of grammatical errors and add some linking words, to

make it better to help my reader understand easily. I am anxious what if my reader doesn't understand my argument. That's why I ask my proof-reader to change my writing",

"sometimes my proof reader asks me to add some of examples of the contents that she couldn't understand, so I put some of examples to help the reader understand easily by using phrases 'for example', 'for instance' and etc."

Thus the use of an English proof-reader by the NNES writers influences the writing style in the NNES texts. Examples of these can be found from the results that the NNES students made more use of additives (e.g. linking words) and significantly more use of code glosses (e.g. *for example, for instance, such as*) compared to their NES counterparts because of their English proof-reader's contributions.

The following are the questions for the second part:

They were asked the questions: **'Is it important that you tell your reader how the assignment is organised?', 'Are you aware of signposts in your academic writing?', 'What sort of signposts do you use for explaining the structure of your writings?'**

All students agree that it is important to signal to their reader how the assignment is organised, which is sequencers in the proposed typology. Some salient views can support this opinion as follows:

"I always try to explain what I am doing and why. This allows the reader to know where they are going to end up and how they will get there, rather than leading them down a dark alley". It is always important to highlight the structure of your work

and to be consistent”. ”It is important especially in the introduction. Your introduction should inform the reader of how your assignment is organised”.

Some of the NNES students mentioned that they try to tell their reader how their assignments are ordered although they do not bother with explanations of how they put their work in order when they write in their own language:

“I use signposts for the sake of the reader, although I prefer not to use them in Chinese writing”. “I try to make the organisation of my writing as clear as possible. But I usually do not explain much about the organisation when I write in my native language”. “I always keep that in mind when I am writing an assignment in English. Sometimes I use listing numbers as a signpost”.

12 out of 14 postgraduate students agree that the use of sequencers is important to signal how propositions or pieces of information are ordered. In fact, 7 out of 7 NNES students agree that they need to use sequencers as signposts in their academic writing.

The interviewees were also asked the following questions: **‘Can you use ‘I’ in your academic writing?’ ‘Will your readers approve of this?’**, **‘What kinds of things can you talk about ‘I’ with?’**

The positive attitude to the use of self-references in their academic writing can be found from most of the comments by NES students, most of whom had not taken the academic writing course recently, although a couple of them learned about academic writing in their previous school:

“My current supervisor uses ‘I’ in his own writing, so I am happy to do so too”.

“Throughout secondary school I was taught not to use ‘I’, but the graduate level I have learned that it is necessary”.

“‘I’ seems to be acceptable in certain contexts and I frequently see it in academic journal articles”.

Five out of the 7 NNES students that were interviewed had taken an academic writing course recently and had a negative attitude towards the use of *I* in their assignments saying,

“Sometimes, I have to use it especially when I talk about my experience but in most cases I prefer not to use it. I use the passive voice instead”.

“I think that it is probably okay to use ‘I’ with anything, although I personally would use it passively (i.e., ‘subjects were interviewed and it was found that’), which I feel sounds more professional”.

“I was always taught not to use it”.

From their comments, they sometimes use ‘I’ in their academic writing, but they try to avoid it because they were taught to avoid the use of self-references in their academic writing (as often described in the literature: Lester & Lester 2002; Day 1998; Hubbuch 1996; Swetnam 2000).

It is apparent that the NES postgraduate students have positive attitudes towards the use of ‘self references’ because there are many scholars who use *I* in their work, whereas the NNES students are reluctant to use *I* in their assignments because of their educational backgrounds (they were taught to avoid the use of self-references). Indeed the NES students employed significantly more self-references than their NNES

counterparts. This different perception of the use of *I* (which is self-references in the present study) between NES and NNES students was supported by the statistical difference in the use of self-references which were more frequently used by the NES writers (MWU test, $z = -2.125$, $p < 0.034$).

There are similar perceptions; both groups of students agree that the use of MD in their academic writing is important to express their propositional content effectively and clearly in their argument.

However there are a few differences in their perceptions; for instance, the NNES speakers were worried about the standard of their English, so most of them have an English proof-reader to improve their writing style. The NES students, however, seemed not to be worried about the standard of their English, thus only one of the seven NES students has a proof-reader to correct minor errors. So the use of an English proof reader by the NNES writers affects to some extent the use of the features of MD (additives and code glosses) in their academic writing.

In addition to this, there are different opinions about the use of MD between the NES and NNES students. For example, the NES students think positively about the use of self-references whereas the NNES students have negative opinions on the use of *I* in their academic writing. These different perceptions have resulted in the statistically frequent use of self-references in the NES texts.

Nevertheless, the majority of the students say that they are aware of the need to use MD in their academic writing to convey their message effectively and to help their reader understand the content easily.

6.7 Summary

This chapter has presented the results from the different group texts as well as the semi-structured interviews, based on the typology of metadiscourse established earlier (Chapter 4). The results from the text analyses, which have statistical differences, are summarised in the tables below:

Table 6- 38: The Summary of the Statistical Differences from the Comparison between the NES and NNES Texts

Categories of MD	Group Comparisons	Results from MWU tests
Total TMD	NES<NNES	$z=-2.732, p<0.006$
Sequencers	NES<NNES	$z=-2.073, p<0.038$
Code Glosses	NES<NNES	$z=-2.940, p<0.003$
Hedges	NES>NNES	$z=-2.706, p<0.007$
Self references	NES>NNES	$z=-2.125, p<0.034$

Table 6- 39: The Summary of Statistical Differences from the Comparison between the Student and PRO texts

Categories of MD	Group Comparisons	Results from MWU tests
Topicalisers	Student>PRO	$z=-2.292, p<0.022$
Concluders	Student<PRO	$z=-1.981, p<0.048$
Hedges	Student<PRO	$z=-2.605, p<0.009$

As shown in Tables 6-38 and 6-39, five statistically significant differences were found (e.g. in the uses of total TMD, sequencers, code glosses, hedges and self references) in the use of MD between the NES and NNES texts while three (e.g. topicalisers, concluders and hedges) came out of the comparison between the student and PRO texts.

Chapter 7: Discussion

7.1 Introduction

The main emphasis of this chapter is on discussing how the results from the data analysis answer the research questions. The three research questions will be considered with the results for MD use in the different group texts, with focus on two variables as shown in the table below.

Table 7- 1: Research Questions and Variables

Research Questions	Variables
1. What are the types and frequencies of MD used: how differently and similarly are they used in non-native English speaker (NNES) student and native English speaker (NES) student writings?	Language difference (NES vs NNES texts)
2. What are the types and frequencies of MD used: how differently and similarly are they used in each genre (i.e. journal article/Master's assignment)?	Genre difference (student vs journal article texts)
3. How do these factors (language or genre?) affect the pattern of metadiscourse in academic discourse?	Language and Genre factors

Firstly a summary of the answer to research question 1 will be given, followed by the discussion of MD usage in the NES and NNES texts in order to find out the effects of difference of 'language' (i.e. native versus non-native English) (research question 1).

Secondly, the MD differences between the PRO and student texts will be reviewed to

see the effects of 'genre' on the journal article texts and MA assignments (research question 2). After that, research question 3 will be answered as a summary of this chapter.

7.2 A Summary of the Answers to Research Question 1

It is evident that there are some similarities in the use of MD between the NES and NNES student texts with similar purposes, helping the reader understand effectively the content of the text by using TMD and the writer's point of view with the use of IMD. Similarities can also be found from the fact that the total amount of TMD is more than the total IMD in both groups' texts (TMD: 86.4 in the NES and 97.6 in the NNES, IMD: 50.7 in the NES and 38.4 in the NNES).

However, there are differences in the frequency of sub-categories of MD, such as sequencers, code glosses from TMD which were more used in the NNES texts (sequencers: MWU test, $z = -2.073$, $p < 0.038$; code glosses: $z = -2.940$, $p < 0.003$), and from the IMD, hedges and self-references were employed more by the NES writers (hedges: MWU test, $z = -2.706$, $p < 0.007$; Self references: $z = -2.125$, $p < 0.034$).

The results from the comparisons indicate that the NES corpus has a higher frequency for IMD in general and the IMD subcategories, 'hedges' and 'self references' in particular than does the NNES (hedges: MWU test, $z = -2.706$, $p < 0.007$; Self references: $z = -2.125$, $p < 0.034$). The TMD has a higher frequency in the NNES corpus, especially in 'sequencers', and 'code glosses' (see Tables 6-30 and 6-34, sequencers: MWU test, $z = -2.073$, $p < 0.038$; code glosses: $z = -2.940$, $p < 0.003$). In other words, the NNES writers are more concerned about helping to inform readers about the explicit organisation and meaning of the content in their academic writing whereas the NES writers employed relatively little TMD for orienting the reader to the structure of propositions compared to their NNES counterparts. This result does not support the claims in the previous literature (Mauranen 1993a; Crismore et al. 1993; Swales 1990)

that NNES writers employ relatively little MD for the explicit organisation of the text.

Metadiscourse (MD), as one of the discourse sensitive features, can be a subject of language variation. Hinds (1987) and Hall (1983) argue that some languages (e.g. English) are writer responsible/low context while other languages (e.g. Japanese and classical Chinese) are reader responsible/high context (as described in the literature review chapter); one can expect therefore that MD will be used more by the former than the latter. The English writer is expected to produce more MD for well-organised and clear discourse with explicit elucidation of the organisation of the content to help the reader understand their message clearly. However the results for the NNES corpus display a higher frequency in TMD use and so go against the expectation since the NES corpus comes from native speakers of English (writer responsible and low context language from Hinds 1987 and Hall 1983's classification) and the NNES data contain texts provided by Japanese, Korean, Chinese and Arabic writers who are from reader-responsible and high context languages (see Hinds 1987 and Hall 1983).

7.3 Language Aspect (NES vs NNES) for Research Question 1

Let us now focus on the 'language' variable (i.e. native English speaker versus non-native English speaker) to see how it influences the use of MD in their academic writing.

Sequencers:

From the high use of sequencers in the NNES texts with the mapping functions to show the reader what will follow (11.67 per 5000 words in the NES and 21.62 in the NNES texts), the NNES writers seem to enhance the reader-oriented text explicitly and achieve the writer-responsible (Hinds 1987) tactic that they are aware of, as came out in

the interviews. This awareness is reflected by the NNES students, who articulated in the interview:

“I try to make the organisation of my writing as clear as possible. But I usually do not explain much about the organisation when I write in my native language”.

“I always keep that in mind when I am writing an assignment in English. Sometimes I use listing numbers as a signpost although I prefer not to use them in Chinese writing”.

This means the NNES respondents are aware of the need to use sequencing devices in academic writing to help readers, which can be reader-friendly and writer-responsible writing. Their awareness of sequencers as signposts for the reader may have resulted in the greater use of sequencers (e.g. *firstly, secondly, in the second section, lastly, finally*) in the NNES texts (MWU test, $z = -2.073$, $p < 0.038$).

Code Glosses:

The NNES students' understanding of the need to follow the writing style, 'writer responsible', which is more common in English writing (Hinds 1987) resulted in the significantly high use of code glosses in the NNES texts (12.76 in the NES and 17.16 in the NNES, MWU test, $z = -2.940$, $p < 0.003$). This is backed up by the NNES respondents, saying

“As for assignments for academic modules, my intended readers are the module leaders who are English, that's why I try to write an essay as a person who minds the reader, which means I try to explain the jargons as explicit as possible although I had been taught that the implicit content (which the reader can infer the message)

is the better way of writing style back in my home”.

“Even if my reader is the module leader who is knowledgeable, I was taught in the academic writing course that I should treat the reader as if it is being introduced for the first time to person who doesn't know anything, so I try to explain meaning of terms and issues with phrases, 'for example', 'defined as', 'such as' and etc.”

This means that they are aware of the need to help readers grasp meanings of propositional content as a way of the writing style ('writer-responsible' and 'reader friendly' in Hind 1987) which is more commonly used in the U.K. where they are studying rather than sticking with their own writing style ('reader-responsible'). This awareness is from their education in the U.K., academic writing courses that they took before or during their Master courses in the university. This awareness influences the high use of 'code glosses' in the NNES texts.

Also, perhaps more importantly as can be seen from the interview, the worry of communication breakdown in the NNES group writers affects the frequency of code glosses (12.76 in the NES and 17.16 in the NNES, MWU test, $z = -2.940$, $p < 0.003$).

This worry about communication conflict is shown by the NNES student's comment,

“I am anxious what if my reader doesn't understand my argument because of my poor English, so I try to explain the contents by using words, 'for example', 'in other words' and etc, to show some examples of the issues in order to help the reader understand”.

Another respondent mentions that the assistance from her English proof-reader reflects the use of code glosses, saying,

“I ask my proof reader to read my assignment to fix grammatical mistakes and change my writing (which is inappropriate) for easy understanding”, “sometimes my proof reader suggests me to add some of appropriate linking words and some of examples of the contents that she couldn’t understand, so I put some of examples of those to help the reader understand easily by using phrases such as ‘therefore’, ‘on the other hand’, ‘for example’, ‘for instance’ and etc.”.

This influences the high use of ‘code glosses’ in the NNES texts.

In fact, from the interview, four in seven NNES students have English proof-readers for their assignments because most NNES students were worried about the standard of their English when writing an assignment. Thus most of them have an English proof-reader who corrects their grammatical mistakes, and advises them to add some linking words (e.g. additives and concessives) and additional explanations with some features of MD (e.g. code glosses) to help the reader grasp the contents easily in order to avoid a breakdown of communication.

The NNES writers’ worry about communication breakdown because of their poor standard of English suggests why the NNES students make more use of ‘code glosses’ than their NES counterparts to elaborate the definition, meaning of jargon, and details of the propositional contents (12.76 in the NES and 17.16 in the NNES, MWU test, $z = -2.940$, $p < 0.003$).

Hedges:

The less frequent use of hedges in the NNES texts with statistical evidence (9.32 in the NES and 5.34 in the NNES, MWU test, $z = -2.706$, $p < 0.007$) may have resulted from the following factors. Firstly, there appears to be a lack of understanding of the

importance of these features to the reader. Secondly, the infrequent use of hedges in the NNES texts might be from the fact that the NNES tend to avoid leaving uncertain statements in their academic writing. These are reflected by an NNES student saying,

“Sometimes, I use ‘might’ and ‘probably’ if I am not sure about if something is correct or not in the content, but I try not to use them, because it sounds a bit irresponsible in academic writing”.

These findings reflect a high degree of uncertainty avoidance from the academic writing produced by the NNES postgraduate students.

Emphatics:

The result was that the NNES students used more emphatics than their NES counterparts (12.76 in the NES and 13.25 in the NNES). This result contrasts, to some extent, with hedges which were more used by the NES writers than the NNES writers. The high use of emphatics in the NNES texts suggests that the NNES tend to be more emphatic in their statements while the NES’ greater use of hedges (9.32 in the NES and 5.34 in the NNES) shows they prefer to express uncertainty or doubts towards the content with significantly frequent use of hedges (MWU test, $z = -2.706$, $p < 0.007$).

Self-references:

The statistically high use of self-references (22.66 in the NES and 14.51 in the NNES, MWU test, $z = -2.125$, $p < 0.034$) and the group preference for *I* rather than *we* in the NES texts and vice versa in the NNES texts reflect the language differences between native English speakers (who are from individualistic backgrounds) and non-native English speakers (who are from collectivistic backgrounds). In other words, the lack of

use of the authoritative singular form (*I*) of self-references by NNES writers in the present study is a product of a culturally and socially constructed view of self which can be a component of the language variable which renders making statements with *I* difficult.

This finding, to some extent, confirms both Ohta's (1991) and Scollon's (1994) work that the use of first person pronouns is largely unacceptable in the traditions of Asian and Arabic cultures due to its association with individual, rather than collective identity.

The present NNES data were produced by students mainly from Asian and Arabic countries which are collectivist cultures where the group is the source of identity and children are brought up as members of the *we* group (Hofstede 1991), as mentioned in the literature review chapter. It is also possible that the data produced by the NNES students come from speakers of pro-drop languages (Chinese, Japanese, Korean, Arabic), which rarely have subject pronouns. Thus the results of the statistically infrequent use of self-references in the NNES texts (22.66 in the NES and 14.51 in the NNES, MWU test, $z = -2.125$, $p < 0.034$) can provide the evidence of language differences from the native English and non-native English speakers.

In sum, the considerable difference in the use of MD is that the NNES writers employed TMD more and IMD less frequently compared to their NES counterparts, while the NES writers used more IMD and less TMD than their NNES counterparts (TMD: 86.4 in the NES and 97.6 in the NNES, MWU test $z = -2.732$, $p < 0.006$; IMD: 50.70 in the NES and 38.40 in the NNES, MWU, $z = -1.466$, $p < 0.143$). In other words, the NNES writers appear to be more concerned about helping readers grasp the structures and meanings of the propositional content whereas the NES writers try more to interact with the reader about their point of view towards the propositional content.

For the NNES postgraduate students who are from reader-responsible languages (Arabic, Chinese, Japanese, Korean), learning a writing style (writer-responsible), which is different from their own style and acceptable in western academic life, to cope with education in the U.K. may affect the use of MD with overt devices in their texts. Examples of this can be found from the significantly high use of the textual metadiscourse in general and significantly more frequent use of sequencers and code glosses in particular (sequencers: 11.67 in the NES and 21.62 in the NNES, MWU test $z = -2.073$, $p < 0.038$; code glosses: 12.76 in the NES and 17.16 in the NNES, MWU test $z = -2.940$, $p < 0.003$).

Thus this finding may support the rival hypotheses of the CR, mainly for Harder's (1984) suggestion, which explains that L2 writings could be a result of composition learning methods, rather than reflecting the native language and culture.

Because of their education at home, the NNES students seem not to signpost their writing for helping the reader to be aware of how propositions are ordered and what the main ideas are in their academic writing (i.e. they do not use the sequencers and code glosses to hint to the reader what is coming next and to help the reader grasp the main ideas clearly). This has been revealed from the interviews:

"I am worried about the fact that I didn't learn how to write an academic writing to help the reader understand how the writing is ordered and what my point is with signposts in my writing before I came to England, so I spent some time to learn academic writing style before I started the Master course in here".

Some of the NNES informants were aware of the difference in the writing style between their own and that of western students. Thus they were aware of the need to learn academic writing style which they think is different from their own writing style:

“I am not used to the English writing style which I suppose a bit different from the way we do in my country”.

“I had to adopt an English writing style to finish my course successfully”.

Consequently most NNES students interviewed took an academic writing course before or during their Master’s courses. As a result as we see from their interviews cited below, they learnt the importance of using sequencers and cohesion markers which can be devices of MD as mentioned in the literature chapter. Also they are aware of the ‘reader-friendly’ or ‘writer-responsible’ approach to guide the reader where they are going and what is happening at different parts of the text, as disclosed by the interviews,

“yes, I did academic writing course as an optional module, and I learnt that academic writing is to some extent, formal, objective, explicit, and responsible by producing a good presentation with a good linking words for consistent arguments”.

“After I took the academic writing course, I always try to explain what I am doing and why. This allows the reader to know where they are going to end up and how they will get there”.

“ It is always important to highlight the structure of your work and to be consistent”.

”It is important especially in the introduction. Your introduction should inform the reader of how your assignment is organised”.

This awareness may lead to the significantly greater use of the devices of TMD in the NNES texts (TMD: 86.4 in the NES and 97.6 in the NNES, MWU test $z=-2.732$, $p<0.006$). Especially the frequent use of the subcategories in the NNES texts, the greater use of sequencers which are almost identical in their introduction section (e.g. *firstly, secondly, in the second section, finally, lastly*) for the explanation of the organising the texts, and the frequent use of code glosses (e.g. *for example, for instance*) to guide the reader towards the meaning of the terms. However the NES writers, who are from the language (i.e. English) which is based on individualism and writer-responsible, make more use of IMD in general (IMD: 50.70 in the NES and 38.40 in the NNES, MWU, $z=-1.466$, $p<0.143$) and hedges and self-references in particular to interact with readers to express the force of the discourse act and their uncertainty towards the content (hedges: 9.32 in the NES and 5.34 in the NNES, MWU test, $z = -2.706$, $p<0.007$; self references: 22.66 in the NES and 14.51 in the NNES, MWU test, $z = -2.125$, $p<0.034$).

The 'genre' aspect which is for research question 2 will be discussed in the next section.

7.4 A Summary of the Answers to Research Question 2

The overall percentages of the main categories in POSCORP (Postgraduate Students Corpus) and RACORP (Research Article Corpus) are similar to each other (68% and 67% for TMD and 32% and 33% for IMD respectively) thus all the figures are non-significant (TMD: 92.30 in the student and 96.80 in the PRO, MWU test $z=-0.541$, $p<0.588$; IMD: 44.19 in the student and 47.66 in the PRO, MWU test $z=-0.322$, $p<0.747$). However, some of the subcategories (concluders and hedges) used less frequently by the student writers compared to the journal article writers seem to show the differences between groups (concluders: 1.03 in the student and 2.05 in the PRO,

MWU test $z = -1.981$, $p < 0.048$; hedges: 7.21 in the student and 10.50 in the PRO, MWU test, $z = -2.605$, $p < 0.009$).

For the difference in the textual categories, the student writers made significantly more use of 'topicalisers' (to guide the reader to what is happening in different parts of the text, such as topic shift and main topic of sentences; 1.49 in the student and 0.64 in the PRO, MWU test, $z = -2.292$ $p < 0.022$) while journal article writers employed significantly more 'concluders' (which indicate that the writer is about to arrive at a conclusion or a summary of what he has been writing; 1.03 in the student and 2.05 in the PRO, MWU, $z = -1.981$ $p < 0.048$).

For the difference in the interpersonal categories, while the student writers make greater use of 'emphatics' (to express certainty and what writers believe to be true; 13.02 in the student and 11.92 in the PRO, MWU test $z = -1.153$ $p < 0.249$), 'evaluatives' (which signal the writer's attitude and point of view on the propositional content; 3.45 in the student and 2.80 in the PRO, MWU test $z = -0.855$ $p < 0.393$), and 'commentaries' (which address the audience directly; 1.75 in the student and 1.10 in the PRO, MWU test $z = -1.902$ $p < 0.057$), the journal article writers make significantly more use of 'hedges' (to express uncertainty or doubts towards the content in order to mitigate the reader's reaction to the writer; 7.21 in the student and 10.50 in the PRO, MWU test $z = -2.605$ $p < 0.009$).

In summary, the impression is that student writers do not use MD devices in the same way as the PRO writers, as reflected in the use of MD devices with a limited range of items in the student texts compared to a broad range of MD features in the RACORP (e.g. concluders, pre/reviews, commentaries and self references).

In addition to this, the purposes the PRO writers use MD devices for seem to be different from those in the POSCORP. Examples of this can be observed from the varied purpose of MD usage by PRO writers in the use of subcategories (e.g. sequencers, code glosses, sources, hedges, emphatics and self references). In fact the PRO writers use code glosses for the same reason as the students do (expounding their propositional content) and also for answering anticipated queries regarding their propositional content (see the extracts from PRO 18 and PRO 17 in section 6.4.7.2) . It also can be found in the RACORP from the use of combining devices of sources with concessives (e.g. *suggest + however, whereas*), which is for not only achieving a similar effect to what the students' writers do (connecting the writer's work with established work by using the feature of sources, *suggest*) but also for signalling that the journal writers will evaluate the established work which they are from, to offer their own conclusion or methodology as an alternative to other researchers' results or methodologies. Examples can be found in the extracts from PRO 14 and PRO 23 in section 6.4.9.2.

The way the PRO/student writers manifest their texts to the reader with the broad/limited usage of MD is different because of the different nature of their writing. There are the differences in the frequency and purpose of MD usage, which might be from the genre aspect because of the fact that there are different intended readers (as mentioned earlier in the section of 'hedges') and different statuses (as a PRO writer, he is supposed to achieve a unique contribution to the discipline while as a student writer, he wants to get a good mark from their module leader as mentioned in the interviews in section 6.6.)

Thus the different pattern of MD usage in the student and PRO texts, to some extent, supports Perelman's (1982) model of argumentation in the modern rhetoric; the way the argument is put forward is determined by two factors as mentioned in the literature review, the nature of knowledge of the reader and the purpose of the argument (refer to chapter 2). The consideration about the readership and the intention of the students' writing can be observed from the statements in the interviews:

"I focus mainly on what the reader's expectations are (e.g. guidelines, course requirements)".

"I think about what our module leader's expectations when I am writing an assignment. As for module assignments, I do not write for general audience".

"I consider my reader when I am writing an assignment and this influences the assignment in some way or another. For example I try to follow his instructions and the things he asked for which might affect the assignment. Thus I can get a good mark".

In other words the students' perceptions about the readership and the goal of their writing may lead to a different way of writing style in general and the different usage of MD in their academic writing in particular compared to their counter-parts (the PRO).

7.5 Genre Aspect (Assignments versus Journal Articles) for

Research Question 2

To elaborate, if a genre variable (i.e. student assignment versus journal article texts) influences the pattern of MD usage, the results of individual subcategories in the following sections will be discussed with a focus on the differences from the preferred

features and purposes in the use of MD (see Tables 6-32 and 6-36).

Sequencers:

The student writers employed a more fixed pattern of sequencers (*firstly, secondly, lastly, etc*) in the introduction of their assignment whereas the PRO writers used these features much less and use the alternative features of sequencers, such as additives, rationales (e.g., *in addition, thus, because of this*) within a running discourse. The results of this may indicate that the student writers are trying to have a more fixed pattern of text structure with explicit features of sequencers than their counterparts.

However the alternative use of sequencers in the PRO texts by using other features of additives and rationales (*in addition, thus, because of*) and a word *both* to imply the logic of the sentences, may suggest how skilled writers create their texts with implication of their intentions and signal from which the reader can tell what comes next and how these are related to each other. Thus the results of these may have reflected the difference of the genre which may stem from the different status of the writers: it might be the students' lack of skill, writing as novices. Also, from the opinions from the interview, 12 out of 14 postgraduate students agree that the use of sequencers is important in showing how pieces of information are ordered, reflecting the fact that the sequencers are more frequently used in the student texts than the journal article texts.

Topicalisers:

The significantly greater use of the overt features of topicalisers (e.g. *with regard to* and *based on*) in the student texts (1.49 in the student and 0.64 in the PRO, MWU test, $z=-2.292, p<0.022$) suggests that the student writers rely heavily on employing the

topic of the sentence and seem to have considered them as the foundation of the context. The results associated with the fact that the PRO writers used topicalisers much more broadly (e.g. sentences with '*developing the ideas presented in Cook (1997), ...*' and '*let me refer to.....*') in their preferred devices for signalling what the topics of sentences are and the changing of topics instead of using the overt devices of topicalisers (e.g. *in regard to, as for, based on*) may suggest how the journal article writers convey their message by signalling and implication of meanings with a varied range of features (topic shift, main topic of sentence). The results of this provide another example of the difference in the use of MD with a broad range of features in the RACORP.

Concluders:

From the result of the significantly frequent use of concluders in the journal article texts (2.05 per 5000 words in the PRO and 1.03 in the student, MWU test, $z=-1.981$, $p<0.048$), it can be speculated that this is derived from the nature of articles, which constructs their novelty of arguments by using concluders. In fact, most devices in this sub-category are used in the RACORP to introduce a summary of what has been said in the previous section, thus providing further support for getting more attention and encouraging the reader to read. Accordingly, the writers may feel the need to add their unique conclusions explicitly to disseminate the originality of their work in their community. However the less frequent use of concluders in the student texts (1.02 less per 5000 words) may represent their status as novices which may restrict the freedom of the writers to add their own conclusions and cause their statements to be left with an overall implicit conclusion.

Pre/reviews:

There are differences in the preferred devices in this sub-category between the two groups. While the PRO writers more frequently used features for guiding the reader with diagrams, tables, and figures (e.g. *table x*, *figure x*), the student writers more frequently employed features for alerting the reader to what has been going on in the previous parts, (e.g. *previously/earlier* and *above*). The results of this may be due to the different nature of writings between the groups. The journal article writers are probably spending more time conducting their research and thus try to explain their results as reliably as possible by providing evidence with figures and tables whereas the student writers seem not to be pressed to produce reliable results as much as the PRO writers are because of their different status. In fact the postgraduate students seem not to spend much time in preparing each assignment because of their limited time schedule with substantial course work (about 7 or 8 assignments for 120 credits) and a dissertation within a year.

Concessives:

The higher frequency of concessives in the RACORP compared to those in the POSCORP (23.63 in the student and 28.57 in the PRO although the difference is not statistically significant), supports Barton's (1993) argument, that experienced academic writers frequently use 'evidentials' of contrast (e.g. *but*, *however*) to problematise their topics to show that they need reconsideration. This is reflected by the high use of concessives in the journal article texts and the three most frequently used features of concessives in the RACORP, '*although*', '*but*' and '*however*' to signal a change of direction in the running discourse in order to modify the preceding proposition.

Code glosses:

The journal article writers employ code glosses more frequently (15.22 in the student and 19.76 in the PRO although statistically non-significant) with varied purposes than do the student writers. This is supported by the fact that the PRO writers use code glosses for the same reason as the students do (giving a detailed explanation of their propositional content) and also as a way of answering anticipated queries regarding their propositional contents and expounding of their contents (see the extract from PRO17 in the section 6.4.7.2). From the results of this, the PRO writers seemed to be aware of possible attack by the readers, while the student writers seemed to try to help their reader (module leader) grasp what they mean by the propositional content rather than preparing for potential attack by the reader.

Sources:

A comparison of the student writings with the PRO ones shows that the two sets of writings are different in the use of sources in terms of frequency and purpose. For the difference in the frequency, the journal article writers used more sources than their student counterparts (6.27 in the student and 7.04 in the PRO although statistically non-significant). In addition to this, there are differences in their preferred choice of devices with a broad range of purpose in the PRO texts. For instance, *x suggested* was employed more frequently by the journal article writers (34 in the student texts and 104 in the journal article texts) to imply that their work is based on earlier research or other researchers' ideas and signal that they will evaluate the work by using the combined devices of concessives (*however, whereas*). Thus they can develop their own methodology and conclusion with originality in order to convey the novelty of their work to the discipline. However '*according to x*' was more frequently used in the

POSCORP (153 in the POSCORP and 73 in the RACORP) to link their work with established works, rather than evaluation of the previous literature.

It is probable that the PRO writers are supposed to present their original results and own conclusions to build a credible writer identity, rather than delivering a repetitive work which is linked and similar to the established work, whereas the writers of student assignments are unlikely to be so concerned about establishing proof of their own academic ability. Consequently, the frequency and the preferred features in the use of sources in the two groups are different from each other (6.27 per 5000 words in the student and 7.04 in the PRO).

Hedges:

The clearest difference between these two groups is found within the IMD subcategory, particularly hedges, which exhibit a statistically higher frequency in RACORP than the corresponding category does in POSCORP (7.21 in the student and 10.50 in the PRO, MWU test, $z = -2.605$, $p < 0.009$, see Table 6-29).

The less frequent use of ‘hedges’ in the student texts (7.21 in the POSCORP and 10.50 in the RACORP) may have resulted from the following factors. Firstly, it might be a lack of the students’ understanding of the importance of these features to the addressed audience since their reader is mainly the module leader (as can be observed from the interview as mentioned earlier in section 7.2):

Sometimes, I use ‘might’ and ‘probably’ if I am not sure about if something is correct or not in the content, but I try not to use them, because it sounds a bit irresponsible in academic writing”

It might thus be because of the 'difference of the readership' (see Swale 1990) between the journal article and student texts. The readership of the journal articles is 'academics', 'students', 'material textbook designers' and similar people who are interested in the topic while those of the student texts is mainly the module leader who assesses the assignments. For the student assignments the readership is more restricted than for journal articles. So the significant infrequency of the hedges in the student corpus (7.21 in the POSCORP and 10.50 in the RACORP, MWU test, $z = -2.605$, $p < 0.009$) suggests that the student writers do not mind if their statements are being attacked because of the lesser use of hedges. The results of this are reflected by Perelman's (1982) model of argumentation; the way of the argument is determined by the difference of the nature of knowledge of the reader and the purpose of the writing (refer to literature review chapter).

Secondly, it could be a result of the absence of accountability and loose ownership of ideas in the student texts, which might have created a state whereby the student writers do not feel committed to validate their statements or express uncertainty (refer to the high degree of uncertainty avoidance from the NNES interview data).

The purpose of the use of hedges in the student texts is to acknowledge the writer's uncertainty about their reader's agreement while they are employed in the RACORP, not only for the acknowledgement of their uncertainty like students' do, but also for justification of their limitations of study and protection from the readers' attack, as can be seen from the extracts in section 6.5.1.2 in the previous chapter (extracts: PRO13, 18, 10, 21, 30). The results of this provide evidence of differences in the purpose of MD usage, that the PRO writers use the MD devices with a broad range of purposes while the student writers use them with a limited range of purposes.

Emphatics:

Devices of emphatics in the student texts are much more frequent compared to those in the journal article texts (13.02 in the student and 11.92 in the PRO). These findings are in line with the literature as Hewings and Hewings (2002) and Hyland (2005a) suggest that the students use more emphatics than their professional counterparts in their academic writing.

However the student writers use them with a limited purpose, mainly for emphasis of the propositional content while the PRO writers employ them for variety of purposes, such as emphasis of the previous sentences, acknowledging that the writers are concerned about the limitations of the methodology by using the feature *always* and demonstrating that they are resisting the anticipated criticism of their work with the use of *clearly* in their texts (see extracts, PRO1, PRO15 and PRO13 in the section 6.5.2.2). This shows how skilled writers express their text effectively to emphasise the propositional contents and acknowledge the limitations of their study by using devices of emphatics with a varied range of purpose.

Commentaries:

Students employ commentaries more frequently, but rely on a limited range of devices, mainly second-person pronouns (e.g. *you, your*) to address the reader directly (1.75 in the student and 1.10 in the PRO). In contrast, the journal article writers tend to use a broad range of devices (e.g. *you, your, imagine, let us, consider* and etc.) which can be used as an effective medium of academic writing in order to enhance the reader-friendliness of the text by using *you* and *you may* and guide the reader to understand the complicated figures with the use of *let us*. The results of this are another example of the

difference from the two different genres; the PRO writers use a broader range of MD devices compared to their student counterparts.

Self-references:

A comparison of the student writing with the journal article writing shows similarities and differences in the use of self-references. The similarities are suggested by the fact that self-references are the most frequent subcategory among IMD in both group texts (see tables 6-36 and 6-37), which might reflect the understanding of the use of self references as rhetorical devices in their academic writing. The differences are suggested by the purpose and frequency of the self-references. The difference in the frequency between groups is reflected by the fact that self-references were used infrequently in the POSCORP (18.33 in the student and 20.87 in the PRO). This would suggest that the students were told to avoid personal pronouns in their academic writing during their education (refer to literature: Lester & Lester 2002; Day 1998; Hubbuch 1996; Swetnam 2000), supported by the interview data such as,

“I was always taught not to use ‘I’ in the academic writing”.

“I learnt that ‘I’ is inappropriate in academic writing”.

Alternatively, the students may have believed that ‘good’ academic writing is free of self- references supported by interview evidence such as-

“I think that it is probably okay to use ‘I’ with anything, although I personally would use it passively, which I feel sounds more professional in academic writing”.

As a consequence, unlike RACORP, neither *I* nor *we* was used frequently in POSCORP.

Thus the reluctance and under-use of authorial pronouns (*I* and *we*) by the student writers might be due to the educational contexts where the students were taught to avoid the use of self-references in their academic writing and might be from their individual perceptions about the use of self-references which were disclosed in the interviews. Both of these possibly influence the students to be reluctant to use self-references in their academic writing.

For differences in the purpose, the most frequent purpose of self-references in the student writing seems to be describing procedural details in their methodology (e.g. *I* chose their model, *I* gave a brief explanation of the test) while the effect of self-references in the RACORP is to construct the writer's status as an expert who tries to be rigorous and careful about methodology by getting rid of the potential limitation of their study by using phrases (e.g. *I* also pay careful attention, *I* also added categories of idiom in order to cover all the errors, *we* will need to consider, *I* believe offers considerable promise).

The professional writers used devices of self-references, such as '*author*' and '*researcher*' to talk about themselves while most uses of '*researcher*' and '*author*' in the student texts were employed to refer to other researchers. Examples can be found in the extracts PRO7, PRO4, PR13, NNES26, and NES14 in section 6.5.6.2. The evidence for this is that the PRO writers use a broad range of MD devices to indicate the force of the discourse act performed by the writer in their texts.

7.6 A Summary of the Answers to Research Question 3

When including the language variable, the difference in the MD usages mainly applies to frequency rather than to purposes. One possible explanation for group differences is

the NNEs students' awareness of the need to use MD as a way of western writing style, 'reader-friendly' or 'writer-responsible' to help readers grasp the propositional contents and structures. As evidence for this we have observed some views from the interview with the NNEs writers such as:

"I always think about the reader whenever I write an essay in English, so I try to explain the details of the contents as possible as I could in order to help my reader understand easily",

"I am not used to the English writing style which I suppose a bit different from the way we do in my country. So I always try to explain what and why I am doing the subject matter at the beginning of my writing".

This awareness might lead them to the high use of TMD in general, and the significantly greater use of sequencers and code glosses in particular. However the high use of IMD in general, and the statistically more frequent use of hedges and self-references in particular in the NES texts suggest that the NES writers try more to express their uncertainty and to show the force of the discourse act by themselves in their academic writing.

Nevertheless, as can be observed from the interviews, both groups of postgraduate students agree on the need to use MD as a good presentation skill in order to deliver their messages effectively and appropriately. Evidence of this can be observed from the interviews:

"I think that a good assignment needs to have a good presentation, such as a clear structure and each section must be clearly linked to the next with some appropriate linking words",

“It is always important to highlight the purpose and topic of your work at the beginning of your writing to inform the reader of what your writing is about and how this is related to the next clearly”.

Their similar perceptions about the need to use of MD in their academic writing might bring about the more frequent use of certain features of MD (e.g. sequencers, additives and topicalisers) in the student texts compared to their PRO counterparts.

When we include the genre aspect, the differences are both in the frequency of the MD usage and the way they use MD. Regarding the differences in the frequency, the PRO writers use significantly more concluders and hedges while the student writers use significantly more topicalisers. And for the differences in purpose, the PRO writers use the MD devices with a broad range of purposes while the student writers use them with a limited range. Examples of this can be found from the use of MD subcategories, code glosses, sources, hedges, emphatics and self-references which were employed with a limited range of purposes in the student texts.

As regards the language aspect from the comparison between the NES and NNES, the differences are mainly in the amount of features in the use of MD subcategories. When it comes to the student and journal article texts, genre variable, the differences are not only in the frequency of MD subcategories but also in the way they use the MD features.

The differences in the amount of MD usage are suggested by the statistical differences in the frequency of MD subcategories between the NES and NNES texts and the student and journal article writings. The differences in the use of MD, apart from the frequency, are reflected by the fact that there is an alternative use of sequencer features

in the PRO corpus (e.g. the alternative use of features in additives and rationales for the function of the sequencers), a broad range of features in the use of topicalisers in the PRO texts (see appendix 6-5-2), a preferred choice of MD features within a subcategory (e.g. the use of sources with a combined feature of concessives, *suggest + however, whereas* in the PRO texts while *according to* was more frequently used in the student texts; the use of pre/reviews, *figure x* and *table x* were employed more in the PRO texts while the greater use of *above* and *previously* in the student texts) and a broad range of purposes in journal article texts (e.g. the additional purposes of the use of MD subcategories compared to those in the student texts; code glosses, sources, hedges, emphatics and self references).

In conclusion, research question 3 (How do these factors (language or genre?) affect the pattern of metadiscourse in academic discourse?) has two different answers. As regards the comparison of texts produced by the NES and NNES postgraduate students, the answer is that the language variable plays a greater role in the use of MD in their academic writing.

This is supported by the differences that the NNES writers use significantly more TMD while the NES writers employ more IMD in general and the statistical differences in the use of MD subcategories in particular. Evidence of this is the significantly high use of subcategories, sequencers and code glosses which were more frequently used by the NNES writers, and hedges and self-references which were more frequently used by the NES writers.

From the comparisons between the student and PRO texts, the genre/expertness variable is also an important factor influencing the pattern of MD usage within the same discipline. This is supported by the differences both in the frequency and the

purpose of MD usage. This is confirmed by the fact that the PRO writers made more use of concessives, concluders, sources, hedges and self references with varied purposes while the student writers made significantly more use of sequencers, and more use of topicalisers, emphatics and commentaries with limited purposes.

Chapter 8: Conclusion

8.1 Introduction

This last chapter of this study is concerned with the contributions made by the study. It is organised in two main parts: the theoretical and empirical perspectives. Then implications of the study will be highlighted for teaching and material development. After that some suggestions for further research will be put forward, deriving from the perceived limitations of this study. This chapter ends with some concluding remarks.

8.2 Theoretical Contributions

After a critical review of the studies related to MD and a survey of the elements used in the different definitions of MD, it appeared that:

- a) the different available definitions of MD are very brief (from different terms which are parallel to MD), too broad and general (especially Beauvais's and Hyland's) and do not help reduce the fuzziness which is the nature of MD
- b) there is overlapping between the subcategories used in most of the typologies as a characteristic of MD is its fuzziness
- c) there is a lack of a clear link between MD definitions on the one hand and the typologies and functions on the other.

Therefore, this study put forward a working definition of MD that is rather extended (taking the term 'MD' which is a broader name, discussing textual as well as interpersonal functions, rather than 'metatext' which is parallel to MD and a narrower one with only textual elements), and linked to the proposed typology used in the study as well as the functions of MD.

The elements of this working definition were selected from the MD definitions proposed by other scholars (Vande Kopple 1985; Crismore 1985, 1990; Crismore et al 1993; Hyland 1998), and their contents were modified to better reflect the purpose of the present study.

This working definition and the typology based upon it led to testing out the following novel ideas in the thesis:

- a) Confirming Hyland and Tse's (2004) findings, TMD is preferred to IMD in the academic writing produced by different group writers. The findings show that there is more TMD than IMD in all three group texts. This indicates that they use more features of TMD that help readers grasp the organising, classifying, and meaning of the texts than those for interacting with the reader about the propositional content with IMD in their academic writing.
- b) Suggesting that the NNES writers are more concerned about helping readers to be informed about the explicit organisation and meaning of the terms in their academic writing with the significantly high use of TMD in general and the significantly frequent use of sequencers and code glosses in particular. This finding is different from the previous claims (Mauranen 1993a; Crismore et al. 1993; Swales 1990) that the NNES writers employ relatively little MD for the explicit organisation of the text.

8.3 Empirical Contributions

This section includes some contributions from the research methods used in the study.

Firstly the study compared three types of writing produced by NES, NNES and PRO writers to apply 'a double-contrastive aspect' for the research. This was achieved by comparisons of the MD usage between the two groups in terms of both 'language aspect' and 'genre aspect', with pedagogical implications for postgraduate students in general and for advanced EFL learners in particular, who want to make academic contributions like the journal article writers do.

Secondly the study utilised semi-structured interviews with actual contributors, unlike other MD studies. So it has explored similar and different perceptions about the use of

MD between the NES and NNES students (e.g. their similar perceptions of the need to use MD in general and particularly sequencers; and their different perceptions of the use of hedges and self references). Thus it can suggest possible reasons for the similarities and differences in the use of MD between groups.

Thirdly the statistical analysis was performed in the study to determine if the differences are meaningful rather than relying on mean frequencies which can distort results by chance if data are not normally distributed. This will hopefully help to develop a rigorous methodology for MD study.

This study takes two variables to find out how language or genre difference influences the pattern of metadiscourse in academic writing. Thus the contributions at the empirical level with two variables are based on three sources: the first is the results from the comparison of MD use in the texts produced by native and non-native English speaking students for the language variable. The second source is the outcome of the comparison in the use of MD between the journal article texts and students' assignments for the genre aspect. The third source is the findings from semi-structured interviews to support the results from the comparisons between the texts.

8.3.1 From the Language Difference

The study provided evidence for the influence of the language difference on MD use. This is confirmed by the differences in MD use between the NES and NNES texts: for example, the NNES writers employed TMD more while the NES writers made more use of IMD. This means the NNES writers are more concerned about the organisation of the text through TMD and the NES writers more with interaction with the reader about the propositional contents with IMD.

The findings of the study also show that learning a writing style which is acceptable in western academic life (e.g. 'writer-responsible') influences the use of MD in the NNES academic writing. As a result of taking an academic writing course, they learnt the importance of using sequencers and cohesion markers which can be devices of MD and they were told to avoid personal pronouns in their academic writing. Also due to taking an academic writing course, they are aware of the importance of a 'reader-friendly' or 'writer-responsible' approach to guide the reader where they are going and what is meant by the terms in their text. Evidence of this can be found from the interview data and the results from the text analysis which show the statistically greater use of TMDs (e.g. sequencers, code glosses) in the NNES texts and the significantly infrequent use of self-references in the NNES texts compared to their NES counterparts. Thus these findings may support the rival hypotheses of the CR, leaving some room for consideration of the educational context.

The study showed a high use of IMD in general and significantly more use of hedges and self-references in particular in the NES texts, indicating that the NES writers tried to involve the readers more in the argument than their NNES counterparts.

The differences between the NES and NNES texts can be attributed to language factors for the following reasons: firstly, the statistically greater use of TMD in the NNES texts provides a solid basis for the influence of language differences (native/non-native English). Secondly, the statistically high use of sequencers and code glosses in the NNES texts and more frequent use of hedges and self references in the NES texts show that 'language' is a variable that affects the pattern of metadiscourse in academic writing in the same discipline.

8.3.2 From the Genre Difference

The study found that the two sets of writing (RACORP and POSCORP) differ in the frequency and purpose (effect) of MD usage. The statistical differences in frequency can be found from the use of MD subcategories, such as topicalisers which were more used by the student writers, and concluders and hedges which were employed more frequently by the PRO writers. Another difference was seen in the use of MD devices with a limited range of items and purposes in the student texts compared to those in the RACORP (e.g. concluders, pre/reviews, commentaries).

In summary, from the results of the two contrastive analyses (language and genre differences), the study found that the language variable plays a larger role in the frequency of MD usage. This is suggested by the statistical difference from the frequency of the main MD categories that the NNES uses significantly more TMD.

This is also confirmed by the statistical difference in the use of MD subcategories, for example, sequencers and code glosses which were more employed by the NNES writers and hedges and self-references which were used more frequently in the NES texts.

In addition to this, the study found that the genre difference explains the statistical difference in the use of MD (e.g. topicaliser, concluders, hedges) and the way they use MD with different intentions where the PRO writers have varied purposes in the use of concessives, concluders, sources, hedges and self-references while the student writers have limited purposes in the use of sequencers, topicalisers and emphatics. So the findings of this study suggest that the genre variable is also an important factor influencing the pattern of MD usage where considerable differences were found

between the student and journal texts in terms of frequency and purpose of the MD usage in their academic writing.

Thus this study confirms that the variables which manipulate the pattern of MD use are both 'language', which affects the significant differences in the frequency of MD usage between the NES and NNES texts, and 'genre' which reflects the significant differences in the frequency and the considerable differences in the purpose of MD usage between the journal articles and student writings within the same discipline.

8.4 Implications

Since this research is involved with MD usage in academic writing, English teachers and people involved in designing teaching materials for both native English speakers and non-native English speakers can use the findings to improve the effectiveness of teaching reading and writing style in English.

8.4.1 Implications for Teaching

More pronounced use of metadiscourse is associated with improved comprehension of reading (Camiciottoli 2003) and the students who had direct teaching of metadiscourse in the composition of their writing skills did better on post tests than the students who were not taught about MD (Steffensen and Cheng 1996). So, instructors can provide exposure to the various features of MD that are considered important and appropriate as rhetorical devices, rather than believing that students will gain this exposure by themselves. Therefore, students should be taught the varied features and purposes of MD in their academic writing for competent writing that conveys their propositional contents effectively to the reader.

So the findings of this study can be used to teach postgraduate students academic writing style. The way the PRO writers use MD can be a model to teach them academic writing since they provide a good example of how skilled writers lead their reader to their texts with implications of the writer's various intentions whereby the reader can tell what comes next when moving from one idea to the next and a modification of proceeding propositions in the text.

For example, they use the frequent use of certain features of MD to signal that they will modify their topic in the following section by using concessives (*however, on the other hand*) within arguments. They also use the alternative use of MD devices (additives and rationale features) to carry out functions of sequencers in order to hint at the writer's intention. So the readers can be notified what comes next although the writers did not use overt features of sequencers (*firstly, secondly, lastly*). All students, who want to use appropriate language to convey their arguments effectively in their academic writing, should be taught how to use these features with their various intentions to convey their propositional contents efficiently and interact with their readers appropriately as they present their research in their academic writing.

In addition, journal writers are more likely to use features of sources combined with concessive devices (*however, whereas*) for the purpose of evaluating the previous work done by other researchers. Thus, showing the origin of ideas by using devices of sources (*suggest*) and evaluating other researchers' work by using the combined devices of concessives (*however, whereas*) in the PRO texts are the skills all advanced postgraduate writers should be taught.

Therefore the awareness of the use of metadiscourse and its appropriate use of implications in their academic writing will help students increase the effectiveness of

conveying messages to a certain degree and learn a successful method of rhetorical practice. These can help the postgraduate students who want to be academic scholars, develop their future career successfully.

8.4.2 Implications for the Development of Materials

According to the literature (Dahl 2004; Harwood 2003; Hyland 1998; 2000), writers' use of MD seems very much related to the subject matter of the piece of writing, materials should be developed to teach the MD feature related to the topic or discipline in order to develop academic writing style as well as the effectiveness of written communication skills of a particular subject area. Thus the findings of the current study can be utilised for the meaningful and new information that EAP textbook writers can use in the area of topic related language, culture and communication.

For example, the findings: TMD outnumbers IMD in all three group texts, and there are similarities in their use of MD; the three groups texts share the top four of TMDs (sequencers, additives, concessives, code glosses) and the three most frequent subcategories in their choice of IMD subcategory (hedges, emphatics, self-references) etc.

The findings of the present study will hopefully be utilised to increase students' awareness of the need to use MDs as rhetorical devices to facilitate effective written communication skills as mentioned above. Accordingly, specialists working in the area of MD need to take language and genre differences into consideration with the relationship between MD use and different factors. It is recommended that these issues be explained to students of academic writing and also discussed with them explicitly, especially at the advanced stages, like postgraduate level, to demonstrate how the use

of MD in the academic writing is related to the factors and how they can use the differences for their written communications.

8.5 Limitations and Suggestions for Further Research

One of the serious problems in carrying out this study on metadiscourse is the classification of the different main categories and subcategories. The review of the existing typologies showed considerable overlap between the subcategories within the same typology. Throughout the different stages of the analysis, I found that many metadiscourse features could be assigned to several subcategories simultaneously (e.g. *in addition*, *moreover* and *because of this* to describe the sequence of the text structure as parts of a process of listing and these can be classified as sequencers, additives and rationales, and *surprisingly* and *essentially* can be read as evaluatives and emphatics, etc). This characteristic creates problems not only for the actual analysis but also for considering metadiscourse as a non-content aspect of discourse and calculating it as MDs. However, for the purpose of the quantification of the frequency in the current study, I followed other researchers, as mentioned in chapter 5, and classified them as only one subcategory which appears to be a primary function.

As far as student writing is concerned, while my student corpora are admittedly small, all the assignments analysed in the same discipline come from the same institution - Newcastle University – and the same school – Education, Communication and Language Sciences. This therefore allows me to make some claims about the departmental/institutional preferences regarding MD use. However their writings are natural and unbiased since the data were produced without knowing anything about the purpose of the study. So the data can be representative of many postgraduate students in the world.

The point is that local/institutional factors can affect preference. Do Newcastle University students use MDs in the same way as students from other institutions? How closely does student MD use in equivalent disciplines score from institution to institution? If time had allowed, then, it would have been interesting to have compared the NES and NNES students' MD use with MD use in other assignments from other UK universities, and with assignments from universities outside the UK. Larger corpora with a greater variety of topics and with writing from both native and non-native speakers of English should be analysed to validate the present findings.

Student writing is a very different genre from that of expert writing (Samraj 2002).

However, I have only studied one genre of student writing (the postgraduate students' assignments) and journal articles. The selection of two different genres of writing, the research articles and the assignments, seems to show different MD practices in general and some of the sub-categories of MD (the PRO writers made more use of concessives, concluders, sources, hedges and self references with a varied function while the student writers made significantly more use of sequencers, topicalisers and emphatics with a limited function) in particular, even within the same discipline. Thus further research is needed to compare and contrast student MD use in the same genre as experts.

8.6 Concluding Remarks

The use of metadiscourse (MD) in academic texts in the same discipline by NES, NNES and professional writers was examined to find out how the main variables affect the pattern of MD use. Both similarities and differences were found in the use of certain MD elements. Writers from different groups produced different/similar frequencies and rhetorical effects through their use of MD elements.

Thus, the findings associated with the similarities and differences can be used to teach academic writing to students who are studying those areas and will hopefully play a part in improving the teaching of writing style in general, and help students develop their rhetorical practice as a resourceful use of metadiscourse knowledge to achieve a desired effect in order to underline their purposes and persuade readers with their intention in their academic argument in particular.

References:

- Alkaff, A.A.R.O. (2000) *Metadiscourse in Texts Produced in English by Yemeni/Arab Writers: A Writer/Reader Oriented Cross-Cultural Analysis of Letters to the Editor*, PhD Thesis, University of Exeter
- Austin, J.L. (1962) *How to Do Things with Words*, Oxford, Clarendon Press
- Barnes, J. (1984) (ed.) *The Complete Works of Aristotle. Vol. 2*, Princeton, N.J., Princeton University Press
- Barton, E.L. (1993) Evidentials, Argumentation, and Epistemological Stance, *College English*, 55/7: 745-769
- Beauvais, P.J. (1986) *A Speech Act Theory of Metadiscourse*, PhD Thesis, Michigan State University
- Biber, D., Conrad, S. & Reppen, R. (1998) *Corpus Linguistics: Investigating Structure and Use*, Cambridge, Cambridge University Press
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999) *Longman Grammar of Spoken and Written English*, London, Longman
- Bloor, M. (1997) Techniques of Validation in Qualitative Research: a Critical Commentary in Miller, G. & Dingwall, R. (eds) *Context and Method in Qualitative Research*, London, Sage Publications, 37–50
- Bondi, M. (1999) *English Across Genres: Language Variation in the Discourse of Economics*, Modena, Edizioni Il Fiorino
- Bruce, B. (1981) A Social Interaction Model of Reading in R. Freedle (Ed), *Discourse Processes*, 4: 273-311
- Bruner, J. (1984) Pragmatics of Language and Language Pragmatics, *Social Research*, 55/4: 969-984

- Bryman, A. & Cramer, D. (2001) *Quantitative Data Analysis with SPSS Release 10 for Windows: a Guide for Social Scientists*, London, Routledge.
- Bryman, A. (2001) *Social Research Methods*, Oxford, Oxford University Press
- Bunton, D. (1999) The Use of Higher Level Metatext in PhD Theses, *English for Specific Purposes*, 18 Supplement: 41-56
- Butler, C. (1997) Repeated Word Combinations in Spoken and Written Text: Some Implications for Functional Grammar in Butler, C., Connolly, J., Gatward, R. & Wismans, M. (Eds.) *A Fund of Ideas: Recent Development in Functional Grammar*, Institute for Functional Research in Language and Language Use, Amsterdam, IFOTT, 60-77
- Camiciottoli, B.C. (2003) Metadiscourse and ESP Reading Comprehension: an Exploratory Study, *Reading in a Foreign Language*, 15/1: 1-16,
<http://nflrc.hawaii.edu/rfl/April2003/camiciottoli/camiciottoli.html>
- Chafe, W. & Nicholas, J. (eds.) (1986) *Evidentiality: The Linguistic Coding of Epistemology*, Norwood, NJ: Ablex.
- Connor, U. (1996) *Contrastive Rhetoric: Cross-Cultural Aspects of Second Language Writing*, New York, Cambridge University Press
- Conover, W.J. (1999) (3rd ed.) *Practical Nonparametric Statistic*, New York, John Wiley & Sons, Inc
- Conrad, S. (1999) The Importance of Corpus-based Research for Language Teachers, *System*, 27: 1-18
- Conrad, S., & Biber, D. (2001) *Variation in English: Multi-dimensional Studies*, London, Pearson
- Cook, V. (2006/2007) The Nature of the L2 User, in L. Roberts, A. Gurel, S. Tatar and L. Marti, ed. *EUROSLA Yearbook 7*, Amsterdam, John Benjamins, 205-220
available at
<http://homepage.ntlworld.com/vivian.c/Writings/Papers/EURO2006.htm>.

- Crismore, A. & Farnsworth, R. (1990) Metadiscourse in Popular and Professional Science Discourse in W. Nash, (ed.) *The Writing Scholar: Studies in Academic Discourse*, Newbury Park, CA, Sage Publications, 119-136
- Crismore, A. (1982) *The Metadiscourse Component: Understanding Writing about Reading Directives* (ERIC Document Reproduction Service No. ED 217 374)
- Crismore, A. (1983) *Metadiscourse: What It is and How It is Used in School and Non-school Social Science Texts*, Centre for the Study of Reading, University of Illinois Technical Report No.273
- Crismore, A. (1984) *Metadiscourse as a Rhetorical Act in Social Sciences Study Texts: Its Effect on Student Performance and Attitude*. PhD Thesis, University of Illinois.
- Crismore, A. (1985) *The Case for a Rhetorical Perspective on Learning from Texts: Exploring Metadiscourse*. Center for the Study of Reading, University of Illinois. Technical Report No. 342.
- Crismore, A. (1989) *Talking with Readers: Metadiscourse as Rhetorical Act*, New York, Peter Lang Publishing
- Crismore, A. (1990) Metadiscourse and Discourse Processes: Interactions and Issue, *Discourse Process*, 13: 191-205
- Crismore, A., Markkanen, R. & Steffensen, M. (1993) Metadiscourse in Persuasive Writing: A Study of Texts Written by American and Finish University Students, *Written Communication*, 10/1: 39-71
- Crismore, A. and Vande Kopple, W.J. (1997) The Effects of Hedges and Gender on the Attitudes of Readers in the United States toward Material in a Science Textbook. In: A. Dusak, Editor, *Culture and Styles of Academic Discourse*, Mouton de Gruyter, Berlin, pp. 223–247.
- Dahl, T. (2004) Textual Metadiscourse in Research Articles: a Marker of National Culture or of Academic Discipline? *Journal of Pragmatics* 36 1807-1825

- Day, R.A. (1998) (5th ed) *How to Write & Publish a Scientific Paper*, Cambridge, Cambridge University Press
- Denzin, N.K. (1989) *The Research Act: a Theoretical Introduction to Sociological Methods*, Englewood Cliffs, NJ, Prentice Hall
- Dillon, G. (1986) *Rhetoric as Social Imagination: Exploration in the Interpersonal Function of Language*, Bloomington, IN, Indiana University Press
- Fraser, B. (1990) An Approach to Discourse Markers *Journal of Pragmatics* 14: 383-395.
- Geertz, C.G. (1973) *The Interpretation of Cultures*, New York, Fontana Press
- Goodenough, W.H. (1971) *Culture, Language and Society McCaleb Module in Anthropology*, Reading, MA, Addison-Wesley
- Grabe, W. & Kaplan, R.B. (1996) *Theory and Practice of Writing*, London, Longman
- Grimes, J. (1975) *The Thread of Discourse*. The Hague: Mouton.
- Gumperz, J.J., & Levinson, S.C. (1996) *Rethinking Linguistic Relativity*, Cambridge, UK, Cambridge University Press
- Hall, E.T. (1983) *The Dance of Life: The Other Dimension of Time*, Garden City, NY, USA, Anchor/Doubleday Press
- Hall, E.T. & Hall, M.R. (1990) *Understanding Cultural Differences: Germans, French, and Americans*, Yarmough, Maine Intercultural Press, Inc
- Halliday M.A.K. (1973) *Explorations in the Functions of Language*, London, Edward Arnold
- Halliday, M.A.K. (1994) *An introduction to functional grammar*. London, Edward Arnold.
- Halliday, M.A.K. & Hasan, R. (1976) *Cohesion in English*, London: Longman.

- Harder, B. (1984) Cultural Attitudes in Discourse, *Canadian Journal of Applied Linguistics*, 29/2: 115-130
- Harris, Z.S. (1959) Linguistic Transformations for Information Retrieval, *Papers in Structural and Transformational Linguistics*. Formal Linguistics Series 1 Dordrecht, Holland, D. Reidel, available at http://books.nap.edu/openbook.php?record_id=10866&page=937 on 29/11/08, 937-950
- Harris, Z.S. (1966/1970). Algebraic Operations in Linguistics Structure, *Papers in Structural and Transformational Linguistics* Formal Linguistics Series 1. Dordrecht, Holland, D. Reidel: 603-611.
- Harwood, N. (2003) *Person Markers & Interpersonal Metadiscourse in Academic Writing: a Multidisciplinary Corpus-based Study of Expert & Student Texts*, PhD Thesis, Canterbury Christ Church University College
- Harwood, N. (2005) 'I Hoped to Counteract the Memory Problem, but I Made No Impact Whatsoever': Discussing Methods in Computing Science Using I, *English for Specific Purposes*, 24: 243–267
- Hewings, M. & Hewings, A. (2002) 'It is Interesting to Note That...': a Comparative Study of Anticipatory 'it' in Student and Published Writing, *English for Specific Purposes*, 21/4: 367-383
- Higher Education Statistics Agency (HESA) (2005) available at <http://www.universitiesuk.ac.uk/faqs/showFaq.asp?ID=7#Whatproportionofstudentsstudyineachsubjectarea> on 20/05/06
- Hinds, J. (1987) Reader Versus Writer Responsibility: A new Typology in U. Connor and R.B. Kaplan (eds) *Writing Across Languages: Analysis of L2 Text*, Reading, MA, Addison Wesley, 141-152
- Hinds, J. (1990) Inductive, Deductive, Quasi-Inductive: Expository Writing in Japanese, Korean, Chinese, and Thai in Connor, U. and Johns, A.M. (eds.) *Coherence in*

Writing: Research and Pedagogical Perspectives, Washington, D.C., TESOL Publication, 89-109

Hoey, M. (1979) *Signalling in Discourse*. Birmingham: University of Birmingham

Hoey, M. (1983) *On the Surface of Discourse*, London: George Allen & Unwin.

Hofstede, G. (1991) *Cultures and Organizations: Software of the Mind*, London, McGraw-Hill

Hollander, M. & Wolfe, D.A. (1999) *Nonparametric Statistical Methods*, New York, John Wiley & Sons, Inc

Hooper, R. (1981) The Taken for Granted, *Human Communication Research* 195-211.

Hubbuck, S.M. (1996) *Writing Research Papers across the Curriculum*, Fort Worth, Harcourt Brace

Huck, S.W. & Sandler H.M. (1979) *Rival Hypotheses: Alternative Interpretation of Data Based Conclusions*, New York, Harper & Row

Hunston, S. (2002) *Corpora in Applied Linguistics*, Cambridge, Cambridge University Press

Hyland, K. (1998) Persuasion and Context: the Pragmatics of Academic Metadiscourse, *Journal of Pragmatics*, 30: 437-455

Hyland, K. (1999a) Talking to Students: Metadiscourse in Introductory Textbooks, *English for Specific Purposes*, 18: 3-26

Hyland, K. (1999b) Disciplinary Discourses: Writer Stance in Research Articles in Candlin, C.N. & Hyland, K. (eds.) *Writing: Texts, Processes and Practices*, Harlow, Longman, 99-121

Hyland, K. (2000) *Disciplinary Discourses: Social Interactions in Academic Writing*, Harlow, UK, Longman

- Hyland, K. (2001) Humble Servants of the Discipline? Self-mention in Research Articles, *English for Specific Purposes*, 20/3: 207-226
- Hyland, K. (2002a) Authority and Invisibility: Authorial Identity in Academic Writing, *Journal of Pragmatics*, 34: 1091-1112
- Hyland, K. (2002b) What Do They Mean? Questions in Academic Writing, *Text*, 22/4: 529-557
- Hyland, K. (2002c) Directives: Argument and Engagement in Academic Writing, *Applied Linguistics*, 23/2: 215-239
- Hyland, K. (2004a) Disciplinary Interactions: Metadiscourse in L2 Postgraduate Writing, *Journal of Second Language Writing*, 13: 133-151
- Hyland, K. (2004b) Graduates' Gratitude: the Generic Structure of Dissertation Acknowledgements, *English for Specific Purposes*, 23: 303-324
- Hyland, K. (2005a) Representing Readers in Writing: Student and Expert Practices Linguistics and Education, *An International Research Journal*, 16/4: 363-377
EJ743779
- Hyland, K. (2005b) *Metadiscourse Exploring Interaction in Writing*, New York, Continuum
- Hyland, K. & Tse, P. (2004) Metadiscourse in Academic Writing: a Reappraisal, *Applied Linguistics*, 25/2: 156-77
- Hyland, K. & Tse, P. (2005) Hooking the Reader: a Corpus study of Evaluative *that* in abstracts, *English for Specific Purposes* 24: 123-139
- Intaraprawat, P. (1988) *Metadiscourse in Native Speakers and ESL Students' Persuasive Essays*, PhD Thesis, Illinois State University
- Intaraprawat, P. & Steffensen, M. (1995) The Use of Metadiscourse in Good and Poor ESL Essays, *Journal of Second Language Writing*, 4/3: 253-272

- Jones, S. & Tetroe, J. (1987) Composing in a Second Language in A. Matsuhashi (Ed.) *Writing in Real Time: Modelling Production Processes*, Norwood, NJ, Ablex, 34-57
- Kachru, Y. (1995) Contrastive Rhetoric in World Englishes: an Analysis of the Cultural Differences in Rhetorical Styles across the Many Varieties of English Worldwide, Particularly as Expressed through Writing, *English Today*, 11/1: 21-31
- Kaplan, R.B. (1966) Cultural Thought Patterns in Inter-Cultural Education, *Language Learning*, 16: 1-20
- Kaplan, R.B. (1982) An Introduction to the Study of Written Texts: the 'Discourse Compact', *Annual Review of Applied Linguistics*, 5: 138-151
- Kaplan, R.B. (1987) Cultural Thought Patterns Revisited in Connor, U. & Kaplan, R.B. (eds.) *Writing Across Languages: Analysis of L2 Texts*, Reading, MA, Addison-Wesley, 9-21
- Keller, E. (1979) Gambits: Conversational Strategy Signals. *Journal of Pragmatics* 3: 219-238.
- Kerr, A.W., Kozub, S.A. & Hall, H.K. (2002) *Doing Statistics with SPSS*, London, Sage Publications
- Labov, W. (1972) Some Principles of Linguistic Methodology, *Language in Society* 1, 97-120.
- Lado, R. (1986) How to Compare Two Cultures in Valdes, J.M. (ed.) *Culture Bound*, Cambridge, Cambridge University Press, 52-63
- Lautamatti, L. (1978) Observations on the Development of the Topic in Simplified Discourse, in Kohonen, V. & Enkvist, N.E. (Eds) *Text Linguistics, Cognitive Learning, and Language Teaching*, Turku, Finland, AFinla, 71-104.

- Leech, G. (1991) The State of the Art in Corpus Linguistics in K. Aijmer & B. Altenberg (eds.) *English Corpus Linguistics- Studies in Honour of Jan Svartvik*, London, Longman, 8-29
- Leki, I. (1991) Twenty Five Years of Contrastive Rhetoric: Text Analysis and Writing Pedagogy, *TESOL Quarterly*, 25/1: 123-143
- Lester, J.D. & Lester J. Jr (2002) (10th ed) *Writing Research Papers: a Complete Guide*, New York, Longman
- Levinson, S. (1985) *Pragmatics*, Cambridge, CUP.
- Liebman, J.D. (1992) Toward a New Contrastive Rhetoric: Differences between Arabic and Japanese Rhetorical Instruction, *Journal of Second Language Writing*, 1/2: 141-165
- Markkannen, R., Steffensen, M. & Crismore, A. (1993) Quantitative Contrastive of Metadiscourse: Problems in Design and Analysis of Data, *Papers and Studies in Contrastive Linguistics*, 28: 137-151
- Matalene, C. (1985) Contrastive Rhetoric: An American Writing Teacher in China, *College English*, 47/8: 789-808
- Mauranen, A. (1993a) A Contrastive ESP Rhetoric: Metatext in Finnish-English Economics Texts, *English for Specific Purposes*, 12: 3-22
- Mauranen, A. (1993b) *Cultural Differences in Academic Rhetoric*, Frankfurt am Main, Peter Lang
- McCarthy, M. (1991) *Discourse Analysis for Language Teachers* Cambridge: CUP
- McEnery, T. & Wilson, A., (2001) (2nd ed.) *Corpus Linguistics*, Edinburgh, Edinburgh University Press
- Meyer, B.J.F., & Rice, G.E. (1982) The Interaction of Reader Strategies and the Organization of Text, *Text, Interdisciplinary Journal for the Study of Discourse*, 2: 155-192

- Meyer, B.J.F. & Rice, G.E. (1981) Organisational Strategies in Prose Comprehension Across the Adult Life Span, Paper Presented at the *Annual Meeting of the American Psychological Association*, Los Angeles, CA.
- Meyer, B.J.F. (1975) *The Organisation of Prose and Its Effects on Memory*. Amsterdam, North Holland
- Meyer, B.J.F., Brandt, D.M. & Bluth, G.J. (1980) The Use of Top-Level Structure in Text: Key for Reading Comprehension of Ninth Grade Students, *Reading Research Quarterly*, 16:172-103.
- Mohan, B.A. & Lo, A.Y. (1985) Academic Writing and Chinese Students: Transfer and Development Factors, *TESOL Quarterly*, 19/3: 515-534
- Myers, G. (1989) The Pragmatics of Politeness in Scientific Articles, *Applied Linguistics*, 10/1: 1-35
- Nash, W. (1992) *'The Meaning of Metadiscourse' an Uncommon Tongue*, London, Routledge
- Nattinger, J. & Decarrico, J., (1992) *Lexical Phrases and Language Teaching*, Oxford, Oxford University Press
- Nystrand, M. (1986) *The Structure of Written Communication: Studies in Reciprocity between Writers and Reader*, Toronto, Academic Press
- O'Keeffe, A. & Farr, F. (2003) Using Language Corpora in Initial Teacher Education: Pedagogic Issues and Practical Applications, *TESOL Quarterly*, 37/3: 89-418
- O'Keeffe, A., McCarthy, M & Carter, R. (2007) *From Corpus to Classroom: Language Use and Language Teaching*, Cambridge, Cambridge University Press
- Ohta, A. (1991) Evidentiality and Politeness in Japanese, *Issues in Applied Linguistics*, 2/2: 183-210

- Ostler, S. (1987) English in Parallels: A Composition of English and Arabic Prose in Connor, U. & Kaplan, R.B. (eds.) *Writing Across Languages: Analysis of L2 Text*, Reading, MA, Addison-Wesley, 169-185
- Perelman, C.H. (1982) *The Realm of Rhetoric*, Translated by Kluback, W. Notre Dam, University of Notre Dame Press
- Plett, H.F. (1985) Rhetoric in Van Dijk, T.A. (ed.) *Discourse and Literature*, Amsterdam, John Benjamins Publication Co, 58-84
- Popovic, A. (1976) Aspects of Metatext, *Canadian Review of Comparative Literature*, 3:225-235.
- Potter, J. (1996) Discourse Analysis and Constructionist Approaches: Theoretical Background in J.T.E. Richardson (Eds.) *Qualitative Research Methods for Psychology and Social Sciences*, Leicester, British Psychological Society, 125-140
- Prior, P.A. (1998) *Writing/Disciplinarity: A Sociohistoric Account of Literate Activity in the Academy*, Mahwah, NJ: Lawrence Erlbaum Associates
- Purves, A.C. (1988) Introduction in Purves, A.C. (ed.) *Writing across Languages and Cultures: Issues in Contrastive Rhetoric*, Newbury Park, California, 9-21
- Ragan, S.L. & Hooper, R. (1981) Alignment Talk in the Job Interview *Journal of Applied communication*, 9:85-103.
- Rahman, M.M. & Govindarajulu, Z. (1997) a Modification of the Test of Shapiro and Wilk for Normality, *Journal of Applied Statistics*, 24: 219-235
- Redeker, G. (1990) Ideational and Pragmatic Markers of Discourse Structure, *Journal of Pragmatics* 14:367-381.
- Salkind, N.J. (2000) *Statistics for People who (think they) Hate Statistics*, London, Sage Publications

- Samraj, B. (2002) Texts and Contextual Layers: Academic Writing in Content Courses in A.M. Johns (ed.), *Genre in the Classroom: Multiple Perspectives*, Mahwah, Lawrence Erlbaum Associates, 163-176
- Schiffrin, D. (1980) Metatalk: Organisational and Evaluative Brackets in Discourse, *Social Inquiry: Language and Social Interaction*, 50: 199-236
- Schiffrin, D. (1987) *Discourse Markers*. Cambridge: CUP.
- Scollon, R. (1994) As a Matter of Fact: the Changing Ideology of Authorship and Responsibility in Discourse, *World Englishes*, 13: 34-46
- Scollon, R. (1997) Contrastive Rhetoric, Contrastive Poetics and Perhaps Something Else?, *TESOL Quarterly*, 31/2: 352-358
- Scott, M. (2001) Comparing Corpora and Identifying Key Words, Collocations, Frequency Distributions through the Wordsmith Tools Suite of Computer Programs in M. Ghadessy, A. Henry & R.L. Roseberry (Eds.) *Small Corpus Studies and ELT: Theory and Practice*, Amsterdam, John Benjamins, 47–67
- Scott, M. (2004) *Oxford WordSmith Tools Version 4.0*, Oxford, Oxford University Press
- Searle, J.R. (1969) *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: CUP.
- Searle, J.R. (1976) A Classification of Illocutionary Acts, *Language in Society* 5: 1-23.
- Selinker, L. (1972) Interlanguage, *International Review of Applied Linguistics*, 10/3: 209-231
- Selinker, L. (1992) *Rediscovering Interlanguage*, New York, Longman
- Skulstad, A.S. (2005) The Use of Metadiscourse in Introductory Sections of a New Genre, *International Journal of Applied Linguistics*, 15/1: 71-86

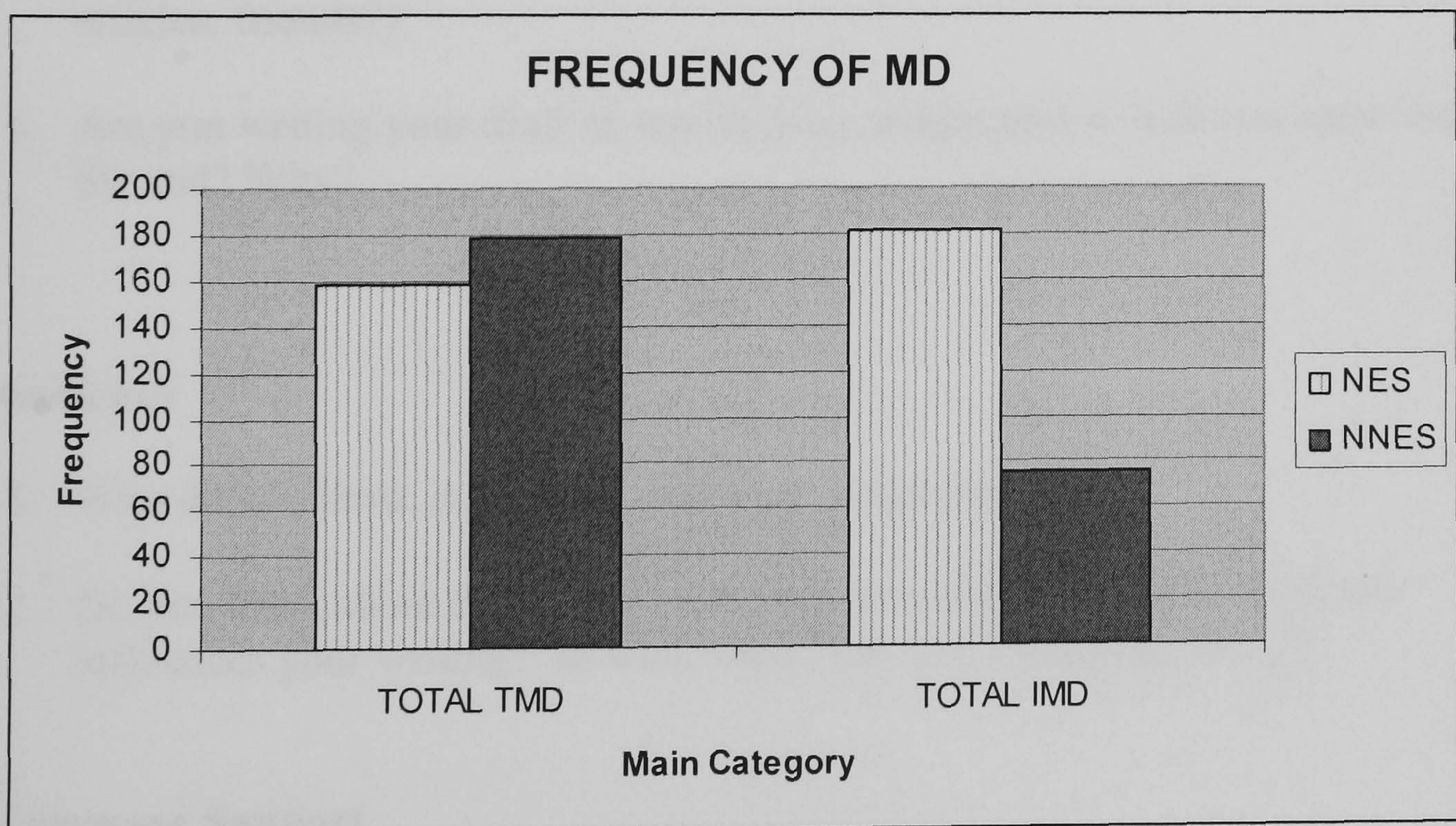
- Stainton, C. (1993) *Metadiscourse and the Analytical Text: a Genre-Based Approach to Children's Discourse*, PhD thesis, University of Manchester
- Steffensen M.S. & Cheng, X. (1996) Metadiscourse and Text Pragmatics: How Students Write after Learning about Metadiscourse in L. F. Bouton (Ed.) *Pragmatics and Language Learning Monograph Series*, 7, 153-171. (ERIC Document Reproduction Service No. ED400709).
- Stubbs, M. (2001) *Words and Phrases*, Oxford, Blackwell
- Swales, J.M. (1990) *Genre Analysis: English in Academic and Research Settings*. Cambridge, Cambridge University Press
- Swetnam, D. (2000) (3rd ed) *Writing Your Dissertation: How to Plan, Prepare and Present Successful Work*, Oxford, How To Books
- Thomas, J. & Short, M. (1996) (eds.) *Using Corpora in Language Research*, London, Longman
- Tierney, R.J., & Raphael, T.E. (April, 1981) *Factors Controlling the Inferences of Fifth Grade Readers: an Extended Examination of the Author-Reader Relationship during Discourse Processing*, paper presented at the Annual Meeting of the American Educational Research Association, Los Angeles
- Tierney, R.J., Raphael, T., Lazansky, J. & Mosenthal, J. (November, 1979) *Readers, Text and Texture: an Extended Examination of the Relationship between Readers and Authors*, National Reading Conference, San Antonio
- Toulmin, S., Rieke, R. & Janik, A. (1979) *An Introduction to Reasoning*, New York, Macmillan
- Toulmin, S.E. (1958) *The Uses of Argument*, Cambridge, Cambridge University Press
- Usunier, J. (1993) *International Marketing: a Cultural Approach*, Harlow, Prentice Hall

- Valdes, J. (1986) (Ed.) *Culture Bound: Bridging the Cultural Gap in Language Teaching*, Cambridge, Cambridge University Press
- Valero-Garces, C. (1996). Contrastive Rhetoric: Metatext in Spanish-English Economics Texts, *English for Specific Purposes*, 15: 279-294
- Vande Kopple, W.J. (1980) *Experimental Evidence for Functional Sentence Perspective*, PhD Thesis, University of Chicago
- Vande Kopple, W.J. (1985) Some Exploratory Discourse about Metadiscourse, *College Composition and Communication*, 36/1: 82-93
- Widdowson, H. (2000) On the Limitations of Linguistics Applied, *Applied Linguistics* 21/1: 3-25
- Williams, J. (1979) Non-linguistic Linguistics and the Teaching of Style in Mcquade, D. (ed.) *Linguistics, Stylistics and the Teaching of Composition* (pp24-40). Akron: University of Akron.
- Williams, J. (1981) *Style: Ten Lessons in Clarity and Grace*, Glenview, IL, Scott, Foresman
- Williams, J. (1990) *Style: Toward Clarity and Grace*, Chicago, Chicago University Press
- Winter, E. (1977) A Clause-Relational Approach to English Texts: A Study of Some Predictive Lexical Items in Written Discourse, *Instructional Science*, 6/1
- Young, R.E., Becker, A.L. & Pike, K. (1970) *Rhetoric: Discovery and Change*, New York, Harcourt

Appendices:

Appendix 5- 1: Results from the Pilot Study

Category	NNES1	NNES2	All (NNES)	NES 1	NES2	All(NES)
Textual (total)			178			158
Sequencers	18	52	70	5	14	19
Topicalisers	2	4	6	0	0	0
Concluders	0	0	0	0	0	0
Previews/reviews	3	4	7	0	5	5
Additives	8	18	26	18	14	32
Concessives	26	1	27	20	19	39
Code Glosses	6	15	21	13	24	37
Rationales	5	5	10	7	4	11
Sources	4	7	11	12	3	15
Interpersonal(total)			75			181
Hedges	5	1	6	10	14	24
Emphatics	19	9	28	17	24	41
Evaluatives	4	1	5	10	3	13
Commentaries	0	0	0	4	2	6
Appeals	1	1	2	0	4	4
Self references	0	34	34	35	58	93



Appendix 5- 2: Questionnaire for Semi-Structured Interview

Preamble

I am conducting a study of academic writing by postgraduate students at the University of Newcastle upon Tyne. This will mainly look at texts, but I want to get an idea of how students see the process. I want to ask you about your views on academic writing.

A. General

1. What is your mother tongue?
2. What do you think makes a good assignment?
3. Do you have any worries about writing up your assignment? What worries you most?
4. How often do you do academic reading in your field?
5. Where do your ideas about assignment writing come from? (books, journal articles, friends?)
6. Are you writing your draft as you do your assignment or will you leave it until the end? Why?

B. Audience

1. Who do you think is going to read your assignment?
2. Do you think about the reader when you are writing? Do you think this influences your writing? In what ways? (do you change anything?)

C. Language Support

1. Do you think the standard of English is important in an assignment?
2. Are you worried about your own standard of English when writing an assignment?
3. Have you done any academic writing course before?

4. Do you know if there is any help you can get when writing, such as the Language Centre, websites, or proof reading by English native-speaker? Do you use any of these?

D. Features of Academic Writing

1. Is it important that you tell your reader how the assignment is organised?
2. Are you aware of signposts in your academic writing?
3. What sort of signposts do you use as connectors, citations, conventions?
4. What can you say if you are not sure if something is correct or not? /if you are uncertain about an idea or about a result? (e.g. might/perhaps/possible/about)
5. What can you say if you are definite or very confident about an idea or a result?
6. Can you use “I” in your academic writing? Will your reader approve of this?
7. What kinds of things can you talk about “I” with? (e.g. your method? your ideas? your results? Acknowledgements?)

E. Attitude

1. Do you think it is important to give your attitude to what you are writing about or should you be neutral?
2. Is it OK for you to express your opinions? Can you express emotions? What kind of emotions?

Thank you very much for your help with the study.

Appendix 5- 3: A List of Journal Articles used in the Professional Corpus (PRO)

Script no	No of words	Details
PRO1	10863	Bell, N. (2005) Exploring L2 Language Play as an Aid to SLL: A Case Study of Humour in NS–NNS Interaction, <i>Applied Linguistics</i> 26/2 192-218
PRO2	7722	North, S. (2005) Disciplinary Variation in the Use of Theme in Undergraduate Essays , <i>Applied Linguistics</i> 2005 26: 431-452
PRO3	7864	Hawkins, R. (2001) The Theoretical Significance of Universal Grammar in Second Language Acquisition, <i>Second Language Research</i> , 17/4, 345-367
PRO4	5409	McKay, P (2005) Research into the Assessment of School-age Language Learners, <i>Annual Review of Applied Linguistics</i> , 25, 243-263
PRO5	4395	Murray, D. (2005) Technologies for Second Language Literacy, <i>Annual Review of Applied Linguistics</i> , 25, 188-201
PRO6	10532	Marsella, A.J. (2005) Culture and Conflict: Understanding, Negotiating and Reconciling Conflicting Constructions of Reality, <i>Int'l Journal of Intercultural Relations</i> , 29, 651-673
PRO7	7832	Juffs, A. (2005) The Influence of First Language on the Processing of <i>Wh</i> -movement in English as a Second Language, <i>Second Language Research</i> , 21/2, 121-151
PRO8	6707	Ward, C. (2006) Acculturation, Identity and Adaptation in Dual Heritage Adolescents, <i>Int'l Journal of Intercultural Relations</i> , 30, 243-259
PRO9	6411	Palthe, J. (2004) The Relative Importance of Antecedents to Cross-Cultural Adjustment: Implications for Managing a Global Workforce, <i>Int'l Journal of Intercultural Relations</i> , 28, 37-59
PRO10	7011	Hinkel, E. (2006) Current Perspectives on Teaching the Four Skills, <i>TESOL Quarterly</i> , 40/1, 109-131
PRO11	6587	Reynolds, D. W. (2002) Causality in the Writing of Middle-grade English Language Learners, <i>Journal of Second Language Writing</i> , 11, 311-328
PRO12	8599	Goldstein, L. M. (2004) Questions and Answers about Teacher Written Commentary and Student Revision: Teachers and Students Working Together, <i>Journal of Second Language Writing</i> , 13, 63-80
PRO13	8184	Barcroft, J. (2004) Effects of Sentence Writing in Second Language Lexical Acquisition, <i>Second Language Research</i> 20/4, 303-334
PRO14	6618	Robinson, P. (2005) Aptitude and Second Language Acquisition, <i>Annual Review of Applied Linguistics</i> , 25, 46-73
PRO15	7970	Tardy, C.(2004) The role of English in scientific communication: lingua franca or Tyrannosaurus rex?, <i>Journal of English for Academic Purpose</i> , 3, 247-269
PRO16	5226	Green, J.A. (2003) The Writing on the Stall: Gender and Graffiti, <i>Journal of Language and Social Psychology</i> , 22/3, 282-296
PRO17	11004	Flowerdew, J. (2001) Attitudes of Journal Editors to Nonnative

		Speaker Contributions, <i>TESOL Quarterly</i> , 35/1, 121-150
PRO18	12634	Boltz, M. G. (2005), Temporal Dimensions of Conversational Interaction: The Role of Response Latencies and Pauses in Social Impression Formation, <i>Journal of Language and Social Psychology</i> , 24/2, 103-138
PRO19	8222	Kramsch, C. (2005) Post 9/11: Foreign Languages between Knowledge and Power, <i>Applied Linguistics</i> , 26, 545-567
PRO20	6759	Johnson, K. E. (2006) The Sociocultural Turn and Its Challenges for Second Language Teacher Education, <i>TESOL Quarterly</i> , 40/1, 235-257
PRO21	11135	Chandler, J. (2003) The Efficacy of Various Kinds of Error Feedback for Improvement in the Accuracy and Fluency of L2 Student Writing, <i>Journal of Second Language Writing</i> , 12, 267-296
PRO22	11312	Williams, J. (2004) Tutoring and Revision: Second Language Writers in the Writing Center, <i>Journal of Second Language Writing</i> , 13, 173-201
PRO23	4707	Hudson, T. (2005) Trends in Assessment Scales and Criterion-referenced Language Assessment, <i>Annual Review of Applied Linguistics</i> , 25, 205-227
PRO24	4432	Cotterill, J. (2004) Collocation, Connotation, and Courtroom Semantics: Lawyers' Control of Witness Testimony through Lexical Negotiation, <i>Applied Linguistics</i> 25, 513-537
PRO25	2530	McKay, S. (2003) Adolescent Risk Behaviors and Communication Research Current Directions, <i>Journal of Language and Social Psychology</i> , 22/1, 74-82
PRO26	4861	Greene, J. O. (2006) Have I Got Something to Tell You: Ideational Dynamics and Message Production, <i>Journal of Language and Social Psychology</i> , 25/1, 64-75
PRO27	5355	Merkin, R. S. (2006) Uncertainty Avoidance and Facework: A Test of the Hofstede Model, <i>Int'l Journal of Intercultural Relations</i> , 30, 213-228
PRO28	8669	Nunan, D. (2003) The Impact of English as a Global Language on Educational Policies and Practices in the Asia-Pacific Region, <i>TESOL Quarterly</i> , 37, 589-613
PRO29	8647	Jarvis, H. (2004) Investigating the Classroom Applications of Computers on EFL Courses at Higher Education Institutions in UK, <i>Journal of English for Academic Purpose</i> , 3/2, 111-137
PRO30	8400	Leki, I. (2006) Negotiating Socioacademic Relations: English Learners' Reception by and Reaction to College Faculty, <i>Journal of English for Academic Purpose</i> , 5, 136-152

Appendix 5- 4: A List of the Assignments used in the NES Corpus

Scrip no.	No of words	MA Programme	Module name
NES1	5191	Applied Linguistics	TESOL
NES2	2916	Applied Linguistics	IAL
NES3	5544	Applied Linguistics	IAL
NES4	5793	Applied Linguistics	TESOL
NES5	3493	CCC	CCC
NES6	5247	Applied Linguistics	TESOL
NES7	5582	Applied Linguistics	DA
NES8	3287	Applied Linguistics	TESOL
NES9	2990	CCC	CCC
NES10	2891	Applied Linguistics	SLA
NES11	3099	Applied Linguistics	SLA
NES12	5604	Applied Linguistics	TESOL
NES13	2546	Applied Linguistics	TESOL
NES14	2320	CCC	CCC
NES15	3563	Applied Linguistics	SLA
NES16	4649	CCC	CCC
NES17	7166	CCC	DA
NES18	3379	Applied Linguistics	TESOL
NES19	4886	CCC	CCC
NES20	4530	Applied Linguistics	IAL
NES21	3548	Applied Linguistics	TESOL
NES22	4495	Applied Linguistics	IAL
NES23	3868	Applied Linguistics	SLA
NES24	3108	Applied Linguistics	IAL
NES25	4891	CCC	CCC

(SLA: Second Language Acquisition, TESOL: TESOL Theory and Practice, DA: Discourse Analysis, IAL: Introduction to Applied Linguistics and C.C.C.: Cross Cultural Communication)

A List of Titles of Students' Assignment:

1. Discuss the Possible Benefits and Dangers of Conflict that can Arise out of Conflict in an Educational Setting.
2. Plasticity and the Infant Brain: Recovery from Acquired Aphasia based on the Lenneberg's Hypothesis
3. Bullying amongst Young People
4. The Teaching of Vocabulary is Relegated to the Filling of Slots in Grammatical Frames
5. A Critical Evaluation of the Journal Article Cross Cultural Communication Competence and Multicultural Team Performance
6. A Reflective Discussion on an ELT Lesson with Its Underpinning Theories and Rationale and Practice
7. Written Discourse Analysis Assignment - Cohesion and Coherence
8. Teaching Vocabulary to EFL students
9. Write a Critical Cross-Cultural Evaluation of at least One Recent Empirical Article in Investigating Politeness
10. Justify or Reject the 20th Century Principle 'The Goal of Language Teaching is to Become as Close as Possible to the Native Speaker' and Decide How far It is

Useful for Teaching in the 21st Century in the Light of Current SLA Research and Theories

11. Chinese Learners' Variability in the Use of the English Article System
12. Teaching Vocabulary to EFL students
13. The Potential of Video for Language Teaching
14. Gift-Giving in Hong Kong and Japan: A Comparison
15. Justify or Reject the 20th Century Principle that the L2 should be used in the Classroom, for the 21st Century Classroom
16. How a Human Activity on Socio-Cultural Context Influences Business Negotiation
17. A Spoken Discourse Analysis
18. A Discussion of the Teaching Potential of Video
19. The Culture and Communication in Israel: The Correlation between Culture and Communication (Hall, 1977) is Evident in the Conflict between the Israeli Jews and the Arabs
20. Effective Teaming in Education
21. Investigation of a Lesson to ESL Learners
22. Evaluation of the Key Issues in 'Error Correction'
23. Acquisition and Variability in Past Morpheme Production – Issues Arising from Mandarin Learners of English
24. Exploration of the Means by which *focus* and *ground* are Evident in Modern Turkish
25. What is The Role of Culture in Language Teaching?: Assessing the Reasons for, and Perceived Problems of, Including a Cultural Element in the Language Classroom

Appendix 5- 5: A List of Assignments used in the NNES Corpus:

Script no.	No of words	Country from	MA Programme	Module name
NNES1	5230	Taiwan	Applied Linguistics	IAL
NNES2	4052	Taiwan	CCC	CCC
NNES3	3963	China	Applied Linguistics	IAL
NNES4	4504	Taiwan	CCC	CCC
NNES5	4048	Saudi Arabia	Applied Linguistics	DA
NNES6	3133	China	CCC	CCC
NNES7	5026	Taiwan	CCC	TESOL
NNES8	4696	Japan	Applied Linguistics	DA
NNES9	797	China	Applied Linguistics	IAL
NNES10	4865	China	Applied Linguistics	DA
NNES11	4733	China	Applied Linguistics	CCC
NNES12	3248	Korea	Applied Linguistics	TESOL
NNES13	4381	Taiwan	CCC	DA
NNES14	4141	China	Applied Linguistics	DA
NNES15	2365	China	CCC	CCC
NNES16	4092	Saudi Arabia	Applied Linguistics	DA
NNES17	4707	Japan	Applied Linguistics	DA
NNES18	1906	China	CCC	CCC
NNES19	4562	China	Applied Linguistics	DA
NNES20	1981	Taiwan	Applied Linguistics	TESOL
NNES21	2932	Kuwait	Applied Linguistics	SLA
NNES22	5450	Libya	Applied Linguistics	TESOL
NNES23	2412	Japan	CCC	CCC
NNES24	3350	China	Applied Linguistics	IAL
NNES25	4290	Libya	Applied Linguistics	TESOL
NNES26	4571	Taiwan	CCC	TESOL
NNES27	3871	Taiwan	Applied Linguistics	DA
NNES28	5334	Taiwan	Applied Linguistics	DA
NNES29	5830	Taiwan	Applied Linguistics	DA
NNES30	4388	China	Applied Linguistics	DA

(SLA: Second Language Acquisition, TESOL: TESOL Theory and Practice, DA: Discourse Analysis, IAL: Introduction to Applied Linguistics and C.C.C.: Cross Cultural Communication)

A List of Titles of Students' Assignment:

- 1 A Critique of a Pair of Research Papers
- 2 Advertising does not only Sell Goods but also Identity
- 3 Needs of Analysis of English Majors in China
4. Investigation of the Values, Traditions and Cultural Dimensions that are Shaping the Manner in which Daily Interaction is Handled.
5. Investigation of the Differences in the Use of Cohesive Devices between Authentic Texts and Educational-oriented Texts
6. Language Teaching in Intercultural Context
7. Evaluation of an English Textbook in Taiwan
8. The analysis of Two Fairy Tales

9. A Critical Evaluation of Nicholas Evans' Article "The Last Speaker Is Dead- Long Live the Last Speaker!"
10. Analysis of the Classroom Interaction with Sinclair and Coulthard's triadic Descriptive Framework
11. The Role of Culture in Business Negotiation
12. The Potential of Video for Language Teaching
13. Spoken Discourse Analysis
14. Written Discourse Analysis in terms of Cohesion and Coherence
15. Discussion of the Reasons and Etiquettes of Gift Giving in People's Daily Life
16. Written Discourse Analysis in an ELT setting
17. Classroom Discourse Analysis
18. Is the 'Sustainable Development' an Oxymoron?
19. Spoken Discourse Analysis
20. A Review on Foreign Language Education Provision
21. Input and Development of SLA
22. A Critical Evaluation and Examination of a Sample of Material which Focuses on the Teaching of Vocabulary
23. An Examination of the 'Apologizing', especially the Difference between Japan and U.S
24. A Brief Comparison between English and Chinese: Phonetics on Three Aspects
25. How Should Reading Skills Best be Taught ? Discuss Ideas and Debates, Backed up with Examples and Activities.
26. A Critical Evaluation of a Textbook focuses on the Teaching of Speaking
27. Written Discourse Analysis
28. Classroom Discourse Analysis with IRF Sequence
29. Written Discourse Analysis
30. To Develop Discourse Competence

Appendix 5- 6: Samples of Data Analysis for the Establishment of the Typology of MD

An example of scoring for a professional text from PRO 10

Hinkel, E. (2006) Current Perspectives on Teaching the Four Skills, TESOL Quarterly 40/1 109-131

Sequencers	0
Topicalisers	To make language learning as realistic as possible (4), Based on (16),
Concluders	0
Previews/reviews	0
Additives	Also (12), additionally (14)
Concessives	However (9) (11) (16)
Code glosses	For example (3) (11) (13), For instance (5), such as (7) (14), also called (7), (and this is not a complete list by any measure) (7), e.g. (11) (12) (13) (16), (see Richards & Rodgers, 2001; Widdowson, 1990,2003) (15), see e.g. (16)
Rationales	0
Sources	(Canagarajah, 2002, 2005) (1), Richards and Rodgers (2001) note (8), some leading...(11), (see Richards & Rodgers, 2001; Widdowson, 1990,2003) (15),
Hedges	Commonly (2), probably (10)
Emphatics	Easily (5), in fact (8), certainly (9), simply inappropriate and impractical (12)
Evaluatives	it is safe to say (9)
Commentaries	0
Appeals	0
Self references	0

(Figures in bracket are sentence no.)

Paragraph 11

(1) In many locations around the world, learning English has the objective of learners' gaining access to technical, educational, or professional opportunities (Canagarajah, 2002, 2005). (2) Commonly accepted perspectives on language teaching and learning recognize that, in meaningful communication, people employ incremental language skills not in isolation but in tandem. (3) For example, to engage in a conversation, one needs to be able to speak and comprehend at the same time. (4) To make language learning as realistic as possible, integrated instruction has to address a range of L2 skills simultaneously, all of which are requisite in

communication. (5) For instance, teaching reading can be easily tied to instruction on writing and vocabulary, and oral skills readily lend themselves to teaching pronunciation, listening, and cross-cultural pragmatics (Hinkel, 2001; Lazaraton, 2001; McCarthy & O'Keeffe, 2004).

Paragraph 12

(6) Integrated and multiskill instruction follows the principles of the communicative approach, with various pedagogical emphases, goals, instructional materials, activities, and procedures playing a central role in promoting communicative language use. (7) At present, the models for integrated teaching with a communicative focus include an extensive array of curricula and types of instructional models, such as content based (including theme based), task based, text based (also called genre based), discourse based, project based, problem based, literature based, literacy based, community based, competency based, or standards based (and this is not a complete list by any measure). (8) In fact, Richards and Rodgers (2001) note that, as long as instruction engages learners in meaningful communication and enables them to attain the curricular objectives, the range of models and teaching materials compatible with integrated language teaching is "unlimited" (p. 165).

Paragraph 13

(9) It is safe to say, however, that few movements, in foreign language (FL) and L2 teaching take place without contest, and integrated language instruction is certainly no exception. (10) Currently, task-based and content based instruction are probably among the most widely adopted integrated models. (11) However, some leading

specialists in L2 teaching and applied linguistics have maintained that the superiority of, for example, task-based instruction over traditional teaching has not been demonstrated empirically and that to date research has had little to say about its effectiveness (e.g., Richards & Rodgers, 2001; Seedhouse, 1999; Swan, 2005; Widdowson, 1990, 1993, 2003). **(12)** Critics also contend that in many ESL and EFL situations worldwide, the implementation of content-based and task-based instruction may be simply inappropriate and impractical (e.g., Swan, 2005; Dr, 1996). **(13)** For example, FL or L2 proficiency cannot be developed when learning is limited to 1-3 hours of classroom instruction and input (e.g., Lightbown, 2000; Lightbown & Spada, 1990). **(14)** Additionally, when instruction in content areas, such as science or math, is carried out in English in EFL settings, teachers often find it difficult to maintain expertise in both English and the subject matter, and learners who need to prepare for examinations often concentrate only on school subjects without much interest in learning the language. **(15)** In task-based, multiskill instruction, with its focus on the development of language fluency, issues of content or linguistic accuracy are of secondary importance, thus limiting the usefulness of the task-based model for schooling and academic preparation (see Richards & Rodgers, 2001; Widdowson, 1990,2003). **(16)** Based on their experience, however, many L2 teachers and curriculum designers believe that integrated FL/L2 instruction can increase learners' opportunities for L2 purposeful communication, interaction, real-life language use, and diverse types of contextualized discourse and linguistic features, all of which have the goal of developing students' language proficiency and skills (for detailed discussion, see, e.g., Ellis, 2003; Fotos, 2001, 2002; Snow, 2005).

An example of scoring for the non-native English speaker's writing from

NNES 12 –

Sequencers	0
Topicalisers	In relation to this (10),
Concluders	0
Previews/reviews	As mentioned above (9), this paper (14)
Additives	0
Concessives	However (6)(12), whereas (11)
Code glosses	For example(6)(12)
Rationales	0
Sources	(Singleton, 1989) (1), (Ellis, 1995) (2) (5), According to this hypothesis (4), (Lennenberg, 1967) (5), (Singleton, 1989, Flege, 1987 and Patkowski, 1990) (6), some studies in (7), the work done in ... (8), (Fathman, 1975) (8) (9), (Bongarerts et al., 1989) (11), some research has suggested (12), Swain, 1981 and Fathman, 1975 (12), (Singleton, 1989) (13)
Hedges	0
Emphatics	Constantly (1), it is a commonly (10),
Evaluatives	0
Commentaries	0
Appeals	0
Self references	0

(Figures in bracket are sentence no.)

Paragraph 1

(1) The age factor has been a constantly recurring theme in the great expansion of language acquisition studies of the last few decades (Singleton, 1989). (2) There is a belief that younger L2 learners generally do better than older learners in learning language (Ellis, 1995). (3) This is supported by the 'critical period hypothesis' (Lennenberg, 1967). (4) According to this hypothesis, the optimum period for language acquisition falls within the first ten years of life, when the brain retains its plasticity. (5) This period has been equated with the period required for lateralisation of the language function to the left side of the brain (Ellis, 1995). (6) However, the particular issue of the effect of the learner's age on the acquisition of a second language

(L2) is one that remains controversial. (For example, Singleton, 1989, Flege, 1987 and Patkowski, 1990).

Paragraph 2

(7) Some studies in second language learning (LL2) have attempted to identify similarities between first and second language acquisition. (8) The work done in first language research has provided new perspectives for research dealing with L2 acquisition processes (Fathman, 1975). (9) As mentioned above, it has been noted that there is limited ability to learn a first language (L1) after a certain age and it has also been noted that young children possess a special ability to learn a L2 (Fathman, 1975). (10) In relation to this, it is a commonly held belief that children are very successful L2 learners compared to adults. (11) After being exposed to another language, children seem to pick up the new language without much effort, whereas their parents experience great difficulty in achieving L2 proficiency (Bongarerts et al., 1989). (12) However, some research has suggested that the young children are not necessarily better L2 learners than adults (For example, Swain, 1981 and Fathman, 1975). (13) From these studies it appears that the principle criteria used to test the role of age in L2 learning have been rate of acquisition and ultimate success (Singleton, 1989).

Paragraph 3

(14) Thus, this paper examines and discusses theory and research concerning the role of age in determining L2 learning rate and success, in particular, between children of various ages and adults, with reference to the experiences of myself and my friends as L2 learners.

An example of scoring for the native English speaker's writing from NES 18

Sequencers	0
Topicalisers	As Family Fortunes is (1)
Concluders	0
Previews/reviews	(line 32, 53) (4), on line 19 (5), an example is given in appendix (7), (line34) (11), shown in appendix (13), on line 25 (9), above (13)
Additives	0
Concessives	However (2)
Code glosses	e.g. (2), (see....)(2), (line 32,53) (4), an example is given in appendix (7), (especially when...)(6), (i.e.....) (10), (line34) (11), (compiling all....) (13)
Rationales	Because (10), because of (12) (13)
Sources	(Higa 1963) (2)
Hedges	Might be (4)
Emphatics	The most obvious (1), that is already quite well known (3),
Evaluatives	Worth highlighting(4), there are something interesting (4), another interesting feature(5),
Commentaries	0
Appeals	0
Self references	I (8)

(Figures in bracket are sentence no.)

Paragraph 14

(1) As Family Fortunes is itself a kind of vocabulary game, the most obvious vocabulary application of the extract is to use it as a model, and then play the game with the students, using questions that naturally elicit lexical sets that the students have learned in previous lessons. (2) However, a considerable body of evidence has built up to indicate that words with semantic similarities (e.g. parts of the body, numbers, days of the week) should not be learned together, as to do so will create an interference problem (see Higa, 1963; Tinkham, 1997; Nation, 2000). (3) Thus, current research suggests that the game should only be played to recycle vocabulary that is already quite well known, and not as part of the teaching of new words.

Paragraph 15

(4) There are some interesting “chunks” of vocabulary which might be worth highlighting in class, such as “I’m not gonna go there” (line 32) and “might as well” (line 53). (5) Another interesting feature can be found on line 19, when Les uses “yeah” as a question tag. (6) Students who have grappled despairingly with English question tags for years will delight in discovering that in some contexts (especially when genuinely seeking factual information) we can also use “yeah” or “yes”. (7) An example is given in appendix E.

Paragraph 16

(8) The last vocabulary feature I wish to discuss is the subordinate and superordinate relationships between lexical items. (9) On line 25, the contestant says, “sharks”. (10) This is accepted as a correct answer because the word shark has a subordinate relationship to the word “fish”. (i.e. a shark is a type of fish.) (11) Something similar happens a few seconds later when a contestant says, “ships” (line34). (12) This is accepted because of the relationship between “ships” and “boats”. (13) Students could be encouraged to keep records of subordinate relationships using diagrams like those shown in appendix E. (compiling all the words in a diagram at the same time would be a bad idea, because of the interference factor discussed above.)

Appendix 5- 7: List of Potential MDs to be Searched for in the Corpora

Textual Metadiscourse:

For TMD, items referring to the text to organise and classify for the reader, the following search items were used:

Sequencers: First, firstly, finally, first of all, lastly, listing (a,b,c,), next, second, secondly, subsequently, third, thirdly

Topicalisers: In regard to/regarding, based on, move on, turn to, with regard to, as for, in relation to/respect to, bearing in mind

Concluders: In conclusion, in sum, in summary, on the whole, overall, so far, thus far, to conclude, to sum up, to summarize

Pre/reviews: Previously/earlier, (in) part X, (in) (the) section X, (in) this chapter, (in) this part, (in) this section, page X, figure X, table X, above, below, this work/paper

Additives: Accordingly, also, furthermore, in addition/additionally, in the same way, moreover, therefore, likewise

Concessives: Although, but, by contrast, however, in spite of, on the contrary, on the other hand, whereas, yet

Code Glosses: Defined as, e.g., for example, for instance, I mean, i.e., in other words, known as, namely, such as, that means, which means

Rationales: The reason for this, as a result of, because (of), as a consequence/consequently, thereby

Sources: According to X, X suggests, X argues, X notes, X states, X proposes

Interpersonal Metadiscourse:

The search items for IMD, used to interact with the reader about the propositional content, are:

Hedges: Approximately, as far as, assumed, broadly, might, perhaps, possibly, presumably, probably, seems, tend to, uncertain, unclear

Emphatics: Actually, already, always, certainly, clearly, commonly, constantly, definitely, doubtless, easily, especially, in fact, no doubt, obviously (obvious), of course, really, simply, surely, truly, undeniably, undoubtedly

Evaluatives: Admittedly, appropriately, astonishingly, correctly, desirable, desirably, disappointingly, dramatically, equally important, more importantly, essentially, fortunately, hopefully, importantly, inappropriately, interestingly, it is safe to say, meaningful, preferably, remarkably, strikingly, surprisingly, understandably, unexpectedly, unfortunately, unusual, worth, worthy,

Commentaries: Allow me, you might be missing, you might add, consider (you might), you might be call, imagine, you will, (the) reader's, let us, you, your

Appeals: What should I ignore at this time?, How should I respond?, Will they learn from my commentary for future writing?, Will each student understand?, Do you...?, Does it....?, How does it...?/how do people....?, What circumstances....?, (When) is it....?, Do we....?

Self references: I, we, me, my, our, us, author, researcher

Appendix 5- 8: Samples of Counting the MD Frequency for 'Below' in the Subcategory of Pre/Reviews

After collecting all the instances of all the MD features, each feature was searched separately. To illustrate how MD frequency was established for each feature, we can take the case of 'below'.

In the NES data; a search of the NES corpus yielded 11 instances of 'below'. These were examined by the researcher according to the working definition of MDs in Chapters 3 and 4 and those that did not fit were excluded from the count.

The screenshot shows the Concordance software window with a menu bar (File, Edit, View, Compute, Settings, Windows, Help) and a toolbar. The main window displays a table with three columns: 'Concordance', a central column with vertical bars representing the search term 'below', and 'File'. The table lists 11 instances of the word 'below' from various files in the NES data corpus. The bottom of the window shows a taskbar with the Start button, system tray icons, and the time 6:34 PM.

Concordance		File
1 quining the article system. The example below illustrates this. Line 10) A: Actual		orrection\nes data\11 data\nes 11.doc
2 These preceding statements (elaborated below) underscore the analytical framew		orrection\nes data\11 data\nes 17.doc
3 4, p.49) Gropper found that "students of below average ability also profited most fr		orrection\nes data\11 data\nes 18.doc
4 which\af12\loch\ team. Listed below are the eight roles with a brief des		\correction\nes data\11 data\nes 20.rtf
5 T) with the student's Interlanguage (see below), though not with\rtch\cs1 \af\		\correction\nes data\11 data\nes 22.rtf
6 giving the assignment? As will be seen below in the correction of the\rtch\cs1		\correction\nes data\11 data\nes 22.rtf
7 aggregate results are shown in Table 1 below (see appendix I for individually tab		orrection\nes data\11 data\nes 23.doc
8 d telic. The results are shown in Table 2 below (individually tabulated in appendix I		orrection\nes data\11 data\nes 23.doc
9 Results The results are given in Table 4 below (see appendix II for individually tab		orrection\nes data\11 data\nes 23.doc
10 le to replace the underlined section here below. \par \par \af12\loch\		2\correction\nes data\11 data\nes 4.rtf
11 option B, and has been drawn in the box below. \par \pard \trpar\q\ \i\i\knowidc		2\correction\nes data\11 data\nes 7.rtf

This chart shows a typical screen-shot from WordSmith for 'below'. The 'file' in the right hand column shows the text number for each example of 'below' in the NES data. The left hand side (Concordance, one-item-per-search program) gives a KWIC (Key Words in Context) display set for 40 characters before and after the search item.

In order to weed out irrelevant occurrences and ensure that the word is functioning as MDs, the context was expanded to include a larger context up to paragraph level before or after the word being searched for, as we see here.

Concordance

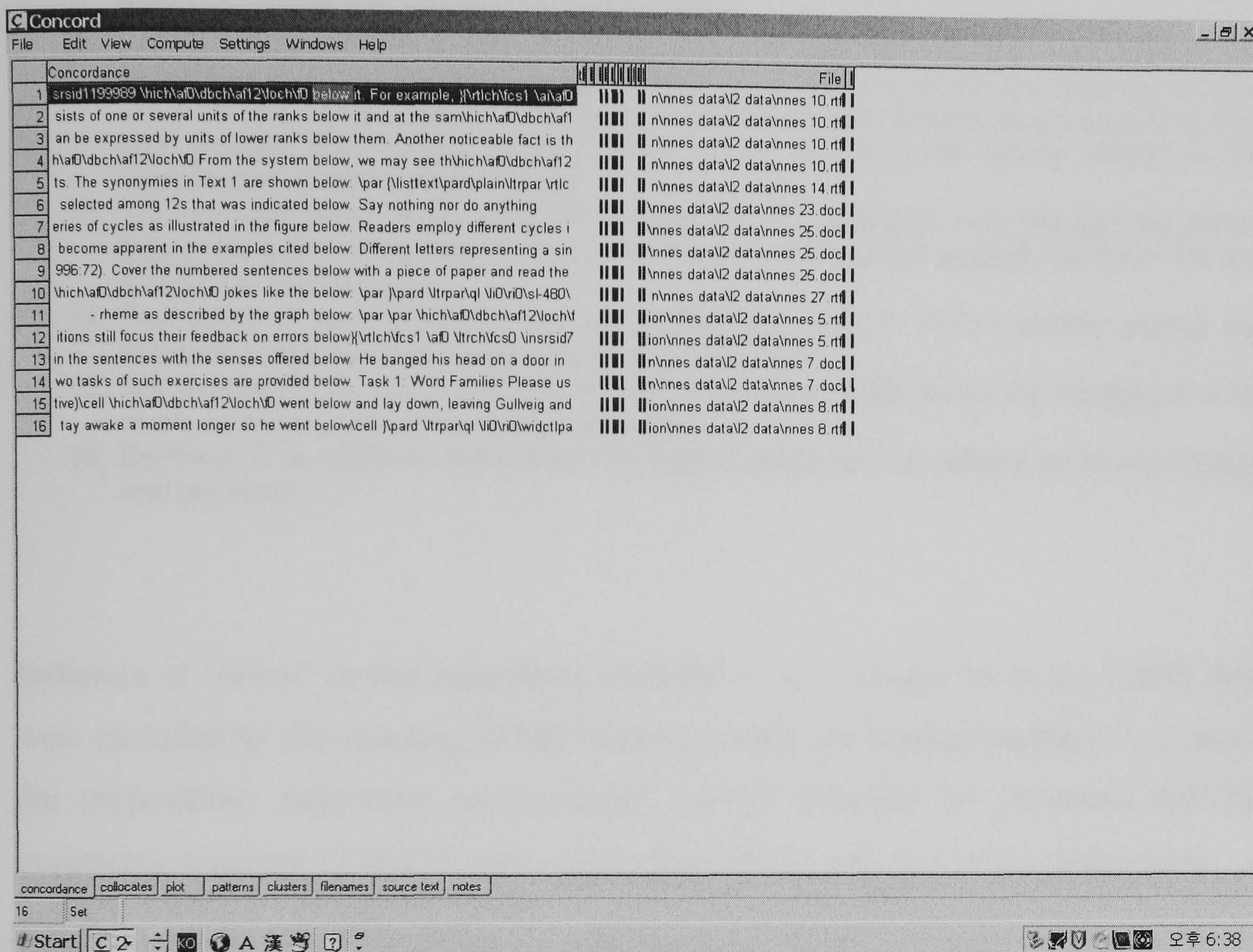
- 1 ...does illustrate that the speaker is still in the developmental process of acquiring the article system. The example below illustrates this. Line 10) A: Actually, the arrow is just line, you know. Line 11) A: Ok, under the arrow, there
- 2 ...prosodic intonation, and turn-taking mechanics, just to name a few. These preceding statements (elaborated below) underscore the analytical framework employed in the current paper. It is Important to note that the....
- 3 the film compared to the other two instructional methods" (Wetzel et ai, 1994,p.49). Gropper found that "students of below average ability also profited most from the visual information". There is admittedly a generalisability issue with...
- 4believes there are eight roles in order to fonn a productive team. Listed below are the eight roles with a brief description of their traits.
- 5 'Transfer Analysis' incorporates the comparison of Mother Tongue (MT) with the student's Interlanguage (see below) though not with the Target Language. In its various guises, this adaptation has gained some
- 6 What points are stressed by the teacher prior to giving the assignment? As will be seen below in the correction of the stress during teacher presentation can directly impact writing output as students seek to meet.....
- 7 ...testing for production of [t] and [d] in word-initial, medial and final positions. The aggregate results are shown in Table 1 below (see appendix I for individually tabulated results). Table 1: % production of voiced and unvoiced [+ coronal]
- 8 I concluded that a non-target choice on perfective or imperfective would indicate support for assuming a correlation between perfective and Atelic. The results are shown in Table 2 below (individually tabulated in appendix). Table 2: % non-target correlation of lexical and grammatical aspect Atelic
- 9 ...to participants' morphological and phonological access. The results are given in Table 4 below (see appendix II for individually tabulated results, identification of target and non-target forms and transcripts).
- 10 ...so that today's students, who will be the custodians of our language, will be able to replace the underlined section here below. The advent of a mass audience, coupled with the advent of a naturalistic style, may dangerously invite a reduction of language.....
- 11 Lines 33-39 culd be said to contain another pattern, one which I have not found described in the literature. It is a kind of mirror image of McCarthy's option B, and has been drawn in the box below.

Most of these occurrences of '*below*' in the NES data were indeed functioning as MDs according to the definition of pre/review (p.98). For example occurrence numbers 1 and 2 clearly fit the definition of MD as pre/reviews while occurrence numbers 4,5,6,7,8,9,10 and 11 are clearly MD by definition of pre/reviews (i.e. they alert the reader as to what the writer is going to do in the following sentence). However the instance in the occurrence no 3 was used to expand the propositional information ('below average') of a text as a primary discourse. It is therefore excluded from the count. Details of the results of the examination in each occurrence from the extended sentences are as follows:

Instance no.	Functioning as MD	Functioning as non-MD	Text no.
1	✓		NES11
2	✓		NES17
3		✓	NES18
4	✓		NES20
5	✓		NES22
6	✓		NES22
7	✓		NES23
8	✓		NES23
9	✓		NES23
10	✓		NES4
11	✓		NES7

Thus there are 10 occurrences of '*below*' functioning as MDs (subcategory of the pre/reviews) and 1 instance of '*below*' which has a propositional function (non-MD) in the NES texts.

In the NNES data, there are 16 instances of 'below'. All of the instances were examined according to the definition of pre/reviews and those that did not fit were excluded for the counting of MD frequency.



The extended contexts for each instance were as follows:

Concordance

1. In their rank scale model, a unit is given a certain rank. Except for the lowest rank, each rank can be realized in forms of units below it. For example,....
2. Each group consists of one or several units of the ranks below it and at the same time one or more units at group rank themselves are made up
3. We may see that higher ranks in each level can be expressed by units of lower ranks below them.....
4. From the system below, we may see that the largest unit lesson has an equal status with paragraph at the level of grammar because it fails to be expressed in a form of structure..
5. Cohesion using similar or near similar terms is seen throughout the texts. The synonymies in Text 1 are shown below:
6. They are consists of 12 scenarios which are given according to the difference of seriousness. Easier one is 'caused a traffic accident by lent-a-car, and got a slight scratch on the surface', and the heavy one is 'caused traffic accident, and gave a serious injure to a walker'. The most suitable answer was selected among 12s that was indicated below.
7. This can be usefully represented as a series of cycles as illustrated in the figure below. Readers employ different cycles.....
8. The inconsistencies and irregularities become apparent in the examples cited below: Different letters representing a single sound:

9. ...From the context as illustrated in the following example from Nuttall (1996:72). Cover the numbered sentences below with a piece of paper and read them one by one. After you have read the first one, make a note of what...
10. Activity I: the aim is to exercise and practise student's coherence. Providing two jokes like the below: (the following jokes are adapted from the web site.....
11. The first paragraph has the simple construct of theme-rheme followed by another theme-rheme as described by the graph below.
12. Unfortunately, many ESL/EFL teachers upon receiving their students' compositions still focus their feedback on errors below, clause-level, e.g. lack of subject-verb agreement, incorrect use of verb tenses, misspelled words and alike. Discourse analysis therefore can broaden this narrow scope and move the focus to the issues related to the organisation of text and coherence.
13. Try to match head or the phrases in italics in the sentences with the senses offered below. He banged his head on a door in the dark. The snake erected the head. He has a head for mathematics...
14. Two tasks of such exercises are provided below. Task 1: Word Families, please use this as the root word and enter its correct derivatives.
15. But eventually he could not stay awake a moment longer so he went below and lay down, leaving Gullveig and Agnar on deck.
16. Sentence 6 is negative evaluation that Sigurd could not stay awake so he went below and lay down.

Instances of '*below*' in the occurrence numbers 1, 2, 3, 15 and 16 in the NNES data were excluded for the counting of MD frequency since the instance numbers 1, 2, and 3 are prepositions expressing propositional content followed by pronouns and the occurrence numbers 15 and 16 were collocations of the verb 'go' with a preposition.

Instance no.	Functioning as MD	Functioning as non-MD	Text no.
1		✓	NNES10
2		✓	NNES10
3		✓	NNES10
4	✓		NNES10
5	✓		NNES14
6	✓		NNES23
7	✓		NNES25
8	✓		NNES25
9	✓		NNES25
10	✓		NNES27
11	✓		NNES5
12	✓		NNES5
13	✓		NNES7
14	✓		NNES7
15		✓	NNES8
16		✓	NNES8

Thus there are 11 instances of 'below' functioning as MDs and 5 instances as non-MDs in the NNES data. Included occurrences, such as 4,5,6,7,8,9,10,11,12,13, and 14 therefore fit the definition pre/reviews.

The excluded sentences of '*below*' (non-MDs) are as follows:

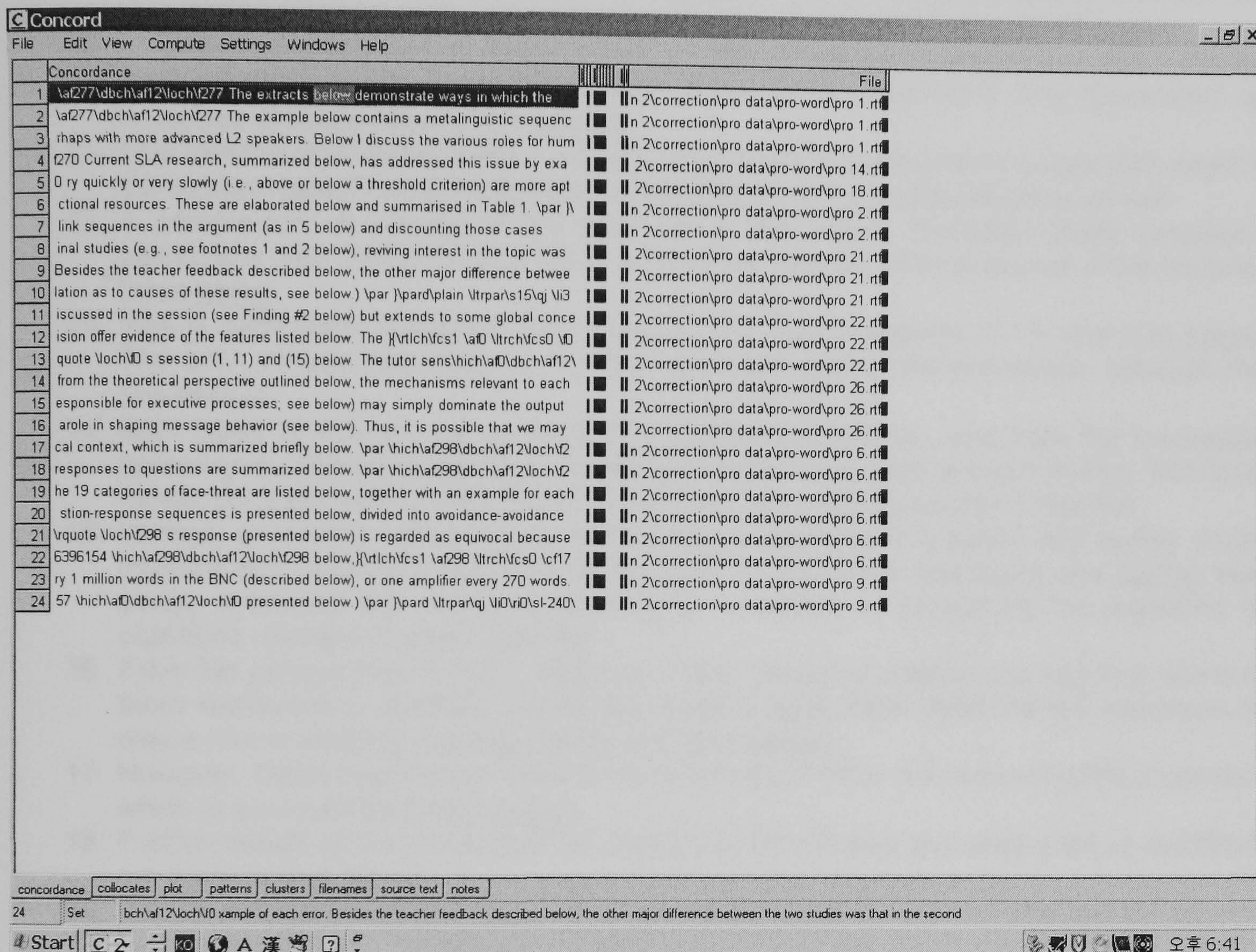
Sentence no. 1- Except for the lowest rank, each rank can be realized in forms of units below it
Sentence no. 2- Each group consists of one or several units of the ranks below it.

Sentence no. 3 - We may see that higher ranks in each level can be expressed by units of lower ranks below them

Sentence no. 15 - But eventually he could not stay awake a moment longer so he went below and lay down, leaving Gullveig and Agnar on deck

Sentence no. 16 - Sentence 6 is negative evaluation that Sigurd could not stay awake so he went below and lay down

In the PRO data; there are 24 instances of 'below' and all the instances were checked according to the criteria of the subcategory (pre/reviews) to ensure that they are functioning as MDs and those that did not fit were excluded from the counting of MD frequency.



Examples of context with the searched feature 'below' are as follows for the checking of MD frequency:

Concordance

1. The extracts below demonstrate ways in which the three participants used voicing and manipulated L2 resources in humorous conversation with varying degrees of success.
2. The example below contains a metalinguistics sequence in which the NS interlocutor is the one who initiates the lengthy metalinguistic discussion.
3. Below I discuss the various roles for humorous L2 language play that arose from the data presented here and suggest how future research might begin to determine whether or to what extent language play actually....
4. Current SLA research, summarized below, has addressed this issue by examining the extent to which individual differences (IDs) in cognitive abilities..
5. But those who speak very quickly or very slowly (i.e., above or below a threshold criterion) are more apt to be perceived in a negative fashion.
6. These are elaborated below and summarised in Table 1. They are concerned with ways of organising discourse to....

7. Those cases where they are being used to link sequences in the argument (as in 5 below) and discounting those cases where they are being used to express relations between processes (as in 6):..
8. Although the conclusions Truscott drew from the findings of previous research were sometimes unsupported by the data in the original studies (e.g., see footnotes 1 and 2 below), reviving interest in the topic was beneficial since teacher...
9. Each student received a sheet listing the abbreviations and an example of each error. Besides the teacher feedback described below, the other major difference between the two studies was that in the second group students were asked to write 40 pages,...
10. ... underlining resulted in more errors on the subsequent assignment even though students made fewer errors on revisions of the same assignment. (For speculation as to causes of these results, see below).
11. This trend is most easily seen in writer's responses to surface-level issues discussed in the session (see Finding #2 below) but extends to some global concerns as well.
12. A comparison of several of the sessions in the corpus illustrates these conditions particularly well. Sessions that led to substantial revision offer evidence of the features listed below.
13. This is particularly evident in the excerpts from Abby's session (1,11) and (15) below. The tutor senses that Abby is still having trouble seeing the connection between the two essays...
14. Both types of thought are governed by ideational dynamics, and from the theoretical perspective outlined below, the mechanisms relevant to each are very similar, but here, I shall focus primarily on the former, what I will call 'message-content ideation'.
15. Instead, according to the theory, decrements in performance speed and quality under multiple-task (including coalitions responsible for executive processes; see below) may simply dominate the output representation, impeding or preventing the assembly of coalitions relevant to other activities.
16. From the perspective of AAT2, however, many coalitions exist at any moment, some of them representing abstract conceptual content, and these need not be conscious to play a role in shaping message behaviour (see below).
17. However, these hypotheses have to be understood within the relevant political context, which is summarized briefly below.
18. Further details of the procedures or identifying face-threats and responses to questions are summarized below.
19. The 19 categories of face-threat are listed below, together with an example for each of the 3 superordinate categories of face-threat; the full typology is presented elsewhere (Bull et al., 1996).
20. A detailed analysis of the 17 question-response sequences is presented below, divided into avoidance-avoidance conflict and no necessary threat questions.
21. Blair's response (presented below) is regarded as equivocal because he does not say whether he considered most of what old Labour stood for was wrong.
22. The content of his responses can be summarized in a series of propositions, which are presented below, together with a consideration of their significance in terms of face management.
23. Together, in total, they occur about 3,700 times in every 1 million words in the BNC (described below), or one amplifier every 270 words. Both maximizers and boosters are open-class words,
24. (Inevitable errors in automatic tagging and retrieval, and revisions in the tagging program used in the later releases of the corpus, do not have a substantial effect on the results but may account for a possible small margin of error in the rank orderings and MI scores presented below.)

'Below' in the instance no. 5 ('above or below a threshold criteria') was excluded for the counting of MD frequency because it was used as a preposition.

Instance no.	Functioning as MD	Functioning as non-MD	Text no.
1	✓		PRO1
2	✓		PRO1
3	✓		PRO1
4	✓		PRO14
5		✓	PRO18
6	✓		PRO2
7	✓		PRO2
8	✓		PRO21
9	✓		PRO21
10	✓		PRO21
11	✓		PRO22
12	✓		PRO22
13	✓		PRO22
14	✓		PRO26
15	✓		PRO26
16	✓		PRO26
17	✓		PRO6
18	✓		PRO6
19	✓		PRO6
20	✓		PRO6
21	✓		PRO6
22	✓		PRO6
23	✓		PRO9
24	✓		PRO9

Thus there are 23 'below's functioning as MDs and 1 'below' functioning as a non-MD in the PRO corpus. A typical MD use is occurrence numbers 1-4 and 6-24 which fit the definition of pre/reviews. Occurrence no 5 was excluded because it is a preposition followed by noun phrase and expresses propositional content.

Appendix 5- 9: Parametric versus Nonparametric Tests

Most of the statistical tests are based on the assumption that the data are normally distributed. The parameters of this distribution, the mean and variance, are estimated. Since parametric tests are based on the assumption that we know certain characteristics of the population from which the sample is drawn, they are called parametric tests. Thus, parametric statistics deals with the estimation of parameters (e.g., means, variances) and the test of hypotheses regarding parameter values – usually under the assumption that variables are normally distributed (which is the form of a symmetric bell-shaped curve).

According to Bryman and Cramer (2001) and Kerr, Kozub and Hall (2002), it is only appropriate to use parametric tests when the data fulfil the following three conditions:

1. the level or scale of measurement is of equal interval or ratio scaling, that is, more than ordinal;⁴
2. the distribution of the population is normal; and
3. the variances in the different groups are equal or homogeneous (also see Salkind 2000).⁵

In addition to this, Bryman and Cramer (2001) suggested that parametric tests should not be applied where both the size of the samples and the variances are unequal and also where distributions of scores are non-normal, and that nonparametric methods should be preferred in those cases. It may also be more desirable to use nonparametric tests when the size of the samples is small.

In cases where the assumption of normality cannot be employed, nonparametric methods may be appropriate. Nonparametric methods do not rely on the estimation of

⁴ All data are generated by one of four scales of measurement: nominal, ordinal, interval, and ratio. Hence, all statistical analyses are conducted with nominal, ordinal, interval, or ratio data. The scale of measurement is nominal if the data is simply labels or categories used to define an attribute of an element. Nominal data may be numeric or nonnumeric. The scale of measurement is ordinal if the data can be used to rank, or order, the observations. Ordinal data may be numeric or nonnumeric. The scale of measurement is interval if the data have the properties of ordinal data and the interval between observations is expressed in terms of a fixed unit of measure. Interval data must be numeric. The scale of measurement is ratio if the data have the properties of interval data and the ratio of measures is meaningful. Ratio data must be numeric.

⁵ The Levene test is extensively used to test homogeneity of variances because it is fairly robust to departures from normality. Often when variances are unequal, distributions also are skewed.

parameters (such as the mean or the variance) describing the distribution of the variables of interest in the population and for this reason these methods are also called distribution-free methods. In general, a statistical method is nonparametric if it satisfies at least one of the following criteria.

1. The method may be used on data with a nominal scale of measurement.
2. The method may be used on data with an ordinal scale of measurement.
3. The method may be used on data with an interval or ratio scale of measurement, when no assumption can be made about the population probability distribution.

Hollander and Wolfe (1999) stated that the rapid and continuous development of nonparametric statistical procedures over the past several decades is due to the following advantages enjoyed by nonparametric techniques:

1. Nonparametric methods require few assumptions about the underlying population from which the data are obtained. In particular, nonparametric procedures forgo the traditional assumption that the underlying populations are normal.
2. Nonparametric techniques are often (although not always) easier to apply than their normal theory counterparts.
3. Nonparametric procedures are often quite easy to understand.
4. Although at first glance most nonparametric procedures seem to sacrifice too much of the basic information in the samples, theoretical investigations have shown that this is not usually important. In practice, the nonparametric procedures are only slightly less efficient than their normal theory competitors when the underlying populations are normal, and they can be sometimes be significantly more efficient than these competitors when the underlying populations are not normal.

Another factor that often limits the applicability of tests based on the assumption that the sampling distribution is normal is the size of the sample of the available data for the analysis. It can be assumed that the sampling distribution is normal even if we are not sure that the distribution of the variable in the population is normal, as long as the sample size is large enough (e.g., 100 or more observations). For example, Conover

(1999)⁶ stated that some parametric tests, such as the one-sample t test or the two-sample t-test, are known to be robust against the assumption of normality, especially if the sample sizes are reasonably large. Tests which are able to withstand some violations regarding the normality assumption are described as being ‘robust’ (Bryman and Cramer 2001). If the sample size is very small, however, then those parametric tests can be used only if we are sure that the variable is normally distributed.

The parametric tests were not employed in the present study since the present sample size (below 30 in each group) is not large enough to employ them and the size of the samples is unequal (25 for NES, 30 for NNES and 30 for PRO).

Testing for Normality

As described above, most parametric methods are based on the normality assumption because the theory behind the test can be worked out with normal population distribution. The resulting procedures are efficient and powerful for normally distributed data. Before using a test based on the normality assumption, however, it is necessary to take a preliminary step to detect normality of the data. A variety of statistics for testing normality are available including the Kolmogorov-Smirnov test, the Shapiro-Wilk test, the Stephens’ test, and the Lilliefors test. The Shapiro-Wilk test is recommended for small and medium samples up to $n = 2000$. For larger samples, the Kolmogorov-Smirnov test is recommended (Rahman and Govindarajulu, 1997). Hence the Shapiro-Wilk test was employed for testing the normality of each subcategory of MD in the different groups since the size of scripts in the present study is below 30 for each group.

The results of the Shapiro-Wilk tests are summarised in the appendix (see appendix 5-10). SPSS Version 14, was used to make these tests. For comparison, the tests were carried out for the whole sample, the students vs the PRO, the NNES vs the NES, the NES vs the PRO and the NNES vs the PRO corpus, respectively. The Shapiro-Wilk tests reject the hypothesis of normality for most of the variables in the present study. The normal distribution assumption cannot be excluded only in a minority of variables.

⁶. Conover (1999) pointed out that the most important reason for preferring nonparametric methods is that they are often more powerful than the parametric methods if the assumptions behind the parametric model are not true.

For those variables which are assumed to be normally distributed, parametric independent t-tests were applied to test for the differences between the two groups. The same results were found with nonparametric Mann Whitney U tests.

It can be observed that the distributions of each variable (sub-category of MD in each group) were not normally distributed from the Shapiro-Wilk test which is normality test. Hence non-parametric tests (Mann Whitney U tests) were employed to examine the statistical differences between the two groups.

Appendix 5- 10: Shapiro-Wilk Tests:

Shapiro-Wilk Sig - if this is bigger than 0.05 then the data is normally distributed

Tests of Normality - Student vs Pro

StuPro		Kolmogorov-Smirnov(a)			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Sequencers	Student	.170	54	.000	.782	54	.000
	PRO	.104	30	.200(*)	.947	30	.139
Topicalisers	Student	.209	54	.000	.765	54	.000
	PRO	.299	30	.000	.596	30	.000
Concluders	Student	.227	54	.000	.703	54	.000
	PRO	.231	30	.000	.740	30	.000
PreReviews	Student	.172	54	.000	.864	54	.000
	PRO	.165	30	.036	.926	30	.038
Additives	Student	.125	54	.034	.919	54	.001
	PRO	.127	30	.200(*)	.950	30	.170
Concessives	Student	.068	54	.200(*)	.937	54	.007
	PRO	.113	30	.200(*)	.913	30	.018
CodeGlosses	Student	.130	54	.024	.904	54	.000
	PRO	.224	30	.000	.761	30	.000
Rationales	Student	.187	54	.000	.830	54	.000
	PRO	.087	30	.200(*)	.971	30	.576
Sources	Student	.093	54	.200(*)	.951	54	.027
	PRO	.166	30	.033	.882	30	.003
Hedges	Student	.132	54	.021	.924	54	.002
	PRO	.162	30	.043	.868	30	.002
Emphatics	Student	.073	54	.200(*)	.971	54	.207
	PRO	.146	30	.105	.926	30	.039
Evaluatives	Student	.144	54	.007	.882	54	.000
	PRO	.139	30	.147	.899	30	.008
Commentaries	Student	.289	54	.000	.592	54	.000
	PRO	.381	30	.000	.553	30	.000
Appeals	Student	.450	54	.000	.464	54	.000
	PRO	.394	30	.000	.670	30	.000
IllocutionaryMarkers	Student	.179	54	.000	.802	54	.000
	PRO	.257	30	.000	.682	30	.000

* This is a lower bound of the true significance.

a Lilliefors Significance Correction

Tests of Normality - NES vs NNES

NES vs NNES		Kolmogorov-Smirnov(a)			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Sequencers	NES	.182	25	.032	.792	25	.000
	NNES	.180	29	.017	.842	29	.001
Topicalisers	NES	.182	25	.031	.828	25	.001
	NNES	.215	29	.001	.753	29	.000
Concluders	NES	.150	25	.149	.884	25	.009
	NNES	.262	29	.000	.659	29	.000
PreReviews	NES	.239	25	.001	.788	25	.000
	NNES	.129	29	.200(*)	.915	29	.023
Additives	NES	.110	25	.200(*)	.961	25	.428
	NNES	.159	29	.059	.916	29	.024
Concessives	NES	.139	25	.200(*)	.956	25	.335
	NNES	.134	29	.194	.899	29	.009
CodeGlosses	NES	.199	25	.012	.896	25	.015
	NNES	.112	29	.200(*)	.930	29	.056
Rationales	NES	.201	25	.010	.815	25	.000
	NNES	.187	29	.011	.856	29	.001
Sources	NES	.158	25	.106	.920	25	.052
	NNES	.073	29	.200(*)	.979	29	.804
Hedges	NES	.192	25	.018	.881	25	.007
	NNES	.104	29	.200(*)	.948	29	.166
Emphatics	NES	.129	25	.200(*)	.953	25	.290
	NNES	.119	29	.200(*)	.957	29	.278
Evaluatives	NES	.128	25	.200(*)	.889	25	.011
	NNES	.161	29	.052	.904	29	.012
Commentaries	NES	.245	25	.000	.814	25	.000
	NNES	.305	29	.000	.572	29	.000
Appeals	NES	.448	25	.000	.561	25	.000
	NNES	.458	29	.000	.388	29	.000
IllocutionaryMarkers	NES	.140	25	.200(*)	.919	25	.048
	NNES	.247	29	.000	.660	29	.000

* This is a lower bound of the true significance.

a Lilliefors Significance Correction

Appendix 5- 11: Informed Consent Form

Statement of Informed Consent to Access Data

I, the undersigned, understand the following:

- That I authorise Seowon Lee to access my assignment for her study entitled, ‘The use of metadiscourse in professionals’ and students’ writings’, and the purpose of this research is to examine the certain features of the writing style of different groups of people in academic contexts.
- That my assignment will be examined anonymously or under a pseudonym, and no record will be kept of my name.
- That the data will be accessible only to Seowon Lee, will be used only for linguistic analysis, and I may contact Seowon Lee at seowon.lee@ncl.ac.uk if I have any questions or concerns relating to this study.

By signing below, I certify that I have read and understood the foregoing terms and conditions, and that I agree to authorise, in accordance with them, in the above-named study.

Student Number and Name

Signature

Date

Appendix 5- 12: Ethical Issues from British Association for Applied Linguistics

Responsibilities to Informants

Responsibilities and relations with informants will sometimes vary according to the type of inquiry carried out, and occasionally the dilemmas and tensions between, for example, confidentiality and the public's right to know, or between anonymity and the safety of other people, will need to be negotiated case by case. The points below generally apply to all Informants, whatever their social position, but particular care needs to be taken with those who have less power to negotiate their rights.

General Responsibility to Informants

Applied linguists should respect the rights, interests, sensitivities, and privacy of their informants. It is important to try to anticipate any harmful effects or disruptions to informants' lives and environment, and to avoid any stress, undue intrusion, and real or perceived exploitation. Researchers have a responsibility to be sensitive to cultural, religious, gender, age and other differences: when trying to assess the potential impact of their work, they may need to seek guidance from members of the informants' own communities. In certain types of contract research, respect for informants cannot be guaranteed, and in these cases, researchers should consider carefully whether they should continue with the project.

Obtaining Informed Consent

Relationships with informants should be founded on trust and openness. They should be informed about all aspects of research that might reasonably be expected to affect their willingness to participate. The information given to efforts at the outset of a project should cover the objectives of the research, its possible consequences, and issues of confidentiality and data security. When informants differ from the researcher in the social groups they belong to, it is worth seeking guidance on social, cultural, religious and other practices which might affect relationships and the willingness to participate. In cases where the research continues over a long period, the informed consent obtained at the start of the project may no longer be adequate, and consent may need to be renegotiated. Researchers should try to obtain the real consent of children and of adults with impairments in understanding. When children under sixteen are

acting as main informants, it is also necessary to obtain the consent of parents or other adults acting *in loco parentis*.

Respecting a Person's Decision Not to Participate

Informants have a right to refuse to participate in research. But applied linguists need to be aware that the power relations between themselves and their potential informants can sometimes be inadvertently misused to pressurise people to participate. It is also important to respect an informant's wish to withdraw from the study, particularly if it is not conducted in the way explicitly agreed in advance.

Confidentiality and Anonymity

Informants have the right to remain anonymous. Their confidentiality should be respected, and an attempt made to anticipate potential threats to both anonymity and confidentiality (e.g. by anonymising the data, making it secure, and sometimes even destroying it). But it is important to let informants know that it is not always possible to conceal identities completely, and that anonymity can sometimes be compromised unintentionally. Recognition of this should inform their consent.

Deception and Covert Research

This is an area of particular concern in applied linguistics. Covert research and deliberate deception are unacceptable to the extent that they violate the principle of informed consent and the right to privacy. However, in some research - concerned for example with phonological variation and pragmatic variation in naturally occurring speech - there are compelling methodological reasons for informants not being fully informed about the precise objectives of the research.

In such cases, defensible options would be to

- withhold the specific objectives of the research without deliberately misleading or giving false information (for example, informing doctors and patients that the research concerned the structure or progress of doctor-patients interviews without specifying that the aim was to study pause phenomena as an index of power);

- ask informants to consent to being deceived at some unspecified time in the future, on the grounds that the research could not be done otherwise. After the event, informants should then give their permission for the data, to be used;
- (if there are no methodological alternatives) present the objectives of the research to informants immediately after the data has been collected, guaranteeing anonymity if consent is given and destroying the data if it is withheld.

A distinction is sometimes made between deception and distraction. In contrast to the former, distraction is generally accepted as ethical, and it can be illustrated either in, for example, the introduction of multiple activities in a psycholinguistic experiment to prevent informants monitoring themselves or, alternatively, in situations of participation observation, in which informants come to accept the researcher as one of the community.

Observation in public places is a particularly problematic issue. If observations or recordings are made of the public at large, it is not possible to gain informed consent from everyone. However, post-hoc consent should be negotiated if the researcher is challenged by a member of the public.

A useful criterion by which to judge the acceptability of research is to anticipate or elicit, post hoc, the reaction of informants when they are told about the precise objectives of the study. If anger or other strong reactions are likely or expressed, then such data collection is inappropriate.

Consulting Informants on Completion of the Research

Wherever possible, final project reports should be made available in an accessible form to informants, and informants should have the right to comment on them.

As the discussion in section 5.1 suggested, some types of research, evaluation and consultancy make a good deal of space for informants' own priorities and perspectives. In such contexts, informants are more appositely described as 'participants'. All of the 'responsibilities to informants' described above apply to people who are more actively involved as participants in research. But some additional considerations also need to be borne in mind:

Balanced Participation

The practical consequences of the kinds of inquiry often designated action research, evaluation and consultancy, are usually much more immediate than they are in traditional research, affecting the distribution of power and resources in more obvious ways. In situations like this, where (a) participants have a significant degree of control over the research process, and (b) the political stakes are quite high, the notion of academic independence needs to be reformulated. In setting the agenda, in accessing and analysing the data, and in writing up the findings, the applied linguist may be happy to relinquish the autonomy entailed in traditional research, but she/he should take steps to avoid uncritically partisan alignment with any one interest group. In addition to the responsibilities outlined in 6.1 to 6.6, a number of checks and balances should be built into the research process to prevent it turning into advertising or propaganda:

- investigators should attend to a wide variety of perspectives on the issue, to the diverse claims made about it, to its context and history;
- no party should have privileged access to the data; the right to wholly determine the focus of the inquiry; sole access to project reports; or a unilateral veto over their contents;
- all participants should have the right to comment on the fairness, relevance and accuracy of project reports;
- all major interest groups should be represented on steering groups or management committees.

These recommendations are taken from the web site:

http://www.baal.org.uk/about_goodpractice_full.pdf (accessed on 21.11.2008)

Appendix 6- 1: Raw Frequencies of TMD Subcategories in the Data

Group	NS	NW	SQ	TP	CC	PR	AD	CO	CG	RA	SC
NES	1	5191	7	1	1	0	7	20	5	5	3
	2	2916	4	0	1	4	16	32	8	8	6
	3	5544	5	0	0	0	18	20	10	7	12
	4	5793	14	0	0	5	14	19	24	4	3
	5	3493	22	1	1	1	15	19	4	1	8
	6	5247	14	4	1	1	26	42	9	8	4
	7	5582	18	2	1	9	27	41	18	4	5
	8	3287	4	0	1	1	8	13	3	5	2
	9	2990	5	1	1	2	12	16	13	6	5
	10	2891	4	0	1	0	10	17	9	4	8
	11	3099	7	0	2	8	7	24	4	14	8
	12	5604	10	1	1	1	9	11	7	0	1
	13	2546	7	1	0	0	16	21	5	5	5
	14	2320	4	0	0	0	9	14	5	3	1
	15	3563	4	2	2	0	14	28	1	6	5
	16	4649	12	5	1	0	16	18	5	8	21
	17	7166	18	0	6	1	14	52	41	8	4
	18	3379	10	1	0	5	15	11	14	5	8
	19	4886	4	2	1	0	29	33	8	22	12
	20	4530	11	0	1	2	13	11	6	2	8
	21	3548	7	1	1	0	8	13	3	3	1
	22	4495	29	1	0	5	7	24	6	5	1
	23	3868	14	1	1	7	11	20	6	3	3
	24	3108	4	1	0	2	11	14	6	3	4
	25	4891	6	1	2	0	7	13	17	3	17
NNES	1	5230	22	3	5	0	19	20	7	0	1
	2	4052	6	0	1	2	14	7	10	6	9
	3	3963	29	0	0	0	13	27	16	11	7
	4	4504	8	2	1	3	31	34	23	9	0
	5	4048	26	0	5	3	22	19	14	12	10
	6	3133	6	2	1	1	29	24	23	2	3
	7	5026	18	2	0	3	8	26	5	6	4
	8	4696	19	1	1	5	35	31	37	8	7
	9	797	4	0	0	2	13	36	7	2	0
	10	4865	8	3	0	12	19	8	10	3	8
	11	4733	10	0	0	0	17	26	17	26	3
	12	3248	17	2	6	1	10	15	19	2	2
	13	4381	4	4	2	2	27	12	14	18	5
	14	4141	21	0	1	6	9	38	23	8	5
	15	2365	3	0	0	0	12	38	9	17	6
	16	4092	20	0	2	3	10	21	9	3	1
	17	4707	52	4	0	4	18	1	15	5	7
	18	1906	7	0	0	1	6	2	4	2	3
	19	4562	7	2	0	4	17	25	37	9	4
	20	1981	4	0	0	2	7	17	5	6	3
	21	2932	3	2	1	1	20	20	10	6	6
	22	5450	27	3	1	7	49	18	45	12	7
	23	2412	2	0	0	1	22	14	2	12	6
	24	3350	0	1	0	1	5	22	2	3	4
	25	4290	17	0	0	6	24	22	19	1	12
	26	4571	14	1	0	0	9	28	8	5	8
	27	3871	32	0	1	0	17	20	17	6	3
	28	5334	19	11	1	1	24	31	12	12	6
	29	5830	89	1	2	7	25	7	28	2	4

	30	4388	20	2	1	7	9	8	12	1	8
PRO	1	10863	20	1	1	12	47	68	19	21	19
	2	7722	17	1	3	6	34	47	23	13	8
	3	7864	20	1	1	8	16	40	16	11	15
	4	5409	12	1	1	1	17	13	11	10	8
	5	4395	4	0	1	1	25	46	30	13	7
	6	10532	18	4	2	9	40	67	33	30	34
	7	7832	40	1	3	1	19	41	23	25	11
	8	6707	18	0	3	0	23	33	23	3	6
	9	6411	8	1	2	3	15	30	64	10	22
	10	7011	3	7	1	1	19	23	97	7	12
	11	6587	26	0	4	2	24	34	10	5	19
	12	8599	11	0	3	9	62	29	11	19	8
	13	8184	19	1	7	9	19	27	29	0	6
	14	6618	12	1	2	2	31	37	30	6	8
	15	7970	19	0	2	0	8	31	23	10	7
	16	5226	6	0	6	0	26	44	41	7	3
	17	11004	4	1	12	2	51	79	30	17	11
	18	12634	43	2	28	0	39	89	56	17	12
	19	8222	8	1	0	0	17	58	15	18	9
	20	6759	3	0	2	1	20	35	18	7	6
	21	11135	34	0	10	3	36	67	19	15	8
	22	11312	64	1	1	9	28	70	27	18	9
	23	4707	8	0	1	0	13	36	18	10	15
	24	4432	15	2	2	0	14	21	21	2	2
	25	2530	0	0	0	0	9	33	5	1	1
	26	4861	15	0	1	11	6	38	16	8	6
	27	5355	5	0	1	1	20	16	23	16	10
	28	8669	12	1	0	5	23	44	25	9	11
	29	8647	28	1	5	0	30	57	75	19	25
	30	8400	9	1	1	0	25	42	22	2	1

NS: number of script

NW: number of words

PR: p/reviews

RA: rationales

SQ: sequencers

AD: additives

SC: sources

TP: topicalisers

CO: concessives

CC: concluders

CG: code glosses

Appendix 6- 2: Frequency per 5000 words for TMD

Group	Script no	No of words	Sequencers	Topicalisers	Concluders	Pre/reviews	Additives	Concessives	Code Glosses	Rationales	Sources
NES	1	5191	6.74	0.96	0.96	0.00	6.74	19.26	4.82	4.82	2.89
	2	2916	6.86	0.00	1.71	6.86	27.43	54.87	13.72	13.72	10.29
	3	5544	4.51	0.00	0.00	0.00	16.23	18.04	9.02	6.31	10.82
	4	5793	12.08	0.00	0.00	4.32	12.08	16.40	20.71	3.45	2.59
	5	3493	31.49	1.43	1.43	1.43	21.47	27.20	5.73	1.43	11.45
	6	5247	13.34	3.81	0.95	0.95	24.78	40.02	8.58	7.62	3.81
	7	5582	16.12	1.79	0.90	8.06	24.18	36.73	16.12	3.58	4.48
	8	3287	6.08	0.00	1.52	1.52	12.17	19.77	4.56	7.61	3.04
	9	2990	8.36	1.67	1.67	3.34	20.07	26.76	21.74	10.03	8.36
	10	2891	6.92	0.00	1.73	0.00	17.30	29.40	15.57	6.92	13.84
	11	3099	11.29	0.00	3.23	12.91	11.29	38.72	6.45	22.59	12.91
	12	5604	8.92	0.89	0.89	0.89	8.03	9.81	6.25	0.00	0.89
	13	2546	13.75	1.96	0.00	0.00	31.42	41.24	9.82	9.82	9.82
	14	2320	8.62	0.00	0.00	0.00	19.40	30.17	10.78	6.47	2.16
	15	3563	5.61	2.81	2.81	0.00	19.65	39.29	1.40	8.42	7.02
	16	4649	12.91	5.38	1.08	0.00	17.21	19.36	5.38	8.60	22.59
	17	7166	12.56	0.00	4.19	0.70	9.77	36.28	28.61	5.58	2.79
	18	3379	14.80	1.48	0.00	7.40	22.20	16.28	20.72	7.40	11.84
	19	4886	4.09	2.05	1.02	0.00	29.68	33.77	8.19	22.51	12.28
	20	4530	12.14	0.00	1.10	2.21	14.35	12.14	6.62	2.21	8.83
	21	3548	9.86	1.41	1.41	0.00	11.27	18.32	4.23	4.23	1.41
	22	4495	32.26	1.11	0.00	5.56	7.79	26.70	6.67	5.56	1.11
	23	3868	18.10	1.29	1.29	9.05	14.22	25.85	7.76	3.88	3.88
	24	3108	6.44	1.61	0.00	3.22	17.70	22.52	9.65	4.83	6.44
	25	4891	6.13	1.02	2.04	0.00	7.16	13.29	17.38	3.07	17.38
NNES	1	5230	21.03	2.87	4.78	0.00	18.16	19.12	6.69	0.00	0.96
	2	4052	7.40	0.00	1.23	2.47	17.28	8.64	12.34	7.40	11.11
	3	3963	36.59	0.00	0.00	0.00	16.40	34.07	20.19	13.88	8.83
	4	4504	8.88	2.22	1.11	3.33	34.41	37.74	25.53	9.99	0.00
	5	4048	32.11	0.00	6.18	3.71	27.17	23.47	17.29	14.82	12.35
	6	3133	9.58	3.19	1.60	1.60	46.28	38.30	36.71	3.19	4.79
	7	5026	17.91	1.99	0.00	2.98	7.96	25.87	4.97	5.97	3.98
	8	4696	20.23	1.06	1.06	5.32	37.27	33.01	39.40	8.52	7.45
	9	797	25.09	0.00	0.00	12.55	81.56	225.85	43.91	12.55	0.00
	10	4865	8.22	3.08	0.00	12.33	19.53	8.22	10.28	3.08	8.22
	11	4733	10.56	0.00	0.00	0.00	17.96	27.47	17.96	27.47	3.17
	12	3248	26.17	3.08	9.24	1.54	15.39	23.09	29.25	3.08	3.08
	13	4381	4.57	4.57	2.28	2.28	30.81	13.70	15.98	20.54	5.71
	14	4141	25.36	0.00	1.21	7.24	10.87	45.88	27.77	9.66	6.04
	15	2365	6.34	0.00	0.00	0.00	25.37	80.34	19.03	35.94	12.68
	16	4092	24.44	0.00	2.44	3.67	12.22	25.66	11.00	3.67	1.22
	17	4707	55.24	4.25	0.00	4.25	19.12	1.06	15.93	5.31	7.44
	18	1906	18.36	0.00	0.00	2.62	15.74	5.25	10.49	5.25	7.87

	19	4562	7.67	2.19	0.00	4.38	18.63	27.40	40.55	9.86	4.38
	20	1981	10.10	0.00	0.00	5.05	17.67	42.91	12.62	15.14	7.57
	21	2932	5.12	3.41	1.71	1.71	34.11	34.11	17.05	10.23	10.23
	22	5450	24.77	2.75	0.92	6.42	44.95	16.51	41.28	11.01	6.42
	23	2412	4.15	0.00	0.00	2.07	45.61	29.02	4.15	24.88	12.44
	24	3350	0.00	1.49	0.00	1.49	7.46	32.84	2.99	4.48	5.97
	25	4290	19.81	0.00	0.00	6.99	27.97	25.64	22.14	1.17	13.99
	26	4571	15.31	1.09	0.00	0.00	9.84	30.63	8.75	5.47	8.75
	27	3871	41.33	0.00	1.29	0.00	21.96	25.83	21.96	7.75	3.87
	28	5334	17.81	10.31	0.94	0.94	22.50	29.06	11.25	11.25	5.62
	29	5830	76.33	0.86	1.72	6.00	21.44	6.00	24.01	1.72	3.43
	30	4388	22.79	2.28	1.14	7.98	10.26	9.12	13.67	1.14	9.12
PRO	1	10863	9.21	0.46	0.46	5.52	21.63	31.30	8.75	9.67	8.75
	2	7722	11.01	0.65	1.94	3.89	22.02	30.43	14.89	8.42	5.18
	3	7864	12.72	0.64	0.64	5.09	10.17	25.43	10.17	6.99	9.54
	4	5409	11.09	0.92	0.92	0.92	15.71	12.02	10.17	9.24	7.40
	5	4395	4.55	0.00	1.14	1.14	28.44	52.33	34.13	14.79	7.96
	6	10532	8.55	1.90	0.95	4.27	18.99	31.81	15.67	14.24	16.14
	7	7832	25.54	0.64	1.92	0.64	12.13	26.17	14.68	15.96	7.02
	8	6707	13.42	0.00	2.24	0.00	17.15	24.60	17.15	2.24	4.47
	9	6411	6.24	0.78	1.56	2.34	11.70	23.40	49.91	7.80	17.16
	10	7011	2.14	4.99	0.71	0.71	13.55	16.40	69.18	4.99	8.56
	11	6587	19.74	0.00	3.04	1.52	18.22	25.81	7.59	3.80	14.42
	12	8599	6.40	0.00	1.74	5.23	36.05	16.86	6.40	11.05	4.65
	13	8184	11.61	0.61	4.28	5.50	11.61	16.50	17.72	0.00	3.67
	14	6618	9.07	0.76	1.51	1.51	23.42	27.95	22.67	4.53	6.04
	15	7970	11.92	0.00	1.25	0.00	5.02	19.45	14.43	6.27	4.39
	16	5226	5.74	0.00	5.74	0.00	24.88	42.10	39.23	6.70	2.87
	17	11004	1.82	0.45	5.45	0.91	23.17	35.90	13.63	7.72	5.00
	18	12634	17.02	0.79	11.08	0.00	15.43	35.22	22.16	6.73	4.75
	19	8222	4.86	0.61	0.00	0.00	10.34	35.27	9.12	10.95	5.47
	20	6759	2.22	0.00	1.48	0.74	14.80	25.89	13.32	5.18	4.44
	21	11135	15.27	0.00	4.49	1.35	16.17	30.09	8.53	6.74	3.59
	22	11312	28.29	0.44	0.44	3.98	12.38	30.94	11.93	7.96	3.98
	23	4707	8.50	0.00	1.06	0.00	13.81	38.24	19.12	10.62	15.93
	24	4432	16.92	2.26	2.26	0.00	15.79	23.69	23.69	2.26	2.26
	25	2530	0.00	0.00	0.00	0.00	17.79	65.22	9.88	1.98	1.98
	26	4861	15.43	0.00	1.03	11.31	6.17	39.09	16.46	8.23	6.17
	27	5355	4.67	0.00	0.93	0.93	18.67	14.94	21.48	14.94	9.34
	28	8669	6.92	0.58	0.00	2.88	13.27	25.38	14.42	5.19	6.34
	29	8647	16.19	0.58	2.89	0.00	17.35	32.96	43.37	10.99	14.46
	30	8400	5.36	0.60	0.60	0.00	14.88	25.00	13.10	1.19	0.60

Appendix 6- 3: Raw Frequencies of IMD in the data

Group	NS	NW	HD	EM	EV	CM	AP	SR
NES	1	5191	7	8	2	1	0	5
	2	2916	5	5	5	0	0	9
	3	5544	10	17	10	4	0	35
	4	5793	14	24	3	2	4	58
	5	3493	7	5	8	1	1	27
	6	5247	20	13	4	1	0	8
	7	5582	21	15	7	4	0	37
	8	3287	5	8	4	0	0	1
	9	2990	11	10	4	1	0	26
	10	2891	5	9	3	3	1	20
	11	3099	5	9	3	0	0	17
	12	5604	4	9	5	4	0	14
	13	2546	2	3	2	0	0	3
	14	2320	4	7	7	0	0	6
	15	3563	2	17	1	3	0	33
	16	4649	3	14	1	1	0	0
	17	7166	8	9	1	2	0	9
	18	3379	12	9	3	2	0	31
	19	4886	4	6	2	0	0	1
	20	4530	5	5	1	4	0	19
	21	3548	3	8	0	0	1	17
	22	4495	15	19	3	0	0	24
	23	3868	7	16	1	0	3	26
	24	3108	3	2	0	0	1	2
	25	4891	13	20	1	0	0	46
NNES	1	5230	6	25	1	0	0	1
	2	4052	3	13	3	1	0	15
	3	3963	1	15	5	0	0	0
	4	4504	12	11	1	10	0	60
	5	4048	8	6	8	1	0	3
	6	3133	7	14	5	2	0	16
	7	5026	5	20	4	0	1	0
	8	4696	3	9	3	2	0	3
	9	797	0	3	3	0	0	2
	10	4865	8	14	5	3	0	2
	11	4733	7	11	2	0	0	2
	12	3248	5	9	4	0	0	0
	13	4381	2	6	1	2	0	2
	14	4141	4	6	3	1	0	5
	15	2365	0	5	1	0	0	10
	16	4092	1	3	2	0	0	16
	17	4707	1	9	1	0	1	34
	18	1906	2	4	1	0	0	5
	19	4562	0	15	1	0	0	24
	20	1981	6	6	3	0	0	14
	21	2932	1	18	0	1	4	14
	22	5450	9	5	2	0	0	19
	23	2412	3	12	1	9	2	57
	24	3350	1	1	0	7	0	2
	25	4290	7	8	3	4	1	10
	26	4571	7	10	5	0	0	4
	27	3871	6	6	0	0	0	0
	28	5334	4	26	1	0	0	15

	29	5830	6	16	2	1	0	1
	30	4388	2	9	2	1	0	9
PRO	1	10863	27	37	1	0	0	65
	2	7722	22	16	2	8	0	46
	3	7864	18	11	5	0	0	54
	4	5409	7	6	4	0	2	11
	5	4395	3	8	6	0	2	9
	6	10532	26	28	11	6	0	20
	7	7832	10	9	5	0	0	1
	8	6707	7	7	8	0	0	0
	9	6411	38	30	4	0	0	2
	10	7011	12	9	14	0	1	1
	11	6587	11	11	6	0	0	4
	12	8599	15	18	3	9	5	213
	13	8184	30	21	6	4	0	134
	14	6618	10	6	5	0	0	5
	15	7970	16	24	4	1	0	13
	16	5226	10	10	4	0	0	4
	17	11004	37	31	6	1	0	127
	18	12634	23	23	3	0	0	1
	19	8222	8	11	1	0	2	46
	20	6759	4	12	12	0	0	5
	21	11135	17	34	2	0	0	23
	22	11312	28	47	3	18	2	24
	23	4707	6	4	6	0	0	14
	24	4432	4	19	0	0	1	1
	25	2530	3	8	0	0	0	3
	26	4861	22	13	0	2	2	45
	27	5355	11	9	2	1	0	7
	28	8669	10	10	1	0	0	30
	29	8647	15	35	2	0	3	31
	30	8400	26	33	1	0	1	7

NW: number of words
EV: evaluatives
SR: self references

HD: hedges
CM: commentaries

EM: emphatics
AP: appeals

Appendix 6- 4: Frequency per 5000 words for IMD

Group	NS	NW	HD	EM	EV	CM	AP	SR
NES	1	5191	6.74	7.71	1.93	0.96	0.00	4.82
	2	2916	8.57	8.57	8.57	0.00	0.00	15.43
	3	5544	9.02	15.33	9.02	3.61	0.00	31.57
	4	5793	12.08	20.71	2.59	1.73	3.45	50.06
	5	3493	10.02	7.16	11.45	1.43	1.43	38.65
	6	5247	19.06	12.39	3.81	0.95	0.00	7.62
	7	5582	18.81	13.44	6.27	3.58	0.00	33.14
	8	3287	7.61	12.17	6.08	0.00	0.00	1.52
	9	2990	18.39	16.72	6.69	1.67	0.00	43.48
	10	2891	8.65	15.57	5.19	5.19	1.73	34.59
	11	3099	8.07	14.52	4.84	0.00	0.00	27.43
	12	5604	3.57	8.03	4.46	3.57	0.00	12.49
	13	2546	3.93	5.89	3.93	0.00	0.00	5.89
	14	2320	8.62	15.09	15.09	0.00	0.00	12.93
	15	3563	2.81	23.86	1.40	4.21	0.00	46.31
	16	4649	3.23	15.06	1.08	1.08	0.00	0.00
	17	7166	5.58	6.28	0.70	1.40	0.00	6.28
	18	3379	17.76	13.32	4.44	2.96	0.00	45.87
	19	4886	4.09	6.14	2.05	0.00	0.00	1.02
	20	4530	5.52	5.52	1.10	4.42	0.00	20.97
	21	3548	4.23	11.27	0.00	0.00	1.41	23.96
	22	4495	16.69	21.13	3.34	0.00	0.00	26.70
	23	3868	9.05	20.68	1.29	0.00	3.88	33.61
	24	3108	4.83	3.22	0.00	0.00	1.61	3.22
	25	4891	13.29	20.45	1.02	0.00	0.00	47.03
NNES	1	5230	5.74	23.90	0.96	0.00	0.00	0.96
	2	4052	3.70	16.04	3.70	1.23	0.00	18.51
	3	3963	1.26	18.93	6.31	0.00	0.00	0.00
	4	4504	13.32	12.21	1.11	11.10	0.00	66.61
	5	4048	9.88	7.41	9.88	1.24	0.00	3.71
	6	3133	11.17	22.34	7.98	3.19	0.00	25.53
	7	5026	4.97	19.90	3.98	0.00	0.99	0.00
	8	4696	3.19	9.58	3.19	2.13	0.00	3.19
	9	797	0.00	18.82	18.82	0.00	0.00	12.55
	10	4865	8.22	14.39	5.14	3.08	0.00	2.06
	11	4733	7.39	11.62	2.11	0.00	0.00	2.11
	12	3248	7.70	13.85	6.16	0.00	0.00	0.00
	13	4381	2.28	6.85	1.14	2.28	0.00	2.28
	14	4141	4.83	7.24	3.62	1.21	0.00	6.04
	15	2365	0.00	10.57	2.11	0.00	0.00	21.14
	16	4092	1.22	3.67	2.44	0.00	0.00	19.55
	17	4707	1.06	9.56	1.06	0.00	1.06	36.12
	18	1906	5.25	10.49	2.62	0.00	0.00	13.12
	19	4562	0.00	16.44	1.10	0.00	0.00	26.30
	20	1981	15.14	15.14	7.57	0.00	0.00	35.34
	21	2932	1.71	30.70	0.00	1.71	6.82	23.87
	22	5450	8.26	4.59	1.83	0.00	0.00	17.43
	23	2412	6.22	24.88	2.07	18.66	4.15	118.16
	24	3350	1.49	1.49	0.00	10.45	0.00	2.99
	25	4290	8.16	9.32	3.50	4.66	1.17	11.66
	26	4571	7.66	10.94	5.47	0.00	0.00	4.38
	27	3871	7.75	7.75	0.00	0.00	0.00	0.00
	28	5334	3.75	24.37	0.94	0.00	0.00	14.06

	29	5830	5.15	13.72	1.72	0.86	0.00	0.86
	30	4388	2.28	10.26	2.28	1.14	0.00	10.26
PRO	1	10863	12.43	17.03	0.46	0.00	0.00	29.92
	2	7722	14.25	10.36	1.30	5.18	0.00	29.79
	3	7864	11.44	6.99	3.18	0.00	0.00	34.33
	4	5409	6.47	5.55	3.70	0.00	1.85	10.17
	5	4395	3.41	9.10	6.83	0.00	2.28	10.24
	6	10532	12.34	13.29	5.22	2.85	0.00	9.49
	7	7832	6.38	5.75	3.19	0.00	0.00	0.64
	8	6707	5.22	5.22	5.96	0.00	0.00	0.00
	9	6411	29.64	23.40	3.12	0.00	0.00	1.56
	10	7011	8.56	6.42	9.98	0.00	0.71	0.71
	11	6587	8.35	8.35	4.55	0.00	0.00	3.04
	12	8599	8.72	10.47	1.74	5.23	2.91	123.85
	13	8184	18.33	12.83	3.67	2.44	0.00	81.87
	14	6618	7.56	4.53	3.78	0.00	0.00	3.78
	15	7970	10.04	15.06	2.51	0.63	0.00	8.16
	16	5226	9.57	9.57	3.83	0.00	0.00	3.83
	17	11004	16.81	14.09	2.73	0.45	0.00	57.71
	18	12634	9.10	9.10	1.19	0.00	0.00	0.40
	19	8222	4.86	6.69	0.61	0.00	1.22	27.97
	20	6759	2.96	8.88	8.88	0.00	0.00	3.70
	21	11135	7.63	15.27	0.90	0.00	0.00	10.33
	22	11312	12.38	20.77	1.33	7.96	0.88	10.61
	23	4707	6.37	4.25	6.37	0.00	0.00	14.87
	24	4432	4.51	21.44	0.00	0.00	1.13	1.13
	25	2530	5.93	15.81	0.00	0.00	0.00	5.93
	26	4861	22.63	13.37	0.00	2.06	2.06	46.29
	27	5355	10.27	8.40	1.87	0.93	0.00	6.54
	28	8669	5.77	5.77	0.58	0.00	0.00	17.30
	29	8647	8.67	20.24	1.16	0.00	1.73	17.93
	30	8400	15.48	19.64	0.60	0.00	0.60	4.17

Appendix 6- 5: Raw Frequencies of Sub-categories in TMD

Appendix 6-5- 1: Sequencers:

Group	No.	First	firstly	finally	first of all	lastly	listing(abc)	next	second	secondly	subsequently	third	thirdly	Total
NES	1	3	1						2	1				7
	2	4												4
	3	3		1				1						5
	4	5						2	6			1		14
	5	10		2					8			2		22
	6	5							9					14
	7	10						1	6			1		18
	8	1		1	1					1				4
	9	2	1					1			1			5
	10	1		1					2					4
	11	4							2			1		7
	12	4		1				1	3	1				10
	13	5		1					1					7
	14								2			2		4
	15	4												4
	16	9		1					1			1		12
	17	3						1	10		4			18
	18	4		2			1	2	1					10
	19	3		1										4
	20	5		2				2	1			1		11
	21	1					1	1	1			3		7
	22	11						5	8		2	3		29
	23	6		1					4			3		14
	24		1	1					1			1		4
	25	6												6
SUM		109	3	15	1	0	2	17	68	3	7	19	0	
NNES	1	10	1					1	10					22
	2			4					1			1		6
	3	9	2					1	9	2		5	1	29
	4	5	1							2				8
	5	12							7		4	3		26
	6		3	2						1				6
	7	9						2	5			2		18
	8	7		1	2				6			2	1	19
	9	1						1	1			1		4
	10	4		1					2	1				8
	11	7	1							1			1	10
	12	2	5	1				2	1	4			2	17
	13	2		1				1						4
	14	8						7	4			2		21

	15								2			1		3
	16	8		3				1	5			3		20
	17	17				1		3	18			13		52
	18	2	1	1		1		1	1					7
	19	1						4	2					7
	20	3								1				4
	21	3												3
	22	11	1	5					3	1		6		27
	23							1	1					2
	24													0
	25	8		3				2	3			1		17
	26	6			1			3	1	1		1	1	14
	27	5	2	5				11	5	2		2		32
	28	10		2				2	3			2		19
	29	34	1	3			1	7	25		1	17		89
	30	10						4	6					20
SUM		194	18	32	3	2	1	54	121	16	5	62	6	
PRO	1	10		5				3	2					20
	2	6	2	3				3		2	1			17
	3	11							6	2			1	20
	4	5					1	3	1		1	1		12
	5	2							2					4
	6	7		1				1	3		4	2		18
	7	20		1				8	7	2		1	1	40
	8	5		7				1	3			2		18
	9	6		1					1					8
	10	2							1					3
	11	7		2				3	8			6		26
	12	4						3	3			1		11
	13	9						1	4	2	2	1		19
	14	4		3								5		12
	15	7		2				3	3		1	3		19
	16	3		1					2					6
	17			3								1		4
	18	14		1			5	2	17		1	3		43
	19	7							1					8
	20			2				1						3
	21	2						12	16			4		34
	22	5		4	1			1	52			1		64
	23	3							1		2	2		8
	24	7		3					4			1		15
	25													0
	26	2							9		1	3		15
	27	1		3					1					5
	28	7		1					1			3		12
	29	13		2				1	6		1	5		28
	30	4		4					1					9
SUM		173	2	49	1	0	6	46	155	8	14	45	2	

Appendix 6-5- 2: Topicalisers

Group	No.	in regard to/regarding	move on	turn to	with regard to	as for	based on	bearing in mind	in relation/respect to	Total
NES	1				1					1
	2									0
	3									0
	4									0
	5			1						1
	6				3		1			4
	7	1								1
	8									0
	9	1								1
	10									0
	11									0
	12				1					1
	13						1			1
	14									0
	15			1						1
	16	4				1				5
	17									0
	18					1				1
	19	1								1
	20									0
	21			1						1
	22	1								1
	23					1				1
	24					1				1
	25			1						1
SUM		8	0	4	5	4	2	0	0	23
NNES	1	1			1	1				3
	2									0
	3									0
	4				1				1	2
	5									0
	6						2			2
	7						1	1		2
	8					1				1
	9									0
	10						2	1		3
	11									0
	12				1				1	2
	13			1	1		2			4
	14									0
	15									0
	16									0
	17		1				3			4
	18									0
	19						2	2		4

	20									0
	21	1						1		2
	22	1			1	1				3
	23									0
	24					1				1
	25									0
	26	1								1
	27									0
	28				5	6				11
	29						1			1
	30			1			1			2
SUM		4	1	2	10	10	14	3	4	46
PRO	1						1			1
	2				1					1
	3				1					1
	4				1					1
	5									0
	6		1		3					4
	7				1					1
	8									0
	9						1			1
	10	2					5			7
	11									0
	12									0
	13							1		1
	14				1					1
	15									0
	16								1	1
	17				1					1
	18									0
	19			1						1
	20									0
	21									0
	22		1							1
	23									0
	24						1			1
	25									0
	26									0
	27									0
	28				1					1
	29									0
	30						1			1
SUM		2	2	1	10	1	8	1	1	26

Appendix 6-5- 3: Concluders

Group	No.	in conclusion	in sum	in summary	on the whole	overall	so far	thus far	to conclude	to sum up	to summarize	Total
NES	1	1										1
	2						1					1
	3											0
	4											0
	5					1						1
	6						1					1
	7							1				1
	8					1						1
	9					1						1
	10					1						1
	11					2						2
	12					1						1
	13											0
	14											0
	15	1						1				2
	16					1						1
	17					6						6
	18											0
	19							1				1
	20					1						1
	21					1						1
	22											0
	23					1						1
	24											0
	25	1					1					2
SUM		3	0	0	0	17	3	3	0	0	0	
NNES	1					2	2		1			5
	2					1						1
	3											0
	4						1					1
	5					5						5
	6				1							1
	7											0
	8								1			1
	9											0
	10											0
	11											0
	12					6						6
	13					1				1		2
	14				1							1
	15											0
	16					2						2
	17											0

	18											0
	19											0
	20											0
	21						1					1
	22								1			1
	23											0
	24											0
	25											0
	26											0
	27									1		1
	28					1						1
	29					2						2
	30										1	1
SUM		0	0	0	2	20	4	0	3	3	0	
PRO	1					1						1
	2		1			1			1			3
	3					1						1
	4						1					1
	5					1						1
	6	1				1						2
	7					3						3
	8			1	1	1						3
	9				1		1					2
	10					1						1
	11			1		3						4
	12		3									3
	13					6	1					7
	14				1		1					2
	15					2						2
	16					6						6
	17		1		3	8						12
	18		4			22		2				28
	19											0
	20					2						2
	21			1		8				1		10
	22					1						1
	23		1									1
	24							1			1	2
	25											0
	26								1			1
	27					1						1
	28											0
	29					3	1	1				5
	30		1									1
SUM		1	11	3	6	72	5	4	2	1	1	

Appendix 6-5- 4: Previews/reviews

Group	No,	Previously/earlier	(in) part X	(in) section x	(in) the section X	(in) this chapter	(in) this part	(in) this section	Page x	figure x	table x	above	below	this work/paper	Total
NES	1	0													0
	2			3								1			4
	3														0
	4							2				2	1		5
	5											1			1
	6	1	1												1
	7			3								5	1		9
	8		1												1
	9											2			2
	10	2													0
	11	5		5								2	1		8
	12	1										1			1
	13														0
	14														0
	15	1													0
	16														0
	17	2											1		1
	18			1			1					3			5
	19	4													0
	20									1		1	1		2
	21	3													0
	22											3	2		5
	23			4							1		3		7
	24			1								1			2
	25														0
SUM		19	2	17	0	0	0	1	2	1	1	22	10	0	
NNES	1	1													0
	2											2			2
	3														0
	4											3			3
	5	1						1				1	2		3
	6											1			1
	7							1					2		3
	8											5			5
	9			1								1			2
	10											12	1		12
	11														0
	12											1		1	2
	13							1			1	1			2
	14					1					1	4	1		6
	15														0
	16							2			2	1			3
	17							1				3			4
	18											1			1

	19			2								2			4
	20				1							1			2
	21			1											1
	22			1			1		1			3		1	7
	23												1		1
	24											1			1
	25								1			3	3		6
	26														0
	27												1		0
	28	1										1			1
	29	1		1				3				3			7
	30			1								6			7
SUM		4	0	7	1	1	1	9	1	1	4	56	11	2	
PRO	1							1				8	3		12
	2								1	6	2	2			6
	3			4						6	3			1	8
	4					1				5					1
	5											1			1
	6											4	6		9
	7								3	4	1				1
	8									1					0
	9										9	1	2		3
	10	1			1										1
	11	1									2	2			2
	12											9			9
	13			3				2		1	3	3		1	9
	14	1				1							1		2
	15	1									10				0
	16										2				0
	17								1			2			2
	18	2									5				0
	19										2				0
	20											1			1
	21	1										11	3		3
	22											5	6	3	9
	23														0
	24									5					0
	25														0
	26											8	3		11
	27										3	1			1
	28							1			2	4			5
	29										1				0
	30	1													0
SUM		8	0	7	1	2	0	4	0	16	72	56	23	2	

Appendix 6-5- 5: Additives

Group	No	accordingly	also	furthermore	in addition/additionally	in the same way	moreover	therefore	likewise	Total
NES	1	1	5					1		7
	2		14					2		16
	3		17					1		18
	4		8			1		5		14
	5		13			1		1		15
	6	1	17	3		1		4		26
	7		19					7	1	27
	8		4			1		3		8
	9		2	3				7		12
	10		9					1		10
	11		2	1				4		7
	12		5			1		3		9
	13		8	2	1		1	4		16
	14		8		1					9
	15		4	2	2		3	3		14
	16		7		2		4	3		16
	17	3	4	2			1	4		14
	18		11	1				3		15
	19		22	1				5	1	29
	20		9					4		13
	21		7	1						8
	22		5					2		7
	23		8					3		11
	24	1	5		1		1	2	1	11
	25		6			1				7
SUM		6	219	16	7	6	10	72	3	
NNES	1	1	9		3		2	4		19
	2	1	12					1		14
	3		7		1		1	4		13
	4	1	24	1	3			2		31
	5	1	6	1			3	11		22
	6	5	16	1	2		1	4		29
	7		6		1		1			8
	8		9	3	7			16		35
	9		2	9	1			1		13
	10	1	16		2					19
	11		6	1	1		2	7		17
	12		6		2			2		10
	13		20	1	3		1	2		27
	14	2	4					3		9
	15		4	1	7					12
	16		8				2			10
	17		13		3		1	1		18

	18		3					3		6
	19		12	2	1			2		17
	20	2	2				2	1		7
	21	1	8				1	10		20
	22	2	29	1	2		3	12		49
	23	1	14	1	2			4		22
	24		1	2				2		5
	25		8	1	5		2	8		24
	26		7		1			1		9
	27		6	2	1		5	3		17
	28		18		4		1	1		24
	29		14	1	5			4	1	25
	30		6				1	2		9
SUM		18	296	28	57	0	29	111	1	
PRO	1		37		9	1				47
	2		22		4	1	2	5		34
	3	1	11			2	1	1		16
	4		16					1		17
	5		17		4			4		25
	6		26	6	2			6		40
	7		14		1	1	3			19
	8		17	4	1			1		23
	9		8	1	5			1		15
	10		16	1	1			1		19
	11		14	1	2			4	3	24
	12		47		10			5		62
	13		9		1		1	8		19
	14		25		1			5		31
	15		8							8
	16		24		1		1			26
	17		45		3			3		51
	18	2	26		8			3		39
	19		8	1	3	1		2	2	17
	20		12		2		2	4		20
	21		17		6	2	3	8		36
	22		19	1	2		1	5		28
	23		6		4			3		13
	24		13		1					14
	25	1	7		1					9
	26		6							6
	27		11	1			1	7		20
	28		15	1	4			3		23
	29		26					4		30
	30		20		4			1		25
SUM		4	542	17	80	8	15	85	5	

Appendix 6-5- 6: Concessives

Group	No	Although	but	by contrast	however	in spite of	on the contrary	on the other hand	whereas	yet	Total
NES	1	2	11		7						20
	2	3	15		12					2	32
	3		13		4					3	20
	4	4	9		1			1	1	3	19
	5	1	11		7						19
	6	10	22		10						42
	7	4	19		14	1		2	1		41
	8		8		2					3	13
	9	1	5		8	1				1	16
	10	1	4		12						17
	11	1	18		5						24
	12	1	4		5				1		11
	13	1	12		6					2	21
	14	3	7		4						14
	15	8	4		14	2					28
	16	3	9		5				1		18
	17	7	20		9		1	1		14	52
	18	1	6		4						11
	19	6	15		10					2	33
	20		7		4						11
	21		7		6						13
	22		18		3					3	24
	23	3	11		6						20
	24	3	5		1			1	1	3	14
	25	5	1		4					3	13
SUM		68	261	0	163	4	1	5	5	39	
NNES	1	2	4		9		1		4		20
	2	2	3		2						7
	3		24	1					1	1	27
	4	3	25		5			1			34
	5	4	5		3			1	6		19
	6		13	1	3	1		4	1	1	24
	7	5	11		7	2		1			26
	8	1	12		7	1		2	8		31
	9	3	12		19					2	36
	10		5		2			1			8
	11	2	13		10			1			26
	12	2	5		6					2	15
	13		8		4						12
	14	7	8		22					1	38
	15	6	22		8		2				38
	16	1	15		1			3		1	21
	17				1						1
	18		1		1						2
	19	8	6		6		5				25

	20	1	11		5						17
	21		12		6			1	1		20
	22		10		5	1		2			18
	23	1	6		7						14
	24	2	13		6					1	22
	25	1	7		4			5	1	4	22
	26	5	11		10			1	1		28
	27	3	10		5			1		1	20
	28	1	19		11						31
	29		3		4						7
	30	3	1		4						8
SUM		63	295	2	183	5	8	24	23	14	
PRO	1	11	32		16			2	2	5	68
	2	4	27		11	1		3		1	47
	3	5	20	1	8				6		40
	4	3	5		3			1		1	13
	5	10	11		18				4	3	46
	6	8	44		11				4		67
	7	5	17		14			1	3	1	41
	8	6	12		15						33
	9	1	12		8			3	6		30
	10	6	4		10			2	1		23
	11	3	15		13			2	1		34
	12	1	20		4			1		3	29
	13	2	8		13			3		1	27
	14	3	21		6				6	1	37
	15	6	7		13				3	2	31
	16	5	19		11				9		44
	17	7	41		19	3		3	5	1	79
	18	16	32		19			6	15	1	89
	19	17	21		17	2				1	58
	20	14	14		3					4	35
	21	10	24		17	1	2	5	7	1	67
	22	6	39	1	6	4		1		13	70
	23	5	15		9				7		36
	24	1	9		1					10	21
	25	15	8		6			1	3		33
	26	1	31		6						38
	27	1	4		9			1		1	16
	28	12	13		19						44
	29	11	25		18			1		2	57
SUM	30	9	20	2	3	11	2	39	83	58	42

Appendix 6-5- 7: Code Glosses

Group	No.	Defined as	e.g.	for example	for instance	I mean	i.e.	in other words known as	namely	such as	that means	which means	Total
NES	1		1	1			2			1			5
	2	1		6				1					8
	3			2				1	1	6			10
	4			10					1	12	1		24
	5			1			2			1			4
	6			3						6			9
	7		6	4		1	3	1	2	1			18
	8			1			1			1			3
	9		1	3				2		6		1	13
	10		1	2				1		5			9
	11			1					3				4
	12			1			1	2		3			7
	13		1							3		1	5
	14			1	1					3			5
	15			1									1
	16			2						3			5
	17		9	5	4	1	13	5	2	2			41
	18		6	2			4			2			14
	19			1						6		1	8
	20			4						2			6
	21			1				1		1			3
	22		2	1			2			1			6
	23									6			6
	24	1	1	3				1					6
	25		5	6				1		5			17
SUM		2	33	62	5	2	28	12	5	8	76	1	3
NNES	1		1				3	1		2			7
	2		1	6				1				2	10
	3			4	2		2			8			16
	4			7	8					7		1	23
	5	1	1	1	1		4	2		4			14
	6			1	2			4		13		3	23
	7			3						2			5
	8		8	2	2		21	2		2			37
	9		2	2	2					1			7
	10		1	7							1	1	10
	11			4	3			1		8		1	17
	12		11	4				1		3			19
	13	1	1	4				2		6			14
	14	2	15	2	1					2		1	23
	15			2	3					3		1	9
	16	1	1	2						5			9
	17	1	8	2	1					3			15
	18							1		3			4
	19		5	10	6	6	2	1		7			37

	20			2	2								1		5
	21			2			1	5		1	1				10
	22			14	2		1		1		23		4		45
	23		1						1						2
	24				1					1					2
	25		1	4	8				1		5				19
	26		2	3				1			2				8
	27			9	3						4		1		17
	28	1		1			6		2	1	1				12
	29	1	2	8	4	1			1		11				28
	30		4	1	2				1		3		1		12
SUM		8	65	107	53	7	40	20	9	3	129	1	17		
PRO	1		7	9		2							1		19
	2		7	5	1			4		1	5				23
	3		6	5	1		2	1		1					16
	4		1	5			1				4				11
	5	1	1	9			1	1			17				30
	6	1		10	1	5	11		1		4				33
	7		15	2	1			1			4				23
	8		10	2			2		1		8				23
	9		39	6						2	17				64
	10		47	17	7		1				25				97
	11			3			1		1		5				10
	12		1	9			1								11
	13		4	12			7	4	1		1				29
	14		17	1							12				30
	15		5	4			2				12				23
	16		33	7							1				41
	17		1	11		6	1				11				30
	18		18	6			21	1		7	3				56
	19		1	6				2			6				15
	20			5			4	3			6				18
	21		10	2			3				4				19
	22		11	7				1			7	1			27
	23		6	2			4		1		5				18
	24			5			3	3			10				21
	25		2	3											5
	26		7	7			2								16
	27		7	2			8				6				23
	28		3	3	2		1				16				25
	29		38	15			13		1		8				75
	30		4	13				1			4				22
SUM		2	301	193	13	13	89	22	6	11	201	1	1		

Appendix 6-5- 8: Rationales

Group	No	The reason for this	as a result of	Because(of)	as a consequence/consequently	thereby	Total
NES	1		1	3		1	5
	2			8			8
	3			6		1	7
	4			4			4
	5			1			1
	6			7	1		8
	7			4			4
	8		1	4			5
	9			2	4		6
	10		1	3			4
	11		2	12			14
	12						0
	13			5			5
	14			3			3
	15			2	4		6
	16		3	3	2		8
	17		2	2	4		8
	18			5			5
	19		1	20	1		22
	20			2			2
	21	1	1	1			3
	22		1	2	2		5
	23			3			3
	24			3			3
	25		1	2			3
SUM		1	14	107	18	2	
NNES	1						0
	2		2	3	1		6
	3			9	2		11
	4			9			9
	5		3	7	2		12
	6			2			2
	7			5	1		6
	8		2	6			8
	9		1	1			2
	10			3			3
	11		3	22	1		26
	12		2				2
	13		1	17			18
	14			7	1		8
	15			17			17
	16		1	2			3
	17			5			5
	18			2			2
	19		1	8			9
	20		2	3	1		6

	21		4	2			6
	22	1	3	8			12
	23		1	11			12
	24			3			3
	25				1		1
	26		3	1	1		5
	27			5	1		6
	28		1	10	1		12
	29			2			2
	30			1			1
SUM		1	30	171	13	0	
PRO	1			21			21
	2		1	10	2		13
	3			10	1		11
	4			10			10
	5			13			13
	6			23	1	6	30
	7			24		1	25
	8		1	1	1		3
	9			8	2		10
	10			7			7
	11			5			5
	12		4	15			19
	13						0
	14			4	1	1	6
	15			7	2	1	10
	16		1	5		1	7
	17			17			17
	18			11		6	17
	19		1	17			18
	20		1	6			7
	21			15			15
	22		1	17			18
	23			9		1	10
	24			2			2
	25			1			1
	26		4	1	1	2	8
	27		1	15			16
	28			8	1		9
	29		1	17	1		19
	30		2				2
SUM		0	18	299	13	19	

Appendix 6-5- 9: Sources

Group	No	According to x	X suggests	X argues	X notes	X states	X proposes	Total
NES	1	2		1				3
	2		1				5	6
	3	7	2			3		12
	4	3						3
	5	4				4		8
	6	2	1	1				4
	7	2	2			1		5
	8	2						2
	9			2		3		5
	10			7			1	8
	11		1		5		2	8
	12	1						1
	13	3	1	1				5
	14					1		1
	15	2		2		1		5
	16	16		2	2	1		21
	17	1		1	1	1		4
	18		2	4		1	1	8
	19	2	1	2		1	6	12
	20	7		1				8
	21	1						1
	22				1			1
	23		3					3
	24	3			1			4
	25	1	7	9				17
SUM		59	21	33	10	17	15	
NNES	1	1						1
	2	7		1		1		9
	3	6					1	7
	4							0
	5	7	1	2				10
	6	3						3
	7	2	2					4
	8	6	1					7
	9							0
	10	4	2	2				8
	11	2		1				3
	12	2						2
	13	2					3	5
	14	4		1				5
	15	6						6
	16	1						1
	17	4	1	2				7
	18	1		2				3
	19	3		1				4
	20	2	1					3

	21	1		4			1	6
	22	5	1				1	7
	23	6						6
	24	2		1			1	4
	25	7	1	4				12
	26	4	3				1	8
	27	2				1		3
	28	1		1			4	6
	29	3				1		4
	30			1	1	1	5	8
SUM		94	13	23	1	4	17	
PRO	1		9	1	5		4	19
	2	1	5	2				8
	3		8	3			4	15
	4	1	6	1				8
	5		1	1	3		2	7
	6	7		12			15	34
	7	1	5	2			3	11
	8	1	4	1				6
	9		16	4		1	1	22
	10	8		1	1	1	1	12
	11	2	1	13		3		19
	12		6	2				8
	13	1	3	2				6
	14	1	3	2			2	8
	15	4			1		2	7
	16		2	1				3
	17	4	2	5				11
	18	1	9	2				12
	19	2	2	5				9
	20	1	1	4				6
	21	6	1	1				8
	22	2	4	2	1			9
	23	2	4	7			2	15
	24	1		1				2
	25		1					1
	26	4	1				1	6
	27	9		1				10
	28	8	1	1		1		11
	29	6	8	10	1			25
	30		1					1
SUM		73	104	87	12	6	37	

Appendix 6-6- 2: Emphatics

Group	Script	Actually	already	always	certainly	clearly	commonly	constantly	definitely	doubtless	easily	especially	in fact	no doubt	Obviously (obvious)	of course	really	simply	surely	truly	undeniably	undoubtedly	total
NES	1			1		5	1					1											8
	2		3			1						1											5
	3			3	1	4						1	3			2	1	2					17
	4	2		1	2					2		3	6			3	6	2	3				24
	5		2								1	1	1					1					5
	6	1	3		1		1						1			3		3					13
	7		3			4	2			1				2	2						1		15
	8	3				1		1				1				1		1					8
	9			1				1				4	1					2	1				10
	10	1		2						1		3	2										9
	11	2	1	3		1											1	1					9
	12	1	2							2		2	1		1								9
	13											2								1			3
	14	1										1						4					7
	15		2	2	1	2						5	2						1			2	17
	16	1		2	1		1				1	1	2					3		2			14
	17			1		2	1					1				1		2			1		9
	18	1	3			1						2			1			1					9
	19		1	1		1	1					2											6
	20			3		1							1										5
	21	1	1	3		1									1								8
	22	4	1		2	4					1		5					2					19
	23	2	1	6		1						1	2				3						16

Appendix 6-6- 4: Commentaries

Group	Script	allow me	you might be missing	you might add	consider(you might)	you might call	imagine you will	(the) reader's	let us	you	your	total
NES	1									1		1
	2											0
	3					2			2			4
	4	1						1				2
	5							1				1
	6					1						1
	7						4					4
	8											0
	9					1						1
	10							3				3
	11											0
	12							3		1		4
	13											0
	14											0
	15							2		1		3
	16									1		1
	17							1		1		2
	18							2				2
	19											0
	20							4				4
	21											0
	22											0
	23											0
	24											0
	25											0
SUM		1	0	0	0	0	4	0	3	18	3	

Appendix 6-6- 5: Appeals

Group	Script	(?)What should I ...	How should I ?	Will they.....?	Will each student ...?	Do you..?	Does it..?	How does it ..?/how do people...?	What circumstances...?	(When) is it...?	Do we...? total
NES	1										0
	2										0
	3										0
	4		1			1				2	4
	5									1	1
	6										0
	7										0
	8										0
	9										0
	10									1	1
	11										0
	12										0
	13										0
	14										0
	15									1	10
	16										0
	17										0
	18										0
	19										0
	20										0
	21									1	1
	22										0
	23								3		3
	24								1		1
	25										0
SUM		0	1	0	0	1	0	0	0	4	6

Appendix 6-6- 6: Self-references

Group	Script	I	we	me	my	our	us	author	researcher	total
NES	1		2			3				5
	2	4	5							9
	3	7	12	4	8	3	1			35
	4	13	22	2	7	10	4			58
	5	19	3		1		4			27
	6		5			3				8
	7	10	16		6	2	3			37
	8					1				1
	9	6	4	3	9	2	2			26
	10	9	4	3	4					20
	11	12			5					17
	12	6			1	4	3			14
	13		2		1					3
	14	3	1	1		1				6
	15	5	23		2	3				33
	16									0
	17		5			3	1		1	10
	18	5	11	1	8	4	2			31
	19					1				1
	20	15	3			1				19
	21	4			2	7	4			17
	22	4	17		3					24
	23	15		1	10					26
	24	1	1							2
	25	27	10		6	3				46
SUM		165	146	15	73	51	24			
NNES	1				1					1
	2	1	7		2	1	4			15

