

# An Empirical Investigation into the Nature and Degree of Nominality

Alexander Carr

Thesis submitted in partial fulfilment of the degree of  
Doctor of Philosophy

Centre for Language and Communication Research

School of English, Communication and Philosophy

Cardiff University

December 2022

# Abstract

This thesis presents an empirical investigation into the nature and degree of nominality, from a semantic perspective. To achieve an empirical understanding of nominal semantic behaviour, the research fundamentally draws upon the framework of Lexical Aspect (Vender 1967) and applies it in the nominal domain. The thesis is separated into two distinct studies, taking a multi-faceted approach.

The first study extracts 5000 nominal instances from the British National Corpus (BNC 2022) and explores how nominals express events and states, by focusing on the extent to which nominal category features, such as abstract/concrete status, count/mass status, genre, and word formation type, influence semantic expression. The results reveal that while abstract/concrete status holds the greatest influence over semantic behaviour, variable interactions between abstract/concrete status, count/mass status and word formation type are highly responsible for expression of different semantic distributions.

The second study takes a smaller scale approach, extracting 500 instances of four Underived event nominals (UENs) (n=2000) from the Timestamped JSI Web Corpus English (Trampus et al. 2004), and focusing specifically on how definiteness and the process type of the clause (Halliday 1985) influence the semantic behaviour of the four UENs. The results suggest that definiteness and process type are largely uninfluential on the semantic expression of the UENs. However, the divergent semantic behaviour of the UENs signals a difference in semantic flexibility between nominals. Through closer inspection of this semantic behaviour, it is proposed that nominals that share an association with (human) agency may be more susceptible to semantic coercion, due to the capacity of agents to exert control over situations and enact change.

This research concludes that the influence of nominal category features on semantic behaviour is highly interconnected and that while notions such as boundedness, homogeneity and telicity are certainly transferable between nominal and verb domains, each lexical class presents different semantic constraints on how these concepts behave. Furthermore, this thesis concludes that associations of (human) agency may be highly relevant to the extent to which a nominal is semantically flexible.

## Acknowledgements

Firstly, I would like to express an enormous thank you to my supervisor Lise Fontaine, for everything she has done for me over the last five years. The last three years, in particular, have been anything but normal, and weirdly, I have now known Lise longer online than I have in-person, due to the Covid-19 pandemic and hybrid working. I am especially grateful for Lise's unwavering support and guidance and how she has always encouraged me to actively pursue my own research interests. Moreover, Lise has continuously provided me with academic opportunities to develop my research abilities, such as working in separate research projects and running a postgraduate research group. She is evidently a brilliant supervisor, given the number of students she has successfully supervised. Perhaps the thing I value most about Lise's supervision, apart from our relationship as friends, is that Lise kept a distance that allowed me to make my own mistakes and effectively learn from them, making me a more efficient and confident researcher. I could not wish a better supervisor. Thank you once again!

Thank you to all the staff at ENCAP who have supported me throughout my time at Cardiff University. I would like to give a special mention to Dr Katy Jones, Dr Tereza Spilioti and Dr Virpi Yläanne for their encouragement and advice when I was an undergraduate. Moreover, I owe a huge thank you to Dr Andreas Buerki and Dr Sean Roberts for their support with various quantitative methods, and for fostering an enjoyable work environment that has facilitated my development as a seminar tutor.

I would also like to express my gratitude to the staff external to Cardiff University that my work has been influenced by. Thank you to Prof Liesbet Heyvaert for our interesting discussions on the direction of my thesis and for being so helpful when I sadly had to leave KU Leuven, due to the pandemic. Thanks also to Dr Laure Sarda for her interesting contributions in the Nominal Group Group that certainly helped advance my understanding of nominal semantics and the framework of lexical aspect.

Before University, I was very fortunate to have many excellent English teachers. Specifically, I would like to thank Trevelyan May for his sheer enthusiasm and the undeniable effort he puts into his work. It was in his year-six classroom that my interest in language was first truly established. I would also like to thank Nicola Carcone for her incredibly captivating (and

distinctly northern) English Language and Literature classes in college. In these classes, I significantly developed my interest in analysing what people do with language, which has stuck with me ever since.

Thank you to the Economic and Social Research Council (ESRC) for funding this 1+3 project, in addition to generously providing a three-month extension, due to the effect of the pandemic on this work.

To my PhD Peers – Aurora Goodwin, Axel Bergström (the Swedish wizard), Debbie Cabral (Debbster), Ellie Bristow, Emma Collier, Katharine Young and Lucy Chrispin – I am very grateful for all the times we have shared together, and that there is now a solid group of PhD linguistics students that actually hang out and talk to one another! Long may this continue. A special thanks to Lucy Chrispin for providing a second eye to my analyses and for being a helpful mate, as we both trudged through our PhDs whilst forced to work online during the pandemic. I would also like to give special thanks to Emma Collier for checking the clarity of my thesis, and for giving unwanted grammatical analysis on my tweets.

To all my friends outside of the PhD, thank you for all the laughs we have had together. Unfortunately, there are too many of you to name, but you know who you are.

Lastly, I want to give a massive thank you to my mum, dad, and Georgina for always being there for me and for all the encouragement and praise they have given me throughout the PhD process. Specifically, thank you mum for always proofreading my work for typos since college - I hope you have learnt a thing or two about nominal semantics! Dad, unlike mum, you have contributed nothing towards this thesis. If anything, you have actually been a hinderance at times, interrupting my work consistently to tell me irrelevant pieces of news from the Southern Daily Echo. However, I wouldn't have it any other way.

# Contents

<b>ABSTRACT</b> .....	<b>I</b>
<b>ACKNOWLEDGEMENTS</b> .....	<b>II</b>
<b>CONTENTS</b> .....	<b>IV</b>
<b>LIST OF TABLES</b> .....	<b>VIII</b>
<b>LIST OF FIGURES</b> .....	<b>IX</b>
<b>Chapter 1: Introduction</b> .....	<b>1</b>
1.1. Aims of the Thesis .....	3
1.2. Motivating Research into the Nature and Degree of Nominal Semantic Behaviour .....	4
1.2.1. Temporal Semantics in the Nominal Domain.....	5
1.2.2. Lexical Aspect and Nominality .....	6
1.3. Summary of Research Gap .....	10
1.4. Structure of this Thesis .....	11
<b>Chapter 2: Understanding Categorisation</b> .....	<b>14</b>
2.1. Categorisation and Lexical Class.....	15
2.1.1. Category Structure: The Classical Model .....	15
2.1.2. Similar, But Not Exactly the Same .....	16
2.1.3. Category Structure: Prototype Theory .....	17
2.1.4. The Consequence of Oversimplified Feature Lists.....	21
2.2. Distributional Approaches to Nouns and Verbs .....	26
2.2.1. Cognitive Grammar and Useful Terminology .....	26
2.2.2. The Billiard-Ball Model.....	27
2.2.3. An Abstract Characterisation.....	29
2.2.3.1. Nominal Conceptualisation in Cognitive Grammar .....	30
2.2.4. Nouns and Verbs: Distinctions of Discourse Functionality.....	31
2.2.5. Semantic Class and Pragmatic Function.....	33
2.2.6. Categorising Nouns and Verbs Summary.....	34
2.3. Construal and Meaning Potential.....	36
2.3.1. Construal .....	36
2.3.2. Meaning Potential .....	40
2.4. Lexical Aspect .....	42
2.4.1. Grammatical Aspect and Lexical Aspect.....	43
2.4.2. Lexical Aspect Features .....	45
2.4.3. Situation Types .....	47
2.4.4. The Inclusion of Nominals.....	49
2.5. Summary and Aims.....	51
<b>Chapter 3: Methodology for Investigating Category Feature Influence on Nominal Semantics</b> .....	<b>53</b>
3.1. Corpus and Sample .....	54
3.2. The Nominal Annotation Scheme.....	57

3.2.1. SFL Annotation of Nominal Instances .....	57
3.2.2. Lexical Aspect Annotation of Nominal Instances .....	59
3.2.3. Count/Mass and Concrete/Abstract Status Annotation.....	61
3.3. Reductions and Repetition in the Sample .....	64
3.3.1. Reductions to the Sample.....	64
3.3.2. Repetition in the Sample.....	67
3.4. Word Formation Annotation.....	68
3.4.1. Inflection vs Derivation .....	68
3.4.2. Word formation Categories.....	70
3.4.2.1. Hybridity in the Sample.....	72
3.5. Dataset Preparations for Analyses .....	72
3.6. Statistical Analyses .....	74
3.6.1. Fisher-Exact Tests and Cramer’s V Test .....	74
3.6.2. Classification and Regression Tree Analysis.....	77

**Chapter 4: Investigating Category Feature Influence on Nominal Semantics - Results.80**

4.1. Individual Influences on Nominal Semantic Behaviour.....	80
4.1.1. The Relationship Between Word Formation Type and Situation Type.....	81
4.1.1.1. The Influence of Repetition on Situation Type Distribution across Word Formation Types .....	84
4.1.1.2. The Influence of Object expression on Situation Type Distribution across Word Formation Types .....	85
4.1.2. The Relationship Between Count/Mass Status and Situation Type.....	86
4.1.2.1. The Influence of Repetition on Situation Type Distribution across Count and Mass Nominal Instances .....	89
4.1.2.2. The Influence of Object expression on Situation Type Distribution across Count and Mass Nominal Instances.....	89
4.1.3. The Relationship Between Abstract/Concrete Status and Situation Type.....	90
4.1.3.1. The Influence of Repetition on Situation Type Distribution across Abstract and Concrete Nominal Instances .....	92
4.1.3.2. The Influence of Object expression on Situation Type Distribution across Abstract and Concrete Nominal Instances.....	92
4.1.4. The Relationship Between Genre and Situation Type.....	93
4.1.4.1. The Influence of Repetition on Situation Type Distribution across Genre .95	
4.1.4.2. The Influence of Object expression on Situation Type Distribution across Genre 96	
4.2. Simultaneous Influences on Nominal Semantic Behaviour.....	96
4.2.1. CRT Analysis:.....	97
4.2.1.1. The Influence of Repetition on Situation Type Distribution across All Predictor Variables.....	104
4.2.1.2. The Influence of Object expression on Situation Type Distribution across All Predictor Variables. ....	106
4.3. Concluding Comments.....	109

**Chapter 5: Investigating Category Feature Influence on Nominal Semantics –**

<b>Discussion.....</b>	<b>111</b>
5.1. Abstract/Concrete Status: The Most Influential Variable.....	111
5.1.1. Abstract/Concrete Status and Situation Type Distribution.....	112
5.2. Count/Mass Status and Semantic Behaviour .....	118
5.2.1. Recapping Concepts: Count vs Mass.....	118

5.2.2.	Count/Mass Status: An Unequal Influence .....	122
5.2.3.	Count Activities and Count States .....	127
5.3.	Word Formation Type: The Influence of MDA; MDV and TC .....	132
5.3.1.	The Influence of MDA on Situation type Expression .....	135
5.3.1.1.	The Semantic Behaviour of -ness and -ity Suffixation in MDAs .....	135
5.3.1.2.	MDAs and the Expression of Events .....	136
5.3.2.	The Influence of MDV on Situation type Expression .....	138
5.3.2.1.	The Dominance of -er Suffixation in Concrete Count MDVs.....	140
5.3.3.	The Semantic Variation of TC Nominals .....	142
5.3.3.1.	The Situation Type Expression of Abstract Count TCs .....	143
5.3.3.2.	The Situation Type Expression of Concrete TCs .....	145
5.4.	Summary: Events, States, Objects and Evaluating Nominality .....	147
<b>Chapter 6: Underived Nominals: Methodology and Results .....</b>		<b>151</b>
6.1.	Corpus and Sample .....	152
6.1.1.	Reductions across the Four Samples.....	153
6.2.	The Annotation of the Four Underived Nominals .....	154
6.2.1.	Definiteness Annotation.....	154
6.2.2.	SFL Process Type Annotation .....	156
6.2.3.	Lexical Aspect Annotation .....	160
6.3.	Statistical Analysis.....	162
6.3.1.	Multinomial Logistic Regression.....	162
6.4.	Results.....	165
6.4.1.	Game .....	165
6.4.2.	Fire .....	171
6.4.3.	Stream .....	177
6.4.4.	Drought .....	182
6.5.	Towards an Exploration into Semantic Flexibility .....	183
<b>Chapter 7: Semantic Flexibility, Stability, and the Boundaries of Nominality.....</b>		<b>184</b>
7.1.	Meaning Potential and Semantic Structure.....	185
7.2.	The Semantic Flexibility of Game .....	186
7.2.1.	Hyponymy and Rules: The Difference of Accomplishment and Activity.....	187
7.2.2.	Video Games and the Conceptualisation of Lexical Facets.....	190
7.3.	Fires, Droughts, and Semantic Stability .....	193
7.3.1.	Natural Events and Human Agency.....	193
7.3.2.	Accidents, Arson, and Atelicity .....	195
7.4.	Stream: Data Transfer and Contextual Dependency.....	197
7.5.	Chapter Summary .....	199
<b>Chapter 8: Conclusion: Evaluating the Nature and Degree of Nominality.....</b>		<b>203</b>
8.1.	Events, States and UENs.....	204
8.1.1.	Nominal Feature Importance and Interaction .....	204
8.1.2.	The Influence of Word Formation on Nominal Semantics.....	206
8.1.3.	Boundedness, Homogeneity and Telicity in the Nominal Domain .....	208
8.1.4.	The Semantic Behaviour of Underived Event Nouns.....	209
8.2.	The Nature and Degree of Nominality.....	213
8.3.	Limitations .....	216
8.4.	Future research.....	218

<b>REFERENCES.....</b>	<b>220</b>
<b>APPENDIX A: STUDY ONE DATASETS .....</b>	<b>244</b>
<b>APPENDIX B: STUDY ONE LEXICAL ASPECT DISTRIBUTION PERCENTAGE TABLES, POST-HOC FISHER-EXACT TESTS AND SIMPLIFIED CRT ANALYSES .....</b>	<b>245</b>
<b>APPENDIX C: STUDY TWO DATASETS .....</b>	<b>259</b>
<b>APPENDIX D: R SCRIPTS GITHUB.....</b>	<b>260</b>



## List of Tables

Table 1-1. Lexical aspect features of situation types with examples (Smith 1991). .....	8
Table 2-1. Different types of phenomena in the ideation base, provided with examples (Halliday and Matthiessen 1999, pp.48–59). .....	38
Table 2-2. Lexical aspect features of situation types (Smith 1991).....	48
Table 3-1. Distribution of nominal instances collected in relation to their genre.....	57
Table 3-2. DSTs for lexical aspect annotation and their literature source. ....	61
Table 3-3. DSTs for count/mass status and abstract/concrete Status and their literature source. ....	64
Table 3-4. Breakdown of instances incompatible with lexical aspect annotation. ....	65
Table 3-5. The three different datasets used in the study accompanied with the number of instances they each contained. ....	73
Table 4-1. Post-hoc Fisher-Exact test results on comparisons of semantic behaviour expressed by different word formation types.....	84
Table 4-2. Post-hoc Fisher-Exact test results on comparisons of semantic behaviour expressed by different genres.....	95
Table 6-1. Frequency of occurrence for four underived nominals across the Timestamped JSI Web Corpus 2014-2021 English. ....	153
Table 6-2. A selection of syntactic probes from Fontaine (2013, pp.85–91) for the analysis of process type.....	160
Table 6-3. DSTs for lexical aspect annotation and their literature source. ....	161
Table 6-4. Coefficients and std. errors produced for game by situation type ~ definiteness + process type.....	166
Table 6-5. Prediction rate and confidence intervals produced for game by situation type ~ definiteness + process type. ....	168
Table 6-6. Coefficients and std. errors produced for fire by situation type ~ definiteness + process type.....	172
Table 6-7. Prediction rate and confidence intervals produced for fire by situation type ~ definiteness + process type. ....	173
Table 6-8. Coefficients and std. errors produced for stream by situation type ~ definiteness + process type.....	178
Table 6-9. Prediction rate and confidence intervals produced for stream by situation type ~ definiteness + process type. ....	179
Table 6-10. Distribution of situation type across definiteness for drought.....	182
Table 6-11. Distribution of situation type across process type for drought.....	182

## List of Figures

Figure 2-1. Four types of prototypicality effects (Geeraerts 2010, p.189). .....	20
Figure 2-2. The profile of circle (indicated by the circle shape) within its domain of two-dimensional space (indicated by the surrounding box) (Langacker 1987a, p.184). .....	29
Figure 3-1. Three types of experiential metafunction analysis, provided with context and example (Noun 83, All Data without Repetition Dataset, Appendix A). .....	59
Figure 4-1. Barplot of situation type proportion expressed by word formation type. ....	82
Figure 4-2. Barplot of situation type proportion expressed by count and mass nominal instances. ....	88
Figure 4-3. Barplot of situation type proportion expressed by abstract and concrete nominal instances. ....	91
Figure 4-4. Barplot of situation type proportion expressed by genre. ....	94
Figure 4-5. All Data without Repetition decision tree visualising situation type ~ word formation type + count/mass status + abstract/concrete status + genre. ....	99
Figure 4-6. All Data with Repetition decision tree visualising situation type ~ word formation type + count/mass status + abstract/concrete status + genre. ....	105
Figure 4-7. <i>No Object Data without Repetition decision tree visualising situation type ~ word formation type + count/mass status + abstract/concrete status + genre.</i> .....	108
Figure 5-1. Barplot of situation type and object meaning proportion expressed by abstract and concrete nominal instances. ....	113
Figure 5-2. Decision tree visualising situation type ~ word formation type + count/mass status + abstract/concrete status + genre. ....	114
Figure 5-3. Distinction of boundedness between count and mass nominals (Langacker 2008, p.133). ....	120
Figure 5-4. Barplot of situation type proportion expressed by count and mass nominal instances. ....	123
Figure 5-5. Decision tree visualising situation type ~ word formation type + count/mass status + abstract/concrete status + genre. ....	124
Figure 5-6. Barplot of situation type proportion expressed by word formation type. ....	133
Figure 5-7. Decision tree visualising situation type ~ word formation type + count/mass status + abstract/concrete status + genre. ....	134
Figure 6-1. Definite and definite analysis, taken from stream dataset (Stream Dataset, Appendix C). ....	155
Figure 6-2. Prediction rate and confidence intervals of definiteness across situation type for game. ....	169
Figure 6-3. Prediction rate and confidence intervals of process type across situation type for game. ....	169
Figure 6-4. Prediction rate and confidence intervals of definiteness across situation type for fire. ....	174
Figure 6-5. Prediction rate and confidence intervals of process type across situation type for fire. ....	175
Figure 6-6. Prediction rate and confidence intervals of definiteness across situation type for stream. ....	180
Figure 6-7. Prediction rate and confidence intervals of process type across situation type for stream. ....	180

# Chapter 1: Introduction

The classification of lexical items into distinct categories is perhaps one of the oldest practices in the field of linguistics (Baker 2003, p.1). However, their exact classification has proven particularly problematic for linguists due to the dynamic nature of language use (Van Lier and Rijkhoff 2013, p.3). This problem is easily seen in classes such as ‘noun’ and ‘verb’. While it is widely accepted that noun and verb represent two separate lexical classes, the distinction has been shown to be inadequate, due to the dependence of these lexical classes on a combination of morphological, syntactic, and semantic criteria, which do not coincide in every instance (Lyons 1977, p.423). For example, most students who attended a British school will recall being told that a noun is a ‘person, place or thing’ (Derewianka 2020, p.825). In this definition specifically, ‘thing’ can relate to anything, whether that be real or imaginary. As Fontaine (2013, p.26) notes, this semantic conception of nouns, although operative, is quite inefficient, as it provides far too much scope to the semantic behaviour of nominals, which, consequently, produces a rather vague definition.

Throughout the last paragraph, you might have noticed I used the term ‘nominal’ instead of noun. Nominality is a term used to “capture the condition of being nominal” (Fontaine 2017a, p.23). By taking this approach to lexical class, we can talk about nominality, and therefore nominals, as something gradable (more nominal to less nominal) (Aarts 2004b), dependent upon the expression of nominal category features. Such an approach promotes inclusion during the categorisation process and fundamentally relates to perhaps the most popular current theory of categorisation: ‘Prototype theory’.

Prototype theory is a theory of categorisation that was largely formulated in opposition to the ‘Classical model’ of categorisation. Consequently, to understand Prototype theory and the motivation for its formulation, I must very briefly discuss the Classical model. The Classical model is an approach to categorisation that views categories as clearly distinguished from one another and understands classification as a binary process based on a strict set of category features (van Der Auwera and Gast 2011, p.170). As I will discuss at greater depth in Chapter 2, this idea of rigid classification is inherently problematic, because categorisation is rarely a black and white matter. For instance, *house* and *salt* are both morphologically and semantically similar (they are both objects), but they syntactically differ, as *house* is generally expressed as

something countable (e.g., two houses) and *salt* is generally expressed as a mass (\*two salts<sup>1</sup>). Although these lexical items differ, this difference in count/mass status does not dictate that one of these lexical items cannot be classified as nominal.

In opposition to the Classical model is Prototype theory. Prototype theory is a series of ideas that suggest a more flexible classification alternative to the Classical approach, founded on four ‘prototypicality effects’. The first is known as ‘graded centrality’. Pioneered by Rosch and her colleagues (Rosch 1973; Rosch and Mervis 1975; Rosch 1978), graded centrality relates to the idea that not all category members are equally representative of the category they are classified under. As a result, Prototype theory does not view all its members as equal and suggests a graded category structure on which members can be placed from prototypical to least prototypical (Aarts 2004b). The second prototypicality effect is that no category feature is deemed essential for category membership (Geeraerts 2006, p.142). Unlike the Classical model, Prototype theory does not impose strict category boundaries in relation to any category features, so category members can express multiple different features from one another but still belong to the same category.

Building on the second prototypicality effect, the third is that the boundaries of category membership are contemplated as ‘fuzzy’ instead of rigid, allowing for the categorisation of more peripheral category members. The final prototypicality effect is that prototype categories exhibit ‘family resemblance structures’: their semantic structure is composed of a “radial set of clustered and overlapping readings” (Geeraerts 2010, p.187). Essentially, what this effect means is that although every category member will express at least one feature in common with another member, there are very few features, if any, that are shared by all members (Rosch and Mervis 1975, pp.574–575).

By understanding these prototypicality effects, we see that Prototype theory offers a flexible alternative of categorisation that can account for instances that may not express every category feature, but still express some. Moreover, Geeraerts (2006, p.144) suggests Prototype theory is particularly effective, as it is able to capture our pre-theoretical intuitions that certain category members are inherently ‘better’ example members than others. For instance, in relation to the lexical category noun, Ross (1972) suggests a prototypical scale of nominality. On this

---

<sup>1</sup> Throughout this paper, an asterisk will be used to indicate a non-typical use of syntax.

continuum, Ross claims that nouns such as *house* express the most nominality (i.e., are the most prototypical). Alternatively, Ross claims that while ‘*that*-clauses’ (e.g., I said that I told him) do express a certain amount of nominality due to their capacity to occupy the same syntactic roles in clauses as nominal groups, they express the least nominality. Such a scale initially plays to our intuitions about nominal category membership, as it is easy to comprehend that *that*-clauses would be classed as more peripheral category members.

However, once we encounter some more complicated nominal examples, Ross’ scale of nominality begins to unravel. For instance, certain nominals can be morphologically underived but express event meaning (otherwise known as underived event nominals (UENs) which will be explored throughout this thesis), e.g., a *storm* can take place over a period of time. Instances like these are incredibly interesting, as nominality has typically been associated with the expression of atemporal meaning (Givón 2001, p.51; Halliday and Matthiessen 2014, p.236), but many nominals, including UENs, go against that trend. As a result, UENs are extremely hard to accurately place on continuums of nominality, as there exists an incongruence between their nominal form and semantic behaviour. We might initially try to categorise nominals like these as more peripheral. Nevertheless, because of the underspecification (fuzzy nature) of category boundaries, Geeraerts (2006, p.149) suggests that the classification of peripheral instances is always rather ambiguous. In addition to this ambiguity, there have been several other critiques that fundamentally question how prototypicality is established and I will briefly address these in Section 1.2.

Before we get ahead of ourselves on the efficiency of Prototype theory in accounting for nominal category structure, Section 1.1 will introduce the principle aim and sub-aims of this thesis, which will act to better contextualise this research, and lead to a greater understanding of the research motivations detailed in Section 1.2.

## **1.1. Aims of the Thesis**

As I will show in the following section and extensively throughout Chapter 2, there is still a great deal of work needed to refine our classification of nominality, particularly in relation to our classification of its semantic behaviour. Consequently, the central aim of this thesis is to carry out a detailed examination into the expression of nominal semantics and how different temporal meanings are expressed. To achieve this aim, three sub-aims are outlined:

1. To determine how event and state meaning come to be expressed in nominal forms.
2. To examine the potential relationships between the syntactic and semantic properties of underived event nominals (UENs).
3. To determine how the nature and degree of nominality can be evaluated in English.

## **1.2. Motivating Research into the Nature and Degree of Nominal Semantic Behaviour**

As shown by the semantic behaviour of UENs, we find a lack of clarity from lexical class prototype models once we begin to categorise instances that might not be deemed prototypical. Aside from this critique, Prototype theory has been criticised for several different reasons, each concerning the inefficiency of oversimplified feature lists (Croft and Cruse 2004, p.87) to accurately account for how prototypicality is established. Specifically, the critiques most central to this research concern (1) the apparent inability of Prototype theory to account for context in what is understood as prototypical (Croft and Cruse 2004, p.87); (2) the inadequate amount of attention Prototype theory gives to category feature importance in the establishment of prototypicality; and (3) the lack of attention Prototype theory places on category feature interactions in the establishment of prototypicality. These critiques will be discussed in greater detail in Section 2.1.4, where we will cover the adequacy of Prototype theory in accounting for nominal semantic behaviour and the expression of prototypicality. In the following paragraph, I will now discuss how exploration into the semantic behaviour of nominals can provide a beneficial foundation into understanding the nature and degree of nominality. The ‘degree’ of nominality, throughout this thesis, will relate to the extent to which semantic behaviour can range within the nominal domain.

When discussing prototypicality in relation to nominality, the semantic behaviour expressed by nominals presents a particularly interesting research avenue for several reasons. Firstly, as nominals are generally on the front line of making provisions in language for new entities that come into existence (Halliday 1966/2003, pp.53–54), the semantic expression of nominals is highly variable and ever-evolving. However, to my knowledge, there has yet to be a large-scale empirical study that explicitly researches the distribution of semantic behaviour by nominals

or the extent to which certain nominal category features simultaneously influence semantic expression. Secondly, previous work on nominal semantics has suggested that nominals prototypically express atemporal (object) meaning and are highly time-stable (Langacker 1991, pp.13–14; Givón 2001, p.51; Halliday and Matthiessen 2014, p.236). Such claims, while certainly intuitive, lack substantial evidence from corpus data, so are prime for empirical investigation. Moreover, a growing amount of work has shown nominals can express temporal semantics, such as events and states. Accordingly, the semantic behaviour of nominals provides a rich foundation to explore what prototypicality is in the nominal domain and the influence that certain nominal category features hold over the establishment of prototypical semantic behaviour. In the following paragraphs we will now very briefly review the expression of temporal semantic behaviour in the nominal domain, before specifically focusing on the framework of ‘lexical aspect’ (Vendler 1967) as a key resource for the research of this thesis.

### 1.2.1. Temporal Semantics in the Nominal Domain

Perhaps the best-known seminal work into nominal temporal semantic behaviour comes from Grimshaw (1990), who proposed a three-part distinction of deverbal nominals: ‘result nominals’; ‘simple event nominals’; and ‘complex event nominals’. An example of each of these nominals, taken from the work of Alexiadou and Grimshaw (2008, p.2), is presented below in (1a-c). You will notice that *examination* is a particularly interesting instance, as it is three-way ambiguous between the different readings.

- (1) a. The examination on the table. (Result Nominal)  
 b. The examination took a long time. (Simple Event Nominal)  
 c. The examination of the patients took a long time. (Complex Event Nominal)

The fundamental difference between these nominals relies on the expression of events and whether the nominal expresses an ‘argument structure’, i.e., “the lexical representation of grammatical information about a predicate” (Grimshaw 1990, p.1). Result nominals, like in (1a), do not express events and are consequently atemporal. Simple event nominals, like (1b), do express events, but do not express an argument structure, as they do not permit event-related prepositional phrases (Bloch-Trojnar and Malicka-Kleparska 2017, p.12). Lastly, complex event nominals, as in (1c), express events and an argument structure. It should be noted here that these three terms will not be carried on throughout the rest of this thesis, as the semantic

expressions of nominal instances in this research will be analysed using the framework of ‘lexical aspect’. This framework will be briefly outlined in the following paragraphs, but also more comprehensively in Section 2.4.

In addition to the work of Grimshaw (1990) on argument structure and deverbal nominalisations, more recent research has also begun to show that a variation of different nominals, ranging from nominals morphologically derived from adjectives (Arche and Marin 2021) to morphologically underived nominals (Huyghe et al. 2017), are capable of expressing events. In doing so, a small but growing amount of research into the semantic behaviour of nominals has typically drawn upon the framework of lexical aspect as an analytical resource to assess this semantic behaviour.

### **1.2.2. Lexical Aspect and Nominality**

Lexical aspect relates to the ‘inherent’ temporal semantic meaning expressed by individual lexical items (Smith 1997). This semantic meaning has traditionally concerned the existence of temporal restrictions and boundaries within the lexical composition of verbs (Filip 2012, p.721). However, and like we will see in later paragraphs, research has begun to explore lexical aspect expressed by other lexical classes, such as nouns (Balvet et al. 2011; Fábregas and Marín 2012; Huyghe et al. 2017; Heyvaert et al. 2019; Lieber and Plag 2022). The inherent temporal semantics posed by the lexical aspect framework revolve around the expression of three central features of lexical aspect: dynamism; durativity; and telicity. When combined in different formats, these features produce distinct situation types, which will be described later in this section, after detailing the different lexical aspect features.

To help explain these lexical aspect features, I will use three example verbs, shown below in (2a-c). The first feature is dynamism, which distinguishes dynamic situations from static situations. Dynamic situations typically involve some kind of “change, motion or activity”, which typically requires a continuous input of energy from an agent external to the process (Comrie 1976, p.49). Static situations alternatively, represent unchanging situations which hold over time and do not require an input of energy from an external agent to persist. This difference is shown in example (2a-b) below. While (2a) dynamically takes place over time, (2b) does not express dynamism. It, instead, holds over time, as a continuous input of energy is not required for Jack’s love of ice-cream to continue.



- (2) a. Jack is swimming.  
b. Jack loved ice-cream.  
c. Jack won the race.

The second lexical aspect feature is durativity. Durativity marks the distinction between durative and punctual situations. Durative situations transpire over a period of time, while punctual situations are instantaneous, single-phase events. For instance, if we return to (2a), we see that swimming is a durative situation, as the event exists through time and has many distinguishable phases, i.e., each phase of the swimming event is different. We can contrast this example with the punctuality expressed in example (2c). Winning a race is generally an instantaneous event, that only requires one moment for its actualisation.

Lastly, we have the lexical aspect feature telicity. Telicity concerns whether a situation expresses an inherent temporal endpoint in its temporal structure or not, i.e., whether a situation is telic or atelic. Telic situations are regarded as having this inherent temporal endpoint, while atelic situations are conceptualised as continuous, with any ending considered arbitrary. For instance, (2a) represents an atelic situation, as we are unable to determine when *Jack's swimming* will end. Moreover, any ending to his swimming will not be the result of internal temporal structure of the event. Alternatively, (2c) is a telic situation, as once Jack wins the race, the situation is terminated – the situation has reached its inherent temporal endpoint.

Now I have outlined the different features of lexical aspect, we can move on to how they combine to create different situation types. The concept of situation type, or 'aspectual class', was proposed by Vendler (1957; 1967) who discerned four 'time schemata' (situation types), which each encompass a combination of the abovementioned lexical aspect features. Vendler's (1957; 1967) four time schemata are States; Accomplishments; Achievements; and Activities. In addition to these four, Smith (1991) proposed a fifth category, Semelfactives. Based on the situation's dynamism (Smith 1997), the five situation types are divided into two categories. These categories are state situation types (States) and event situation types (Accomplishments, Achievements, Activities and Semelfactives). In the following paragraph, I will describe the differences between each of these situation types, with verb examples provided in Table 1-1.

Firstly, States constitute the only static situation type, as they require the presence of an external agent to induce change. They are also durative, as they typically hold for extended periods of time, although this has been debated (cf. Murphy 2010, p.210). However, they are neither viewed as telic or atelic as States do not take time, they simply hold for a period of time (Smith 1997, p.32). So, even though States are represented as continuous, the initial and final endpoints of a State do not form part of the State itself. Activities, contrastingly, are dynamic events that involve “movement, activity and volition” over a duration of time (Smith 1997, p.24). Activities are also atelic, as these events have no inherent temporal endpoint.

Accomplishments, like Activities, are also dynamic and durative, but they differ in relation to their telicity, expressing telic events. Accomplishments express inherent temporal endpoints and consequently are not construed as continuous. Achievements, on the other hand, are highly dynamic events that are punctual, due to their instantaneous character. They are also telic, since the capacity of the event to expeditiously take place in a single phase naturally requires the expression of an inherent temporal endpoint. The last situation type is Semelfactive, which, according to Smith (1991), represents dynamic, punctual and atelic events. They are single-stage events which do not result in an outcome other than the expression of the event itself (Smith 1997, pp.29–30).

*Table 1-1. Lexical aspect features of situation types with examples (Smith 1991).*

<b>Situation Type</b>	<b>Dynamic/ Stative</b>	<b>Durative/ Punctual</b>	<b>Telic/ Atelic</b>	<b>Example</b>
States	Stative	Durative	N/A	Jack loved ice-cream
Activities	Dynamic	Durative	Atelic	Jack is running.
Accomplishments	Dynamic	Durative	Telic	Jack swam the English Channel.
Achievements	Dynamic	Punctual	Telic	Jack jumped.
Semelfactives	Dynamic	Punctual	Atelic	Jack coughed.

As mentioned above, research into the expression of temporal semantics has only recently begun to explore the relevance of lexical aspect in the nominal domain. Huyghe (2017, p.118) suggests that lexical aspect is not a verb specific concept, and is principally dependent on the expression of eventualities, regardless of lexical class. This claim is supported by the semantic behaviour of certain nominals, as events and states can be expressed in the nominal domain.

State nominals, like their verbal counterparts, are static entities, which typically transpire over time, such as example (3a). On the other hand, event nominals, which encompass the lexical aspect feature combinations of Activities, Accomplishments, Achievements and Semelfactives, are dynamic nominals which occur or take place, such as example (3b). Unlike verbs, however, nominals also display the unique capacity to express ‘Objects’ which are atemporal expressions that do not express lexical aspect (Fábregas and Marín 2012, p.36). An example of an Object is available below in example (3c).

- (3) a. The love.  
b. The swim.  
c. The rock.

Research into the expression of lexical aspect in nominals is a somewhat emerging field, with work largely focused on whether nominals of different word formation type, such as deverbal nominalisations (Balvet et al. 2011; Fábregas and Marín 2012; Lieber and Plag 2022), deadjectival nominalisations (Arche et al. 2021) and morphologically underived nominals (Huyghe et al. 2017), can express temporal semantics. Interestingly, however, little research (with the exception of Heyvaert et al. (2019) and Lieber and Plag (2022)) has worked with corpus data, analysing language in its actual use.

For instance, Huyghe et al. (2017) provide a compelling account of the ability of morphologically underived nominals to express lexical aspect features. From a sample size of 3489 morphologically underived nominals, collected from the *Lexique3* lexicon ([www.lexique.org](http://www.lexique.org)), they identified 283 underived nominals that were able to express at least one lexical aspect event feature (e.g., either dynamism, durativity, or telicity) after analysing each morphologically underived nominal in a series of diagnostic syntactic tests. However, while Huyghe et al. provide clarity on the lexical aspect features certain underived nominals can express, their paper does not investigate potential patterns of situation type expression displayed across the underived nominals. Moreover, as the nominals used in Huyghe et al.’s study were taken from lexicon and were therefore isolated instances, the underived nominals were not analysed for their exact meaning in use. Consequently, a proportion of the lexical aspect features exhibited by certain nominals, while syntactically acceptable, may have been largely uncharacteristic.

### **1.3. Summary of Research Gap**

Accordingly, what we have seen over this chapter is that although prototype categories provide a flexible alternative to more rigid classification, there is a considerable lack of attention given to the different meanings that nominals can express and how prototypicality is accurately established. This issue is prevalent in our categorisation of nominality, as while we are aware that expression of atemporal semantics is prevalent among nominals, we are currently unaware of which category features are more central to prototypical nominality and how category features may combine in the expression of different nominal meanings. Moreover, at this moment, there is also an evident lack of emphasis placed on investigating the semantic boundaries, i.e., the degree, of nominality. This shortfall in the literature results in a loss of accuracy when categorising nominals such as UENs, which express temporal semantics despite expressing prototypical nominal form. As discussed above in Section 1.2, the semantic behaviour of nominals provides an advantageous base from which to explore prototypicality in the nominal domain, as the semantic expression of nominals is highly dynamic, and we also already hold intuitive presuppositions about prototypical nominal semantic behaviour. Therefore, it will certainly be interesting to explore how prototypicality is established semantically in the nominal domain and how different nominal category features influence semantic behaviour.

Moreover, while studies using lexical aspect in the nominal domain have largely focused on the identification of temporal semantics in different nominal forms, significantly less work has been devoted to the examination of the specific character of these temporal semantics across different nominal forms, and the syntactic contexts which support them. For example, the temporal semantics expressed by deadjectival nominalisations may typically express States more than deverbal nominalisations. Furthermore, given their problematic nature in relation to their classification, it is surprising that a study has yet to be composed which explicitly focuses on the temporal semantic expression of morphologically underived nominals in their actual use.

Having outlined the motivation for this research, the following section will now present the structure of the thesis moving forward. The section will cover the specific details of each chapter and will provide a focus on how the sub-aims of this work (noted in Section 1.1) were incorporated into the research design of two distinct studies.

## **1.4. Structure of this Thesis**

In Chapter 2, Section 2.1, I will expand on the notion of categorisation and its relationship with lexical class by first outlining in detail the Classical model of categorisation (2.1.1) and then Prototype theory (2.1.3). I will examine each theory with a critical lens, noting their advantages and disadvantages when applied to lexical class categorisation. After our discussion of Prototype theory and the issues it faces in the nominal domain, Section 2.2 provides a critical overview of perhaps the three most notable examples of Prototype theory in the categorisation of nouns and verbs: Langacker's (1991) theory of Cognitive Grammar; Hopper and Thompson's (1984a) distinctions in discourse functionality; and Croft's (1991) categorisation of semantic class and pragmatic function. The chapter then moves away from theories/models of categorisation, to focus on the semantics of nominals and the dynamism of meaning. Specifically, section 2.3 first picks apart the notion of construal from the perspectives of Langacker's (1987a; 1991) Cognitive Grammar and Halliday's (1985) Systemic Functional Linguistic (SFL) framework, before describing why a combined perspective of construal will be adopted for the rest of the thesis.

Section 2.3 then turns to Hanks' (2013) concept of 'meaning potential', which provides an intuitive heuristic in relation to how meaning comes to be expressed by lexical items in use. Hanks' concept of meaning potential is particularly relevant to the current thesis as it places great emphasis on understanding how different semantic meanings can be expressed and the influence of the surrounding syntactic context on this expression. Section 2.4 then reintroduces the framework of lexical aspect as an empirical resource to facilitate access to the semantic behaviour of nominals. While sections 2.4.2 and 2.4.3 are dedicated to teasing apart different lexical aspect features and situation types, section 2.4.4 places emphasis on how the framework of lexical aspect has so far been applied in the nominal domain. The chapter concludes by drawing attention to the overall lack of coverage lexical aspect has received in the nominal domain and how very few studies have explored the influence nominal category features, such as word formation, may have on the lexical aspect expression of nominals. Most importantly, the end of the chapter notes how an investigation into the lexical aspect expression of nominals would provide empirical data to help inform our understanding of nominal semantic behaviour more accurately.

In Chapter 3, I will discuss how the remainder of the thesis will be structured in relation to two distinct studies. The first study will examine the individual and simultaneous influence of nominal category features on the temporal semantic expression of nominals. The second study will take a smaller-scale approach, by assessing the extent to which the temporal semantic expression of four morphologically underived nominals (*drought*; *fire*; *game*; and *stream*) shares a relationship with surrounding syntactic context (definiteness; process type of the clause). Chapters 3-5 are dedicated to the first study of the thesis, while Chapters 6-7 are dedicated to the second study. This structure is of course slightly unusual. The main motivation for this structure was to provide two distinct quantitative studies that together covered nominal semantic behaviour at both a macro and a micro level. As a result, I was able to attain a more general understanding of nominal semantic behaviour, whilst also investigating the individual behavioural nuances of certain nominals.

Returning back to the chapters of the thesis, Chapter 3 presents the methodology of the first study, in which nominal data from four distinct genres was taken from the British National Corpus 1960-1993 (2022) and was analysed for lexical aspect, abstract/concrete status, count/mass status and word formation type. Section 3.6 then concentrates on the statistical analyses conducted on the annotated nominals, namely Fisher-Exact tests, Cramer's V tests, and Classification and Regression Tree (CRT) analysis. After the methodology has been established, Chapter 4 outlines the results of the statistical analyses. These results will first centre on the individual influence of each independent variable (word formation type; count/mass status; abstract/concrete status; and genre) on the temporal semantic expression of the nominals. Subsequently, the following section will concentrate on the simultaneous influence of these variables on the temporal semantic expression of nominals.

Chapter 5 presents a discussion of these results, which is broken down into three main sections. Section 5.1 first centres on abstract/concrete as the most influential variable on the lexical aspect expression of nominals. Section 5.2 then moves onto the influence of count/mass status, focusing on both the interactive influence of count/mass status and the presence of countable Activities and States in the data. Lastly, Section 5.3 explores the influence of word formation type on situation type expression. This section focuses on three distinct word formation types that each expressed a notable influence over nominal semantic behaviour in the CRT analysis: nominals morphologically derived from adjectives; nominals morphologically derived from verbs, and transcategorised nominals (i.e., conversions).

Chapters 6 and 7 cover the second study of the thesis. Principally, the second study examines the relationship between the syntactic and semantic properties of UENs. Chapter 6 details the methodology of this study, noting how 500 nominal instances of four UENs (*drought*; *fire*; *game*; and *stream*) (n=2000) were taken from the Timestamped JSI Web Corpus 2014-2021 English and analysed for their definiteness, the process type of the clause the nominal was situated in (process type here relates to Halliday's (1985) categorisation of different types of process in the clause – this concept will be covered in greater detail in Section 6.2) and their lexical aspect expression. Section 6.4 outlines the statistical analyses undertaken, which included multinomial logistic regression analysis and likelihood-ratio tests. Afterwards, section 6.5 presents the statistical results of the study and provides a short discussion on each UEN in relation to the influence definiteness and process type holds over the expression of lexical aspect.

Chapter 7 then builds on the findings of Chapter 6, by discussing the situation type distributions of each UEN from a lexical semantic perspective. The chapter dedicates space to the semantic behaviour of each nominal and concentrates principally on the varying semantic flexibility of nominals. After distinguishing *game* and *stream* as semantically flexible, and *fire* and *drought* as semantically stable, sections 7.2 to 7.4 explore why the semantic flexibility of these nominals differs, and how the (non)expression of human agency interestingly influenced the semantic and syntactic behaviour of the four UENs.

Lastly, Chapter 8 compiles everything I have identified about the nature and degree of nominality. Considering the findings from both studies in this thesis, Section 8.1 considers how event and state meaning come to be expressed in the nominal domain and the relationship between the semantic and syntactic properties of UENs. Section 8.2 then covers how we can better evaluate the nature and degree of nominality moving forward. In doing so, the section places emphasis on nominal feature importance and feature interactions in the production of nominal semantics, before highlighting the influence of (human) agency on semantic expression and relationship between boundedness, homogeneity, and telicity as a dividing line between nominal and verb semantic expression. Afterwards, Section 8.3 will acknowledge the limitations of the thesis, while Section 8.4 will present avenues of future research.

## Chapter 2: Understanding Categorisation

To understand how the nature and degree of nominality is currently contemplated in English, this chapter will first establish a comprehensive background to the main theories of categorisation. In doing so, Section 2.1 will critically review the beliefs held by these theories, which will provide clarity to how they attempt to account for the categorisation of lexical classes. This section will also address the issues that arise in each theory, with an overriding focus on the semantic classification of nominals. The chapter will then lead onto a discussion in Section 2.2 on how nouns and verbs have so far been categorised, covering Langacker's (1987a; 1991) proposed distinction of nouns and verbs, rooted in the theory of Cognitive Grammar; Hopper and Thompson's (1984b) distinctions in discourse functionality; and Croft's (1991) approach to semantic class and pragmatic function. In reviewing each of these categorisations, I will explore how each way of categorisation attempts to account for the variable syntactic, morphological, and semantic behaviour expressed by nouns and verbs. The discussion of lexical categorisation is an especially relevant area to investigate as, although it is central to our understanding of nominality in English, it has proven rather problematic territory due to the highly dynamic nature of language use (Van Lier and Rijkhoff 2013, p.3).

After I have established how lexical classes such as noun and verb have typically been categorised, and the existing issues we face with their categorisation, I will then move to focus more explicitly on semantic behaviour as an avenue from which nominals can be categorised and understood. The choice to focus on semantic behaviour was made, because the semantic behaviour of nominals is highly dynamic, and, as a result, has regularly presented issues for our understanding of nominality (see sections 2.1.4; 2.2.2-2.2.5). Through using the framework of lexical aspect (which concentrates on the expression of temporal semantics (see Section 2.4), we can gain empirical access into the semantic behaviour of nominals and assess the extent to which certain category features (e.g., count/mass status) contribute to the dynamic nature of nominal semantics (see Chapters 3; 4; and 5). Accordingly, Section 2.3 will first establish the interpretation of construal adopted in this thesis, before then moving onto Hanks' (2013) concept of 'meaning potential' and how it provides an intuitive heuristic for how meaning comes to be expressed. Afterwards, Section 2.4 will present the framework of lexical aspect, noting its use in the nominal domain and its relevance to the current thesis. The chapter will end with a brief summary of the literature covered and will re-state the sub-aims of the thesis



moving forward. For now, however, we will begin with Section 2.1, which centres on categorisation and its importance in our understanding of lexical class.

## **2.1. Categorisation and Lexical Class**

Categorisation is one of the most important characteristics of human cognition. As Harnad (2017) claims, “to cognise is to categorise”. Fundamentally, categorisation constitutes a cognitive process which facilitates one’s comprehension of the world, through dividing it up, and imposing a systematic order on the world for the mind to more efficiently manage (Taylor 2003, p.xi). This process is often automatic and unconscious (Lakoff 1987, p.6). It is only when we encounter an entity that is ambiguous in regards to its classification that we become aware of the process of categorisation (Lakoff 1987, p.6). For instance, we would not question an apple’s status as a fruit, but continuously discuss whether a tomato should instead be categorised as a vegetable. Ambiguity in relation to the category status of certain entities has sparked debate about the structure of categories within current literature on categorisation. At present, two opposing approaches which attempt to account for category structure remain the most prominent (1) ‘the Classical model’; and (2) ‘Prototype theory’.

### **2.1.1. Category Structure: The Classical Model**

The Classical model can be traced back to Greek antiquity, first devised by Aristotle. Since its conception, it has permeated much of Western thought, and has traditionally been taken for granted in most scholarly disciplines (Lakoff 1987, p.6). The model proposes that categories are clearly distinguished on the basis of their similarities (van Der Auwera and Gast 2011, p.170). In doing so, categories are interpreted as possessing rigid boundaries, where category membership is dependent on a set of “necessary and sufficient conditions” (Van Lier and Rijkhoff 2013, p.6). As a result of this interpretation, each category member is said to hold equal status in relation to one another (Taylor 2003, p.21). Classical interpretations of categorisation subscribe to what Ross (1974, p.111) calls “the law of the excluded middle”, which is the notion that an entity can only either exist inside or outside of a category. From a Classical standpoint then, categorisation is a binary process. Admittance to a category is decided on an ‘all or nothing’ basis. For instance, it is common knowledge that mammals have warm blood. Here, the Classical model argues that if an animal does not have warm blood, then it cannot be classified as a mammal as it does not meet one of the necessary conditions of ‘mammality’.

Within the field of linguistics, categorisation is viewed as principally important. The very foundations on which linguistic theory and analysis are based is dependent on the existence of categories, be that morphological, syntactic, lexical, semantic, etc. Without categorisation, language in its spoken form would be no more than a stream of sounds. Through syntactic categorisation, we can map these sounds onto meanings by separating the sounds into distinct linguistic units, e.g., lexical classes (noun, verb, etc.). From a Classical view of categorisation, lexical classes possess “maximum homogeneity” (Gleason 1965, p.130), meaning lexical items which do not express all the properties understood as necessary for a specific lexical class are denied entry. For instance, a Classical interpretation would claim that a lexical item can only attain the status of a noun if it can be affected by quantity (the lexical item can be either plural or singular). If a lexical item cannot express this feature, then it will not satisfy the criteria required for ‘noun membership’. However, as shown later in this section, definitional criteria cannot always be used to determine status of a lexical item as a noun.

At face value, the Classical model forms an initially attractive and elegant proposition of categorisation. Aarts (2006, p.364) suggests part of the reason why the Classical model has remained popular for such a long period of time is its tie to the supposed human tendency to view the world as organised in terms of distinct categories. Moreover, the use of strict criteria as a gatekeeper to category membership affords language users the ability to refer to a set of lexical items which share homogeneous properties, and provide a structure to an otherwise disordered landscape (Aarts 2006, p.364). This use of strict criteria also enables individuals to discern the properties of a lexical item simply by identifying the lexical class it belongs to (Taylor 2003, p.35). The Classical view of categorisation, consequently, would appear to boast a high degree of cognitive efficiency (although Geeraerts (1985, p.141) disputes this position). Nevertheless, as will be covered below, several critiques of the Classical model have rendered its supposed advantages largely insubstantial.

### **2.1.2. Similar, But Not Exactly the Same**

The Classical model of categorisation was not born out of empirical study. As Lakoff (1987, p.6) points out, it was a philosophical standpoint reached through a priori speculation. Perhaps one of the earliest insights into the inadequacy of the Classical model is found in the work of Wittgenstein (1953), where he questions the importance of necessary features in the

categorisation process, by using the example ‘game’. Wittgenstein discussed how different types of game, e.g., board games, ball games, (and now) video games etc., exhibit considerably different properties, yet are still all recognised as games. What this example nicely demonstrates is that the use of necessary features as a determinant for category status often fails to cover the referential scope of what can be classified as an instance of a given category (Cruse 2011, p.54). In the following paragraph, we will see that this issue of inclusion commonly arises when attempting to classify lexical items into lexical classes.

Traditional lexical classes, such as noun and verb, have been shown to be unsatisfactory, as they are dependent on a mixture of morphological, syntactic and semantic criteria which do not necessarily correspond in particular instances (Lyons 1977, p.423). This position is echoed by Langacker (1987a, p.369), who states that lexical classes are not always open to homogenous characterisation of class membership. To use an example mentioned in Chapter 1, both the lexical items *house* and *salt* share morphological and semantic features. Nevertheless, they differ syntactically, as *house* is a count noun, in that it can be counted (e.g., *two houses*), whereas *salt* is a mass noun, as it cannot be counted (e.g., *\*two salts*). While a construal of two individual grains of salt is not beyond the realms of possibility, *salt* will generally be construed as a mass (Langacker 2004, p.82). Despite what the rigid boundaries of the Classical model indicate, both *house* and *salt* in the above instances can be categorised as nouns. From only looking at the examples provided above, it is apparent from the syntactic variation present that entities cannot always be categorised on necessary and sufficient conditions. This flaw has seen scholars seek alternative interpretations of categorisation, such as Prototype theory, which I will now cover below.

### **2.1.3. Category Structure: Prototype Theory**

Before we engage in a discussion of Prototype theory, two pieces of information are important to note. The first is that there is no one prototype theory. One cannot speak of ‘the Prototype theory’, but only a succession of ideas that have been introduced and (some) discarded by various scholars over time that each fall under the umbrella term of Prototype theory. The second is that prototypicality is not a structural component of a category, but an effect of the categorisation process. For example, given that most cats have fur, the possession of fur becomes a prototypical feature of the category *cat*. By reversing the causality of prototypicality, it allows us to consider ‘frequency of occurrence’ as an influential indicator of

prototypicality (van Der Auwera and Gast 2011, p.174). In Prototype theory, there are four prototypicality ‘effects’ that potentially arise in each prototype category. These prototypicality effects are outlined below.

The first prototypicality effect concerns the concept of ‘graded centrality’, which was pioneered by Rosch and her colleagues (Rosch 1973; Rosch and Mervis 1975; Rosch 1978). Looking back again at Wittgenstein’s (1953) ‘game’ example, we observed that entities within the same category do not always share the same properties. However, what Wittgenstein did not consider was that not all games are equally representative of the category ‘game’. Prototype theory presents a more flexible approach to categorisation, viewing category structure as scalar, and expressing what is defined by Aarts (2004b) as ‘Subjective Gradiance’ (SG). SG is the idea that category members exist on a continuum ranging from ‘core’ members which express all typical features associated with the given category, to more ‘peripheral’ members which do not express all of the typical features associated with the given category, but still share some similarity with the core members (Aarts 2004b, p.6). Accordingly, Prototype theory does not view all category members as equal, and poses a graded internal category structure, from less typical class members to the most typical, i.e., ‘prototypical’ (Rijkhoff and Van Lier 2013, p.3). This leads neatly into the second prototypicality effect: that Prototype categories do not contemplate the features associated with a particular category as essential for category membership (Geeraerts 2006, p.142). There is no one feature in a prototype category that a member must express to be a category member.

The third prototypicality effect relates to the ‘fuzzy’ nature of boundaries. By not imposing strict boundaries onto the structure of categories, prototype categories can exhibit a certain fuzziness at their boundaries. This is otherwise known as ‘Intersective gradiance’ (discussed later in this section). Of course, not all categories that exhibit prototype characteristics will exhibit fuzzy boundaries. For instance, while we may accept that a *robin* is a more prototypical *bird* than a *penguin*, we are aware that a *penguin* is still biologically a *bird*, and a *horse* is not. Lastly, prototype categories exhibit ‘family resemblance structures’: their semantic structure is composed of a “radial set of clustered and overlapping readings” (Geeraerts 2010, p.187). In such a structure, members will have at least one, or most likely multiple properties in common with one or more members, but none, or very few properties are common to all members (Rosch and Mervis 1975, pp.574–575). This structure is more simply interpreted as a chain of similarity relations between entities.

Geeraerts (2006, p.144) claims that the category structure proposed by Prototype theory is an appealing characteristic, as it provides a descriptive approach which captures our pre-theoretical intuitions that some category members are better examples of a category than others. An important study related to Prototype theory is Berlin and Kay's (1969) cross-linguistic study on colour terms, which found that although languages and, to an extent, individuals tended to separate the colour spectrum in distinct ways (e.g., the use of different terms for specific colours, i.e., different names for types of green/yellow), very similar shades of colour were generally chosen as the most representative shades of each given colour. For example, the languages of American-English and Vietnamese both indicated that the same shade of *red* was the most representative of the category *red*, despite separating the scope of colours that could be classified as *red* differently (Berlin and Kay 1969, pp.119; 133). These findings support the concept of prototypicality within categories, as they indicate that certain members of a category are 'better', more prototypical instances than others.

By rejecting the notion of essential properties, Prototype theory proposes that the boundaries of categories are not unambiguous and clear cut, but are instead 'fuzzy' (van Der Auwera and Gast 2011, p.172). This fuzziness allows for 'Intersective Gradiance' (IG) (Aarts 2004b). While SG (established above) relates to intra-categorical structure, IG concerns the potential convergence of distinct categories (Aarts 2004b, p.3). By propagating the notion of IG, Prototype theory acknowledges the potential for multiple diverse categories to grade into one another along various parameters, enabling the possibility of a continuous spectrum between categories. An example that demonstrates the potential for fuzziness in boundaries is the word *cup*. In Labov's (1973) study into the linguistic categorisation of objects, he prepared drawings of containers of various shapes. These drawings were shown to subjects who were then asked to name each of the objects portrayed in front of them. From the subjects' answers, Labov found that the boundaries of a *cup* were ambiguous, as the answers from subjects did not establish a clear division between a *cup* and similar container-like objects, such as a *bowl*. The categories *cup* and *bowl* were seen to grade into one another, in contradiction to the Classical model's proposal of clear-cut category boundaries.

The four types of prototypicality effects which together form the basis of Prototype theory are presented in the Figure 2-1 below. These four prototypicality effects are systemically related along two dimensions: 'Extensional' and 'Intensional' characterisation. While the extensional

dimension concerns the scope of a lexical item’s use or an individual sense of a lexical item, the intensional dimension explores the different semantic properties that a lexical item can express and its definition (Geeraerts 2010, p.188). Geeraerts (2010, p.188) claims that the prototypicality effects (a) and (c) in Figure 2-1 relate to the extensional structure of a category, and primarily focus on the example category members that represent the referential boundaries of the category, i.e., prototypical and peripheral category members. Prototypical members at this referential level are known as ‘onomasiologically’ salient, i.e., they are the most salient denotatum for a given entity (Geeraerts 2010, p.200). The prototypicality effects (b) and (d), alternatively, focus on the intensional level, where the definitional structure of the category is the primary focus instead of the referential structure. Prototypical members at this definitional level are known as ‘semasiologically’ salient, i.e., they are the most salient semantic possibility for a given entity (Geeraerts 2000, p.79).

	Extensional characterisation (on the level of exemplars)	Intensional characterisation (on the level of definition)
Non-equality (salience effects, core/periphery)	(a) Differences of typicality and membership salience	(b) Clustering into family resemblances
Non-discreteness (demarcation problems, flexibility)	(c) Fuzziness at the edges, membership uncertainty	(d) Absence of necessary-and-sufficient definitions

*Figure 2-1. Four types of prototypicality effects (Geeraerts 2010, p.189).*

Relating back to the category structure of lexical classes, van Der Auwera and Gast (2011, p.181) assert that if lexical classes cannot be identified on the basis of necessary and sufficient conditions, their status as principles of grammar remains largely questionable. They do state nevertheless that the idea of lexical classes displaying effects of prototypicality provides an attempt to rescue the very idea of lexical class (van Der Auwera and Gast 2011, p.181). Fontaine (2013, p.26) claims that lexical classes do not have strict boundaries, as a lexical item of a given lexical class may not behave identically to another lexical item of the same lexical class.

Following on from the works of Quirk (1965) and Crystal (1967) on lexical class structure, and also Lakoff’s (1970) work on gradience phenomena, Ross (1972; 1973) was perhaps the first to apply the concept of gradience to the structure of lexical class. Focusing on the lexical class

of noun, Ross (1972; 1973) devised a scale of nominality. This scale of nominality ranged from supposed prototypical nouns which express the most nominality, such as underived nouns, as shown in example<sup>2</sup> (4a) below, to more supposed peripheral nouns which express the least nominality but still display nominal properties, such as ‘*that*-clauses’, shown in example (4b). Although we will uncover later in this section why Prototype theory accounts of lexical classes such as Ross’ may not be entirely accurate, for now it is worth acknowledging the ability of Prototype theory to capture our intuitions about the structure of lexical classes and the entities within these structures. Like the example of *cup* examined above, lexical classes also exhibit fuzzy boundaries, and can grade into one another (Aarts 2004a, p.375). For example, nominal gerunds express both verbal and nominal properties, as they display verbal morphology (*-ing*), but syntactically function as nominals (Fonteyn 2016, p.60), shown in example (5). Nominal gerunds are accounted for by Prototype theory, as they obtain a middle space between nominality and ‘verbiness’; a space which is conclusively denied by the Classical model (Aarts 2004a, p.345).

- (4) a. Jack’s table.  
b. That Jack kicked the ball.
- (5) Jack’s running.

Prototype theory therefore provides compelling insight into the behaviour of categories as Prototype theory offers a more inclusive take on category structure than the Classical model. Nonetheless, Prototype theory is not perfect, which I will now discuss below.

#### **2.1.4. The Consequence of Oversimplified Feature Lists**

One overarching critique of Prototype theory is that notions of prototypicality are unnecessarily tied to ‘oversimplified’ lists of features (Croft and Cruse 2004, p.87). These lists of features have been labelled as oversimplified, as they are seemingly unable to account for several additional aspects which influence the status of an entity as a more prototypical or peripheral category member. The first of these aspects is context. Context clearly plays a role in determining the categorical status of entities as either more prototypical or peripheral category

---

<sup>2</sup> Throughout this thesis the instance in question will be underlined within the examples given. Moreover, all examples used throughout chapters 2 and 3 will be self-produced examples, unless stated otherwise.

members. For instance, what may be considered a prototypical *fruit* will most probably vary, dependent on contextual variables, such as country, culture etc. However, in its current state, Croft and Cruse (2004, p.87) argue that Prototype theory is unable to completely reflect the entire influence context may have on categorical status. This inability is also apparent in the categorisation of nominality. For example, as presented in example (6) below, one might initially view certain deverbal nominalisations (nominals which have derived from verbal stems) as more peripheral nominals, due to their ability to express event meaning (this behaviour is discussed later in this section) within their nominal form (Vendler 1967, p.141). However, in the register of ‘Scientific Writing’, this ability is highly regarded, as it allows writers to condense clauses into nominal ‘technical terms’, which can be referred to throughout the rest of their discussion (Thompson 2014a, p.245). Due to this function, deverbal nominalisations might be considered more of a ‘typical’ nominal within the genre of Scientific Writing.

(6) The interaction took place.

In addition to the problems posed by context, traditional notions of Prototype theory do not adequately assess how potential interactions between category features could affect the prototypicality of an entity (Cruse 2011, p.64). Medin and Shoben (1988) in their study into modifications of conceptual knowledge associated with judgments of adjective-noun conceptual combinations show that the combination of certain features can influence the perceived prototypicality of an entity. They found that participants considered small spoons to be more prototypical of the category ‘spoon’ than large spoons, but large spoons to be more prototypical if the spoon was wooden. The prototypicality of a spoon was seemingly dependent on the existence of an interaction between the size of the spoon and its material. Accordingly, the structure of categories such as noun and verb may also rely on the interactions of multiple nominal and verbal features. However, it appears that few attempts have been explicitly undertaken to identify the importance of feature interaction within the categories noun and verb. The influence of feature interactions in establishing different nominal semantic behaviour will be specifically assessed throughout Chapters 3-5.

Sticking with the value of importance, improvements to Prototype theory have been made through the observation of ‘weight’ differentials in feature lists (Rosch 1978:30). While traditional oversimplified feature lists interpret each feature as possessing equal strength in



determining the category status of an entity, the addition of feature weight differentials to Prototype theory allows us to observe that certain features of a category are evidently weighted as more important, i.e., more influential, to category status than others (Taylor 2019, p.42). Watanabe's (1985) 'Ugly Duckling Theorem', based on the well-known fable 'The Ugly Duckling', provides the perfect example of the difference in weight between category features. Watanabe (1985) states that there is no ground to claim that the Ugly Duckling is any more different than the other ducklings, but we perceive the Ugly Duckling to be less similar than the other ducklings as we weight certain features, such as colour, as more central to what might constitute the 'prototypical duckling'.

In relation to lexical categories, literature remains sparse on what features might be weighted more 'nouny' or more 'verby' than others. For instance, is 'having the potential to classify a referent in a referring expression' more of a central nominal feature than 'the ability to express inflectional/grammatical information such as pluralisation'? If so, what are the reasons on which this weighting difference is based? We might initially claim that 'the ability to be affected by definiteness' is more central to nominal status, as most nouns can be affected by definiteness, and not as many can be pluralised. While this statement is true, a common misconception is to equate prototypicality with frequency of occurrence.

As discussed in Fontaine et al. (in-press, p.80), prototypicality may relate more to what language uses *believe* is most typical, as opposed to what empirical corpus evidence would suggest. For example, Taylor (2003, p.56) states that children's drawings typically depict the sky as the colour blue. Yet, in some environments the sky is more frequently grey than blue. If you live in Cardiff like I do, you might understandably feel like the sky is almost never blue. On this account, it would appear that what constitutes a prototypical instance of an entity may not always correlate with the most frequent instance of that entity (Taylor 2003, p.56). The colour blue appears a more readily drawn upon, i.e., a more 'salient' (Geeraerts 2000, p.75), description of the sky, despite occurring less frequently in certain environments. While typicality is evidently associated with frequency of occurrence, it is also associated with a sense of belief, informed by one's individual experience of the world (Fontaine et al. in-press, p.80). Therefore, work remains to provide clarity on what determines the weight of certain features within the categories of noun and verb, and which features are weighted as more central to nominal and verbal status than others.

In addition to providing transparency to the weight distribution of features within categories, future work must also look beyond features alone as the sole contributors to category status. Taylor (2003, p.38) claims that we do not only apprehend categories by reducing them to their individual components, but that we also comprehend and learn most categories “holistically, in the context of our interaction with the world”. Lakoff (1987, p.xv) argues for the notion of ‘Experiential Realism’, which is the idea that our thoughts are closely associated with our bodily experiences. For instance, when we think of a bed, we do not necessarily only think of the features of a bed (e.g., *mattress*; *quilt*), but also the perceptual experiences we might gain from a bed, such as *rest* and *comfort* (van Der Auwera and Gast 2011, p.171). Thus, when we categorise entities, we also draw upon our ‘real-world knowledge’ of how we interact with the entities, and how those entities interact with the rest of the world (Taylor 2003, p.38). It is known that the meanings of lexical items produce similar cognitive activation patterns to their real-world referents (Damasio and Tranel 1993; Martin et al. 1995). Moreover, nominals which express real-world referents, i.e., ‘concrete nouns’, are typically learned earlier in childhood than nominals which do not express real-world referents, i.e., ‘abstract nouns’ (Saxton 2010, p.149). Real-world experiences with entities, in turn, may also influence the status of nouns as either more prototypical or more peripheral nominal instances.

Turning back now to category features for the final critique in this section on Prototype theory, perhaps the most important issue relates to the supposed under-specification of category boundaries. Geeraerts (2006, p.149) argues that by not clearly identifying the boundaries which contribute to the status of an entity as a particular category member, the category status of peripheral instances is often left ambiguous. To better understand this point about ambiguity, I will discuss below the ambiguity present in the previously mentioned Prototype-based theory of nouns proposed by Ross (1972; 1973).

As discussed earlier in Section 2.1.3, Ross (1972; 1973) devised a prototypical scale of nominality, ranging from supposed prototypical nouns which express the most nominality, such as morphologically underived nouns, to supposed peripheral nouns which express the least nominality but still display nominal properties, such as ‘*that*-clauses’. Within such a scale, we can ask where nominals such as *storm* should be placed. According to Ross (1972; 1973), *storm* would constitute a highly prototypical noun due to its morphologically underived status. However, *storm* semantically expresses an event: a *storm* can take place. This expression of an event by *storm* indicates a quality traditionally thought of as verbal, through the nominal’s

capacity to express dynamic temporal meaning, i.e., an event occurring over time. What makes this example particularly interesting is the fact that *storm* can express an event, despite not deriving from a different lexical class (Vendler 1967, p.141). What we can see in the above example is that in attempts by Prototype theory to categorise lexical classes, there exists a certain amount of ambiguity in relation to lexical items whose grammatical form does not correspond to their semantic meaning. This issue, specifically, will be revisited throughout this thesis, as it will become apparent in the next section that underived nouns which exhibit event semantics remain largely unaccounted for within lexical class distinctions between nouns and verbs.

In this section, I examined how categories are currently conceptualised and structured, whilst also exploring their existing imperfections. Most notably, I located several faults with Prototype theory. These included its inability to fully account for the influence of context, category feature interactions, and the influence (weight) of category features on category structure. Furthermore, Prototype theory's inability to account for the influence of real-world experience on category structure, as well as the ambiguity produced by its under-specification of category boundaries were discussed. Nevertheless, Prototype theory is still a useful concept, especially for its ability to capture our pre-theoretical intuitions that some category members are better examples of a category than others (Geeraerts 2006, p.144). Accordingly, despite the above-mentioned critiques of Prototype theory, this thesis will continue to draw upon aspects of Prototype theory but will not use the theory exclusively. In the next section, I will now turn to distributional approaches that have attempted to categorise nouns and verbs. Typically, these distributional approaches have oriented to conceptual-semantic definitions (Langacker 1987b; 1991); the semantic and pragmatic components of nouns and verbs (Croft 1991; 2001), and also their discourse functions (Hopper and Thompson 1984b). By reviewing these distributional approaches, we will see that while prototype attempts at categorising nouns and verbs are seemingly efficient in capturing the behaviour of prototypical category members, there is a notable lack of clarity in these approaches when covering nominal and verbal instances that do not subscribe to prototypical conventions. I will first start with Langacker's (1987b; 1991) conceptualisation of nouns and verbs.

## **2.2. Distributional Approaches to Nouns and Verbs**

To explore Langacker's proposed distinction between nouns and verbs, we will first refresh our knowledge on the linguistic framework of 'Cognitive Linguistics', and also Langacker's (1987a; 1991) theory of 'Cognitive Grammar'.

### **2.2.1. Cognitive Grammar and Useful Terminology**

Cognitive Linguistics (CL) is an approach to the study of language which is rooted in the assumption that our linguistic capabilities are a product of our cognitive abilities and that the meaning we derive from language is a product of our conceptualisation (Dąbrowska and Divjak 2015, p.1). Like the early work from scholars such as Malinowski (1923), Vološinov (1986) [1929], and, slightly later, Halliday (1985), CL also holds the assumption that grammar is moulded by its use. Taking these viewpoints, the framework of CL supports the idea that our knowledge of language resides within the mind, and considers the primary job of linguists to describe the mechanisms within the mind which allow individuals to generate and interpret linguistic expressions (Taylor 2002, p.5). These interpretations are typically achieved through the analysis of linguistic structure in relation to our basic cognitive systems and abilities (understood as interlinked with linguistic structure), such as perception, attention and categorisation (Langacker 2017b, p.32).

'Cognitive Grammar' (CG) (Langacker 1987a; 1991) is a theory of language which belongs to the wider framework of CL. CG places emphasis on the symbolic nature of language and the position conceptualisation plays in social interaction (Langacker 2017b, p.31). Grounded in the view of conceptualist semantics, CG acknowledges the significance of construal (further discussed in Section 2.3.1), i.e., our capacity to understand and represent identical situations in different ways, and views grammar as a "symbolic phenomenon", comprised of patterns for establishing and representing specific schemes of conceptual organisation (Langacker 2017b, pp.31–32). Here, the term 'scheme' refers to an abstract template which constitutes the similarity of the structures it categorises (Langacker 1990, p.59). For instance, the concept of 'musical instrument' shares a relationship of schematicity to notions such as 'guitar' and 'piano'. CG proposes that morphology, syntax and lexis constitute a continuum, which is viewed as groupings of symbolic structures (form-meaning pairings) (Langacker 1987a, p.3). Accordingly, CG views all grammatical expressions as expressing schematic conceptual meaning (Langacker 1991). It is through these conceptual meanings that Langacker (1987b,

p.53) claims that a distinction between nouns and verbs is plausible. In the following paragraphs, I will detail this apparent conceptual distinction between nouns and verbs.

### **2.2.2. The Billiard-Ball Model**

Heyvaert (2003, p.26) claims that as well as the internal functional properties of lexical items, such as morphological, syntactic and semantic criteria, the categorisation of lexical items is also dependent upon ‘intrinsic’ conceptual distinctions. Langacker’s (1991) conceptual definition of nouns and verbs is comprised of two proposals: one which primarily focuses on the prototypical instances of nouns and verbs, and one which attempts to establish an inclusive characterisation of all nouns and verbs, taking a highly abstract approach. We will start with the former, known as the ‘Billiard-ball’ model.

Langacker (1991, pp.13–14) proposes a conceptual distinction between prototypical nouns and verbs based on an “idealised cognitive model”, grounded in everyday experience. To convey this conceptual distinction, the idealised cognitive model draws upon the analogy of a game of billiards. Similar to a pool table, our world is populated by discrete physical objects, which are capable of moving through space and connecting with each other (Langacker 1991, p.13). In this model, the lexical class ‘noun’ typically represents the discrete physical objects on the table, while the lexical class ‘verb’ typically represents the energetic interactions between them.

This distinction is established on four elemental components: ‘Space’, ‘Time’, ‘Material substance’ and ‘Energy’ (Langacker 1991, p.14). Material substance is manifested in space. For example, we can ask ‘where is the table?’. Energy, alternatively, can only be observed through change, so is dependent on time for its manifestation. Prototypical nouns are viewed as instantiated in space as they denote physical objects which are comprised of material substance, and prototypical verbs are instantiated in time as they denote energetic interactions which require change (Langacker 1991, p.14). Langacker (1991, p.14) claims that the physical objects that nouns denote are ‘spatially compact’, in that they are composed of a restricted amount of substance. They are also ‘temporally stable’, as they will generally alter slowly, if at all, through time (Givón 2001, p.51). In opposition to physical objects, the energetic interactions that verbs denote are understood to be ‘spatially expansive’ in that they include the spatial locations of the various participants associated with the interaction over a duration

of time (i.e., the balls across the pool table) (Langacker 1991, p.14). These interactions are also conceived as ‘temporally unstable’, as interactions are transient, so occur at a specific instant in time. Lastly, physical objects and energetic interactions differ in relation to ‘autonomy’. While physical objects are conceptually autonomous, as the conception of a physical object does not presuppose an energetic interaction, the same cannot be said for energetic interactions (Langacker 1991, p.14). Energetic interactions are conceptually dependent, as they inherently require some reference to the entities through which they transpire (Langacker 1991, p.14).

Although this conceptual distinction between nouns and verbs provides perceptive insight into the relationship between lexical classes and their real-world denotations at a prototypical level, it fails to account for less prototypical instances of nouns and verbs. Hopper and Thompson (1984b, pp.705–706) state that Prototype models which employ structured notions of time-stability are difficult to apply to empirical data, as certain nominals can express events, as in example (7a) below, while many verbs denote more stable situations, shown in (7b).

- (7) a. The destruction  
b. Jack knew the answer

Wilcox (2018, pp.61–62) acknowledges the simplicity of this classification, noting that the conception of physical objects and energetic interactions represent the most typical nouns and verbs. For the less core members of each lexical class, Langacker (1987b; 1991) proposes an inclusive characterisation, taking a highly abstract approach. Instead of attempting to distinguish the two lexical classes in relation to objective, truth-conditional factors, Langacker (1991, p.15) claims that a more subjective theory of meaning based on cognitive processing may be more applicable. Specifically, our “capacity for imagery”, i.e., one’s capacity to construe the same conceived entity differently. An entity here refers to “anything one might refer to for analytical purposes” (Langacker 1991, p.16). This can extend from a physical object, relationship, sensation, location etc., to a period of time, or a sound etc. Once this capacity for imagery is realised, nouns and verbs are attributed conceptual import, based upon their construal (Langacker 1991, p.15). In the following section, we will only examine the characterisation of nouns in relation to cognitive processing, as the cognitive processing of verbs goes beyond the scope of this paper which later directs its focus principally on the semantic behaviour of nouns. For the interested reader however, Langacker (1987a; 1991) presents seminal work into the cognitive processing of verbs. In the section below, I will now

explain two central terms to help us understand how nouns are supposedly cognitively processed: ‘Domain’ and ‘Profile’.

### 2.2.3. An Abstract Characterisation

The term ‘domain’ constitutes the cognitive structures that a lexical item presupposes, otherwise known as its conceptual base (Langacker 1990, p.61). The domain is, more simply put, a composition of background knowledge against which the conceptualisation of a lexical item is realised (Taylor 2002, p.195). The term ‘profile’, alternatively, represents a region of that domain, which is the specific entity that the lexical item denotes (Langacker 1990, p.61). For instance, the domain of a *circle* is two-dimensional space, and its profile is viewed as representing a specific region, i.e., a *circle*, within this basic domain. This relationship between a lexical item’s profile and domain is termed its ‘Semantic Pole’ (Langacker 1987a, p.183). A visual representation of the semantic pole of *circle* is shown in Figure 2-2.

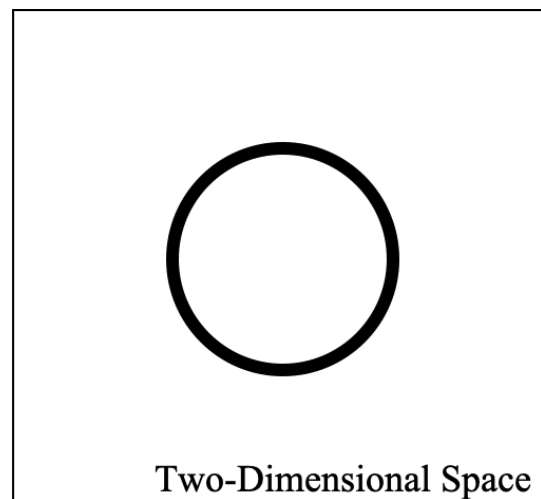


Figure 2-2. The profile of circle (indicated by the circle shape) within its domain of two-dimensional space (indicated by the surrounding box) (Langacker 1987a, p.184).

Some domains are labelled ‘basic’, as they cannot be simplified or broken down any further, e.g., our perception of space and time (Langacker 1987a, p.148). Most domains, nevertheless, are not basic, and are known as ‘abstract domains’ (Langacker 1987a, p.150). These types of domains express cognitive structures of indeterminate complexity, ranging from novel conceptualisations to entire knowledge systems (Langacker 1987b, pp.55–56). A profile is not constrained to one domain and can sometimes require a combination of multiple domains. For instance, the nominal *beep* requires the domains *time* and *pitch* (Langacker 1987a, p.191).

Furthermore, certain lexical items can invoke the same domain, but express different profiles (Langacker 1987b, p.61). This difference in profile can lead the lexical items to differ semantically. For instance, the lexical items *arm* and *leg* are profiles of the same domain *body*, but they contrast semantically as they occupy different regions, i.e., profiles, within that domain. Taking our knowledge of domains and profiles forward, we will now focus on how nominals are conceptualised in CG.

### **2.2.3.1. Nominal Conceptualisation in Cognitive Grammar**

In CG, entities are distinguished into those which profile things and those which profile relationships (Broccias 2013, p.195). In this current section, we will focus on entities which profile things. According to Langacker (1987a, p.183), a nominal is distinct from other lexical categories as it profiles a ‘thing’. The ability to profile a thing is a conceptual property which is exclusive to nominals. The expression of this conceptual property, in turn, is said to simultaneously define the lexical class noun (Langacker 2017b, p.46). But this prompts an obvious question: what is a ‘thing’? Langacker (2017a, pp.268–269; 2017b, p.45) defines a thing abstractly in terms of ‘grouping’, stating that nouns profile groups, where numerous entities function together to represent a singular whole, i.e., a ‘thing’. This whole then represents a region in some domain (Langacker 1987b, p.58). Langacker (2017b, p.45) claims that physical objects represent the most typical nominals, as the conceptual grouping of their constitutive entities is so rudimentary and automatic that, to become conscious of them, we would require a higher level of conceptual analysis. He provides the example of a *rock* to support this claim, asserting that a *rock* is made up of “splotches of material substance that are continuously distributed throughout a certain expanse of space” (2017b, p.45). This is an instance of what is termed ‘conceptual reification’, as the splotches are together conceptualised as a uniform thing; they cannot be individually distinguished nor knowingly acknowledged (Langacker 2000, p.10).

As well as physical objects, abstract nominals also constitute things which exhibit grouping and reification. For example, the abstract nominal *theory* exhibits grouping and reification, as all of the elements of a *theory* unite to perform an overall function which constitutes the *theory* (Langacker 2017a, p.269). Moreover, even abstract nominals, such as those that denote emotions, supposedly constitute things. For instance, the emotion of *anger* is instantiated through the experience of *anger*. It is the individual ‘fragments’ of this experience that together group and reify the experience as a particular type of emotion that represents a region in the



emotive domain (Langacker 1991, p.30). Due to the focus of this chapter, I cannot detail how every nominal is conceptualised. However, the examples above were chosen with the intention to capture a range of how different nominals are conceptualised in Langacker's (1991) conceptual-semantic classification.

Overall, Langacker's more subjective theory of meaning based on cognitive processing presents an intuitive understanding of how individuals conceptualise an entity as nominal. Despite this, the theory lacks empirical psychological evidence to support its claims about conceptualisation. In particular, Broccias and Hollman (2007, p.498) question the understanding of nouns and verbs in CG, stating that, to achieve a strong essence of validity, the theory must be supported by evidence and also be capable of being verified through experimentation (see Langacker (2008b) for his response to this critique). Moreover, while Langacker's Billiard-ball model adopts a prototype type stance, his theory of meaning based on cognitive processing is interestingly more aligned with the Classical model of categorisation. More specifically, the theory proposes an abstract schematic definition that attempts to capture all members of the noun category (Fonteyn 2019, p.78). As a result, and as with generally all attempts at classical categorisation, the theory provides less attention to graded category structure and the intra-categorical nuances of different nominals (such as the expression and interaction of different category features) which help us better interpret the categorical status of nominals. Below I will review two further approaches to the classification of nouns and verbs, before then moving onto Section 2.3, where I will explore how the meaning of lexical items can provide a beneficial foundation from which to capture of divergent behaviour of nominality.

#### **2.2.4. Nouns and Verbs: Distinctions of Discourse Functionality**

Hopper and Thompson (1984b) propose a lexical class distinction on the basis of discourse functionality. Hopper and Thompson (1984b, p.708) claim that the differentiating discourse functions displayed by nouns and verbs are central to their classification, as the semantic distinctions expressed by nouns and verbs are the product of their typical discursive role. They claim that prototypicality is imposed on linguistic forms by discourse, and nouns and verbs form the prototypical instantiations of basic discourse functions (Hopper 1991, p.30).

For nouns, this typical discourse function is ‘referentiality’. According to Hopper and Thompson (1984b, p.711), prototypical nouns function in noun phrases as principally referential entities. Taking from the work of Du Bois (1980) and Givón (1981), Hopper and Thompson (1984b, p.711) assert that nouns typically express a “continuous identity over time”, which is recognised in discourse. This continuity is not always the case however, as Jones (2014) notes that identity can evolve over discourse, using the example of a ‘potato’. A peeled potato is conceptually different from an unpeeled potato, despite the noun remaining the same in the referential expression. Hopper and Thompson also suggest that nouns are generally discrete concrete entities that possess the potential to act as participants or propositions (Hopper and Thompson 1984b, p.711). Prototypical verbs, contrarily, report discourse events, and typically transpire in environments where participants are introduced or events are reported (Hopper and Thompson 1984b, p.709).

In line with the early thinking of Halliday (1961, p.277), Hopper and Thompson (1984b, p.747) claim that lexical items lack ‘categorality’ (i.e., lexical class classification) until a lexical class is enforced upon them by their use in discourse. In doing so, they acknowledge that certain nouns and verbs can express varying categorality, as nouns and verbs perform their respective discourse functions to differing degrees. For nouns, ‘predicate nominals’ (shown in example (8) below) are viewed as expressing the least categorality, as they exhibit an “absence of intention to refer” to an entity (Hopper and Thompson 1984b, p.715). For verbs, participial forms, as in example (9), are regarded as displaying the least categorality, as they typically inflect like nouns, with *-ing* case and number (Hopper and Thompson 1984b, pp.740–741). Croft (1991, p.50) notes that although this line of thought accounts for how lexical items lose their nominal or verbal discourse functionality, the approach lacks an explanation as to why these lexical items express this divergent discourse functionality in certain contexts. Contrarily, this emphasis on the motivation for discourse functionality is somewhat present in their account of nominalisations. Nominalisations, in Hopper and Thompson’s (1984b) approach, express discursive properties typical of both nouns and verbs. They are viewed to denote an event, but one that is being referred to, instead of reported, in the discourse (Hopper and Thompson 1984b, pp.745–746).

While Croft (1991, p.50) notes that this is certainly a step in the right direction for Hopper and Thompson, it would be largely inaccurate to suggest that all nominalisations express event semantics. Fontaine (2017b, p.8) raises examples such as *appliance* and *teacher*, which are

deverbal nominalisations that express object semantics instead of event semantics. Furthermore, Sasse (2001, p.501) asserts that it is not transparent whether the functional prototypes underlying Hopper and Thompson's (1984b) distinction are the correct ones, or the only feasible ones. Languages, according to Sasse (2001, p.501), "possess all kinds of grammatical machinery to signalise indefinite, generic, predicative, and non-referential uses of nouns". So, to establish definite discourse-referent-indicating nouns as the most typical would be to deny the multiple different forms that nouns typically express.

(8) Jack is a builder.

(9) Having eaten his dinner, Jack was now full.

### 2.2.5. Semantic Class and Pragmatic Function

Unlike the two previous approaches to lexical categorisation, Croft's (1991) approach does not directly relate to the lexical categories of noun and verb, but, instead, to semantic classes and pragmatic functions. With regards to semantic class, Croft makes reference to the traditional semantic account of nouns as denoting objects, adjectives as denoting properties and verbs as denoting actions (Brown 1957; Dixon 1982). However, he states that semantic classes based on intuition require more precise semantic definition (Croft 1991, p.62). Specifically, Croft argues that the three major semantic categories, objects, properties, and actions, are distinguished by four semantic properties: 'Valency'; 'Stativity'; 'Persistence'; and 'Gradeability'. As the current focus is on the distinction between nouns and verbs, the following discussion will only discuss these semantic properties with regards to objects and actions.

Valency, somewhat akin to Langacker's (1991) semantic property 'autonomy', relates to whether a concept is "inherently relational" or not, i.e., whether its existence is dependent on the existence of another entity (Croft 1991, p.63). Actions are dependent on the existence of separate entities, so express valency. Objects, alternatively, are typically not dependent on the existence of separate entities for their own existence (this stance could be debated for relational nominals like *father* and *son*), so do not exhibit valency. The second semantic property, stativity, describes whether a concept displays a change of state (Croft 1991, p.63). Prototypical actions exhibit such changes of state, while prototypical objects display an absence of change, so remain stative. The third semantic property is persistence (similar to stativity). Persistence,

like the distinction of temporal stability used by Langacker (1991) and Givón (2001), relates to the temporal length of a concept (Croft 1991, p.64). Prototypical actions display rapid change, so are deemed as ‘transitory’. Alternatively, prototypical objects are assumed to be ‘persistent’, as most will display either slow change, or no change at all. Lastly, the semantic property of gradeability distinguishes properties from objects and actions (Croft 1991, p.65). Prototypical properties can be manifested by degrees, while prototypical objects and actions cannot.

In addition to semantic classes, Croft also establishes the typical pragmatic functions displayed by these classes. These pragmatic functions are defined as “propositional speech acts” that arrange the information expressed by the lexical roots for the purpose of communication, and conceptualise it in a particular way (Croft 1991, p.51). Croft proposes three principle pragmatic functions: Reference; Predication; and Modification. Each of the pragmatic functions is said to typically correlate to a specific semantic class. According to Croft (1991, p.51), reference generally maps onto the semantic class object, as it concerns the identification of entities. Predication, on the other hand, commonly shares an association with the semantic class action, as it relates to what the speaker says about the entity they are speaking about. Croft acknowledges that these relationships between pragmatic function and semantic class are not exclusive but are prototypical. He claims that, through the presence of additional morphemes, marked relationships between pragmatic function and semantic class are created (Croft 1991, p.58). Deverbal nominalisations are one of these marked instances, as they are said to typically express the semantic class of action, while simultaneously expressing the pragmatic function of reference, due to their capacity for identification. The static nature of this model however can lead to complications. For instance, as mentioned previously in Section 2.1.4, with regards to the genre of ‘Scientific Writing’, static conceptions of markedness inadequately compensate for the context in which grammatical configurations are formulated, as contextual factors impose a substantial influence on what configurations are preferred in certain instances (Bloor and Bloor 2013, p.130).

### **2.2.6. Categorising Nouns and Verbs Summary**

What we see from these differing approaches is that prototypical nouns and verbs can be distinguished from one another. The prototypical semantic, conceptual, discursive, and pragmatic features expressed by nouns and verbs are clearly partitioned into two opposing sets.

However, once one wants to investigate nominal and verbal instances that do not subscribe to these typical conventions, a lack of clarity is apparent in all the discussed approaches. While attention has been given to less prototypical nominals such as nominalisations, far less attention has been dedicated to underived nominals which also express event meaning, such as previously used examples like *storm*. What is evident from the above distributional approaches to nouns and verbs is that it is particularly difficult to account for these instances in traditional Prototype-based models, as they generally exhibit prototypical nominal syntax but temporal semantics, leaving their position on the nominal continuum ambiguous. Thus, we cannot predict the exact behaviour of individual nominals by only referring to traditional notions of noun and verb.

The expression of temporal semantics is a particularly interesting area within the nominal domain, as temporality is a concept that has, up until recently, been typically associated with verb semantics (discussed in detail in Section 2.4). By gaining a greater understanding of how nominals can express temporal meaning, we will gain a clearer picture of the semantic boundaries of nominality, and, subsequently, the nature of the lexical class. Accordingly, in the following sections of this chapter, I will direct our focus to the semantic behaviour of nominals and how meaning is expressed. In doing so, I will first present the interpretation of construal that I will adopt in this thesis, touching on the linguistic frameworks of CL and Systemic Functional Linguistics (SFL). Establishing how this thesis views construal is particularly important, as there are various discussion points throughout the following chapters that cover the influence of syntax and real-world (i.e., ontological) knowledge on the semantic expression of nouns. As a result, and as I will discuss in the following section, both a cognitive and a more functional view of construal is required to accurately capture how nominals are able to express different semantic meanings.

Afterwards, we will explore the notion of meaning in more depth and how it is expressed by visiting Hanks' (2013) concept of 'meaning potential'. The decision to focus on meaning potential specifically was chosen, as the concept seemingly provides an intuitive proposal for how word meaning comes to be expressed in different syntactic contexts. In doing so, it can enable an insight into nominal semantics that cannot be gleaned from the base of prototypicality alone. Following on from this section, the chapter will direct its attention to the framework of 'lexical aspect', centring on the capacity of lexical aspect to enable access to the semantic behaviour of nominals in use, and its recent application within empirical studies. It is

this capacity, which is of interest, as lexical aspect can provide a supposedly more objective base from which to empirically study the nature and degree of nominality in English and examine how it can be evaluated.

## **2.3. Construal and Meaning Potential**

### **2.3.1. Construal**

Within the field of linguistics, the term construal has received several different interpretations from various frameworks. Construal has so far been described as our capacity to understand and represent identical situations in different ways (Langacker 1991, p.15) (Section 2.2.1). which largely represents the CG interpretation, since the emphasis is on the relationship between language, thought and experience (Croft and Cruse 2004, p.32). CG views construal as a subjective process, an inescapable product of human conception (Langacker 2017b, p.34). Jackendoff (1983, p.29), in particular, takes the viewpoint that we only have access to a ‘projected’ version of the world, as our perception of the world is unconsciously organised by our minds (a similar standpoint is also briefly risen in Section 2.2.1). As a result, our capacity for construal is limited. We can only talk about things which have successfully gained mental representation through this unconscious organisation (Jackendoff 1983, p.29).

From the CG perspective of construal, autonomy is given to the speaker, noting their ability to actively organise and construct their own world through language (Taylor 1995, p.4). Thus construal within CG largely concerns the association between the speaker (or hearer) and the situation they actively conceptualise and represent (Langacker 1987a, pp.487–488). This relationship is formed through cognitive activity, which is based on the individual’s prior experience, cognitive capabilities, and their state at the time of conceptualisation. In CG, construal is said to encapsulate numerous interrelated factors, which Langacker (2019) broadly organises into five ‘construal dimensions’: ‘Perspective’; ‘Selection’; ‘Prominence’; ‘Dynamicity’; and ‘Imagination’. These dimensions are not discrete, and all become active in a given linguistic expression.

Within the current thesis, I do not have space to explore each of these dimensions in detail, nor will I be taking them forward and explicitly applying them later. For a comprehensive overview of these construal factors, see Langacker (2019). The reason for not taking these dimensions forward is that this thesis will not exclusively adopt the CG understanding of construal. This

choice was made because, as well as a cognitive ability, language is evidently also a key social feature in the interaction between humans (Croft 2009, p.397). One framework that principally focuses specifically on the use of language in its social environment is Systemic Functional Linguistics (SFL) (Hasan 2009, p.37). As I will discuss in the following paragraphs, SFL has been previously seen as a complementary framework to CL, predominantly because of the focus it places on how meanings are constructed in naturally occurring text (Butler 2013, p.204) as opposed to only the cognitive processes involved in the construction of meaning.

SFL is a heuristic theory, which treats language as a semiotic system that speakers can use to create meaning within particular social and cultural contexts (Bloor and Bloor 2018, p.151). This semiotic system is a representation of the meaning potential of language and meaning is conceived as part of a wider stratified system (discussed below). SFL suggests that language is composed of multiple different strata: phonetics, phonology, lexicogrammar (wording), semantics and context (Halliday and Matthiessen 2004, p.25). The semantic stratum, otherwise known as the ‘meaning base’, is a multifunctional system made up of three separate strands that each relate to the different types of meaning that can be expressed in the clause: the ‘ideation base’; the ‘interaction base’; and the ‘text base’ (Matthiessen et al. 2022, p.166). These three different strands map onto what Halliday (1985) names the ‘experiential’, ‘interpersonal’ and ‘textual’ metafunctions (an overview of each metafunction is presented in Section 3.2.1). In this discussion of construal, I will focus only on the ideation base, as according to Halliday and Matthiessen (1999, p.82), it contains the fundamental components of experience and the combinations of these components.

In the ideation base, experience is generalised into what are called ‘phenomena’, which represent “anything that can be construed as part of human experience” (Halliday and Matthiessen 1999, p.48). Essentially, the ideation base is a semantic model of ‘rank scale’, where experience, i.e., phenomena, is expressed in the form of one of three primary options: ‘elements’; ‘figures’; or ‘sequences’ (Taverniers 2019, p.82). Elements either represent processes (elements which embody temporal properties which unfold through time), participants (elements that either induce the process or mediate it) or circumstances (elements which specify any circumstantial information about the process, e.g., its temporal location, its cause etc.) (Halliday and Matthiessen 1999). These elements, when expressed together, combine at the configurational level to represent ‘figures’, i.e., grammatical configurations (Halliday and Matthiessen 2014, p.220). The sequential level then pertains to combinations of

these figures, that, when placed together, form ‘clause complexes’ (Halliday and Matthiessen 1999: 48). Examples for each of these phenomena are provided below in Table 2-1. Phenomena which enter the semantic system through experience are then located along a continuum of ‘delicacy’ in relation to these three levels. In SFL, delicacy refers to a “scale of differentiation”, or depth in detail (Halliday 1961, p.272). The least delicate phenomena obtain the more general, all-inclusive class of ‘phenomenon’, while the most delicate can be realised as codified lexically in English (Halliday and Matthiessen 1999, p.66). These more delicate phenomena are said to be most associated with our senses and day-to-day experience (Halliday and Matthiessen 1999, p.66). As a result, the construal of these more delicate phenomena is more readily drawn upon by the speaker.

*Table 2-1. Different types of phenomena in the ideation base, provided with examples (Halliday and Matthiessen 1999, pp.48–59).*

<b>Phenomena</b>	<b>Components</b>	<b>Example</b>
Element	Process	Run
	Participant	Jack
	Circumstance	In the morning
Configuration	Figure	Jack runs in the morning.
Sequence	A series of figures	Jack runs in the morning whilst Mary plays football.

Through the association of the semantic system with the lexicogrammar in the ideation base, the notion of typicality plays an integral part in the construal of experience, as indicated in the table above. For instance, a process in the ideation base is understood as typically construed by a verbal group in the lexicogrammar (Taverniers 2011, p.1111). This understanding of typicality is highly linked to the rank scale. When a clause such as *Jack ran to the shop* is construed by a nominal group (e.g., *Jack’s run to the shop*), the relationship between the semantics and lexicogrammar is viewed as ‘incongruent’ (nominal groups offer a more condensed construal than clausal realisations) (Taverniers 2019, p.82). This incongruency plays a large role in generating what is known in SFL as ‘Grammatical Metaphor’ (Halliday 1985). While I do not have space to cover this concept here, a full overview on Grammatical Metaphor is available in Taverniers (2003) and Halliday and Matthiessen (2014). The ideation



base, therefore, relates to the different types of entities humans conceptualise, and consequently offers a view of ‘phenomenological’ semantics (Taverniers 2019, p.82).

However, the idea of congruence between the semantic and lexicogrammatical strata has been argued to be underdeveloped due to its preconceived assumptions related to typical construal (Fontaine 2019). In the SFL framework, nouns are considered to construe objects and verbs are considered to construe processes (Halliday 2009, p.117). While this claim holds in most instances, Fontaine (2017b) argues that it does not always hold, since it would be difficult to claim that a noun like *storm* is an incongruent realisation. Foregrounding the indeterminacy surrounding the metaphorical status of event nouns, she suggests that a *storm* may be regarded as a congruent construal even though it denotes an event that transpires in time (as discussed in Section 2.1.4). Despite this critique, the semantic system proposed in SFL presents a multidimensional semantic space for new phenomenological instances to enter. Once enough semantic meaning in the form of phenomena is compiled in the ideational base, new categories can be construed (Halliday and Matthiessen 1999, p.82). Accordingly, the semantic system and the relations within it are consistently modelled and re-modelled, adapting to the ongoing changes in our social environment. Consequently, the speaker is able to bridge the relationship between ontological knowledge and language, and actively construe their experience through language (Taverniers 2011, p.1122).

Moving forward, this thesis will draw upon both CL and SFL understandings of construal when discussing the capacity of lexical items to express meaning. The decision to consider both understandings is motivated by two reasons. The first is that the SFL framework places great emphasis on the influence of one’s social environment on their use of language (Martin 2022). By focusing on the influence context may have on construal, the SFL understanding of construal is able to effectively account for the dynamicity of meaning, through posing a system of meaning that is continuously re-modelled in line with its social environment. In doing so, the framework also provides beneficial information on typical and atypical relations between strata. This sensitivity to context will offer an advantage in the ‘lexical aspect’ analysis (see Section 3.2), as it will provide a means of accurately assessing how nominals semantically function in their context of use.

The second reason is that construal is a naturally subjective process that evidently involves cognition. While the CG interpretation of construal may be critiqued for a lack of focus on the

more social aspects of language use, SFL can be critiqued for the opposing issue, as work in SFL on the cognitive processes used within language remains notably underdeveloped (c.f. the criticisms present in Thibault 2004; Butler 2013). As I will discuss in the following paragraphs, one's cognition and conception play an important role in the process of construal, as it is suggested that the semantic expression of lexical items is influenced not only by the interaction of a lexical item with its surrounding syntactic context, but also our real-world (i.e., ontological) knowledge of how the phenomena expressed by lexical items behave (Croft and Cruse 2004, p.101). By combining the views of both CG and SFL, I will try to achieve an effective balance between more "in-contexts" (e.g., SFL) and more "in-heads" (e.g., CG) (Bateman 2017, p.13) theories of language use in relation to construal.

### 2.3.2. Meaning Potential

Following on from the concept of construal to how the meaning comes to be expressed by lexical items, Hanks' (2013, p.73) posits the view that lexical items, in isolation, do not possess inherent meaning, but "meaning potential". This potential is composed of various semantic components, which are generated from our previous experience of a lexical item's linguistic behaviour and, further, our ontological knowledge of what the lexical item can potentially denote. When a lexical item is used in a given syntactic context, certain semantic features are triggered into activation and others are not, which generates specific meaning (Allwood 2003, p.29). Thus, meaning is viewed as dynamic rather than static (Hanks 2000, p.210). For Hanks, emphasis is therefore placed on the role of lexicogrammatical context for the instantiated meaning. The lexicogrammatical context, according to Hanks (2013, p.83), limits the meaning potential of the lexical item to a specific instance of meaning.

To better understand the concept of meaning potential and the idea of semantic components, we can look at an example originally provided by Alexiadou and Grimshaw (2008) which was repurposed by Grimm and McNally (2013) in their paper on the class status of 'argument-structure' nominals. Grimm and McNally (2013) show that the nominal *examination* can refer to either an object or an event, dependent on its construal. While *examination* in example (10a) below refers to a physical object, e.g., an *examination* printed on a piece of paper, *examination* in example (10b) refers to an event, e.g., an *examination* organised to take place at a specific time (Fontaine 2017b, p.7).

- (10) a. The examination on the table.  
b. The examination took three hours.

Here, there are two principal notes of interest. The first is that, in example (10a), *examination* is not an event being construed as an object. The second is that the potential for *examination* to be construed as either an event or object is seemingly the result of the syntactic construction (i.e., the linguistic co-text) the nominal is present in, and not the nominal itself. Using Hanks' (2013) conception of meaning potential, I suggest that in examples (10a-b) below, different semantic features of *examination* were activated, as a consequence of the nominal's interaction with the different syntactic constructions. For instance, in (10a) I could say that the semantic feature of 'physicality' is activated by the prepositional phrase *on the table*, as we are aware that physicality is generally required for an entity to be present *on* a physical entity, e.g., *the table*. Alternatively, in (10b) the semantic feature of 'temporality' is activated by the temporal post-modifier *took three hours*, as temporality is required for an entity to take time. These two semantic features are of course only two of many different activated semantic features that help distinguish the two meanings of *examination* shown in examples (10a-b). Nevertheless, what they show is that the activation of semantic features, and thus the expression of meaning, is highly dependent upon the influence of surrounding syntactic context (Hanks 2020).

However, the expression of meaning is not solely dependent on the syntactic constructions that lexical items enter. To claim this would be a slight overgeneralisation, as meaning potential is not limitless. It is largely maintained by the "nature of reality" (i.e., our ontological, 'real-world' knowledge) (Croft and Cruse 2004, p.101) and this maintenance can influence the typical use of lexical items. The first time I truly acknowledged the interaction between the influence of syntactic context and ontological knowledge and its effect on the expression of meaning was when I encountered the example in (11) *I didn't jump so quickly* in a colleague's MA thesis. This construction is particularly interesting, as there is a tension between the negated adverbial phrase and the naturally explosive action of jumping. Because of our ontological knowledge, we are aware that a *jump* (especially when carried out by a human agent) is typically an instantaneous event, and, as a result, the adverbial phrase does not appear to influence the meaning expressed by the verb. Instead, I propose that we are more likely to interpret example (11) as the 64-year-old was slower to begin the jump action, as opposed to executing a slow jump.

(11) ...I'm nearly 65 I didn't jump so quickly. (Collier 2021, p.56)

However, the expression of meaning is highly dynamic and, consequently, lexical items do not always adhere to their typical usage and meaning expression. In addition to meaning potential, Hanks proposes the term 'exploitation' which relates to instances that are intended deviations from the conventional use of a lexical item. Over time and through consistent use, these deviations in the meaning of the lexical item can become conventionalised within certain contexts (Hanks 2013). And, as our memory is influenced both qualitatively and also quantitatively by experience, these deviations can become more readily accessible within one's mind (Hudson 2010, p.70), leading to a reduction in cognitive demand for the speaker. This discussion on the influence surrounding syntax has over the semantic expression of lexical items will be revisited in the following section on the framework of lexical aspect and more extensively in Chapter 7, in relation to the semantic flexibility of the nominals *drought*; *fire*; *game*; and *stream*.

Both Hanks' conceptions of meaning potential and exploitation provide intuitive understandings of how meaning is expressed. By detailing how the components of a lexical item's meaning potential can combine, we can economically account for the diverse range in which lexical items are used (Hanks 2013, p.72). Nevertheless, while the acknowledgement of meaning potential and exploitation is certainly helpful in our quest to understand the nature and degree of nominality and how nominals can express certain meanings, it does not provide us with empirical data on how the semantics of nominal lexical items generally behave. One approach to viewing semantics that can facilitate empirical access to the semantic behaviour of these nominals in use is 'lexical aspect'.

## 2.4. Lexical Aspect

Aspect, as a term in the linguistic sense, traditionally concerns the different "ways of viewing the internal temporal constituency of a situation" (Comrie 1976, p.3). This definition distinguishes aspect from 'tense', as tense is defined as a linguistic notion "which denotes the form taken by a verb to locate a situation in the referred time" (Declerck 2006, p.22). While tense attempts to locate a situation to a specific temporal reference point, e.g., 'time of utterance', aspect focuses on the structural features of the situation itself (Rothstein 2004, p.1). Since the field of aspect is extensive, it is often difficult to fully grasp its coverage due to the

range of differing terminology and theories made popular by different scholars, e.g., Vendler (1957; 1967); Smith (1991); Rothstein (2004; 2008a); Declerck (2006); and Croft (2012). To simplify matters, the aspectual framework discussed will predominantly draw from the work of Smith (1997), as her five part situation type distinction provides a foundational understanding of the lexical aspect framework that additionally considers the role of Semelfactive situation type (see Section 2.4.3 for a discussion on situation types and the differences between them).

There are generally considered to be two types of aspect: ‘grammatical aspect’ and ‘lexical aspect’. Work in these two areas has traditionally focused on verbs. However, as briefly discussed in Chapter 1 and later revisited in Section 2.4.4, the focus of lexical aspect has recently begun to stretch into the nominal domain. In the following sections, I will tease these frequently intertwined types of aspect, before focusing on the features of lexical aspect and the ‘situation types’ they form through different feature combinations. This focus on lexical aspect is particularly important, as the framework will be frequently drawn upon throughout this thesis.

### **2.4.1. Grammatical Aspect and Lexical Aspect**

Grammatical aspect primarily explores the verb’s location in relation to a temporal reference point (Li and Shirai 2000, p.11). Often referred to as ‘viewpoint aspect’, grammatical aspect relates to the temporal viewpoint from which one ‘views’ a situation, by focusing one’s attention on a particular time of that situation, e.g., the ongoing process (Van Hout 2016, p.587). Grammatical aspect is what Smith (1997, p.39) would describe as an overt linguistic category, as grammatical aspect is often signalled through the use of inflectional and derivational morphology, auxiliaries, or periphrastic constructions (Li and Shirai 2000, p.11). Grammatical aspect contains three principal viewpoints, which each construe situations differently in relation to time: Perfective (bounded); Imperfective (unbounded) and Habitual. Perfective viewpoints view the situation as a whole, taking into account both its initial and final points (Smith 1997, p.3), as shown in example (12a) below. Imperfective viewpoints, alternatively, only centre on the ongoing nature of the situation, neglecting both its initial and final points (Smith 1997, p.3), as shown in example (12b). Habitual viewpoints, lastly, relate to the construal of a situation as continuously instantiated over a given time frame, as opposed to focusing on one specific instantiation (Binnick 2006, p.248), like in example (12c).

- (12) a. Jack baked a cake.  
b. Jack is running.  
c. Jack runs to work.

Contrary to grammatical aspect, lexical aspect covers the ‘inherent’ temporal structures denoted by individual lexical items (Smith 1997). These structures have traditionally concerned the existence of temporal restrictions and boundaries within the lexical composition of the verb (Filip 2012, p.721). As a result of this emphasis on verbs, most of what we know about lexical aspect stems from the study of verbs. Nevertheless, this pattern has recently begun to shift, with more contemporary work also exploring lexical aspect expressed by other lexical classes, such as nouns (Balvet et al. 2011; Fábregas and Marín 2012; Huyghe et al. 2017; Heyvaert et al. 2019). Lexical aspect, in relation to nouns, will be specifically addressed in Section 2.4.4.

Unlike grammatical aspect, lexical aspect is, according to Smith (1997, p.39), a “covert” linguistic category. By this, Smith draws attention to the idea that the temporal structures expressed by individual lexical verbs are not directly encoded by grammatical morphology but are available in the usage of the verb and the restrictions placed on its construal by our knowledge of the world. These structures take the form of binary features which centre around the verb’s ‘Dynamism’, ‘Durativity’ and ‘Telicity’ (Vendler 1957; 1967). We will return to these features and explicate them later in Section 2.4.2.

Traditionally, lexical aspect suggests that the temporal structures expressed by verbs are a product of their inherent semantics, independent of the verb’s instantiation (Vendler 1957; 1967). Despite this suggestion, it is apparent that the temporal structures expressed are typically not the product of the verb alone, but the product of the complete verb phrase (Verkuyl 1972). Smith (1997, p.17) furthers this position, noting that lexical aspect ultimately resides in “verb constellations”, including the verb’s arguments and other potential predicative elements. The lexical aspect of a verb is largely dependent upon the compositionality of the phrase in which it enters. This is evidenced by the fact that the same verb can exhibit different temporal structures when placed in different syntactic constructions. In the examples below, the same verb *think* can express both a dynamic event (13a below) and an unchanging state (13b). This is known as an instance of “aspectual coercion” (Dickey 2016, p.340), as the alteration in argument structure is able to ‘coerce’ different lexical aspect features in the verb phrase.

Accordingly, throughout this thesis, when referring to verbs (and later nouns) as expressing lexical aspect features, I will be referring to verbal/nominal contextual “situations” (Smith 1991) in which the verbs/nouns are present, encompassing both the verb/nouns and its surrounding syntactic context.

- (13) a. Jack sometimes thinks about his friends.  
b. Jack thinks that he is clever.

The temporal structures expressed by these situations provide evidence for treating grammatical aspect and lexical aspect as distinct linguistic concepts that regularly intersect. Olsen (1997, p.16) suggests that grammatical aspect should be interpreted as an “overlay” on lexical aspect, which seems to describe the relationship between the two appropriately. The way they regularly intersect appears to stem from the role of different viewpoints in influencing the lexical aspect features of a given situation. For instance, while in example (14a) below, the imperfective viewpoint draws our attention to the ongoing dynamic nature of washing the dishes, the perfective viewpoint in example (14b) construes a dynamic event as a whole which has reached its endpoint. The distinction between the two examples concerns the lexical aspect feature of ‘telicity’, which will be defined and explained in the next section, along with the lexical aspect features of ‘dynamism’ and ‘durativity’.

- (14) a. Jack is washing the dishes.  
b. Jack washed the dishes.

### **2.4.2. Lexical Aspect Features**

The feature of dynamism distinguishes situations that are dynamic from those that are static. Dynamic situations typically involve some kind of change, motion or activity which requires a continuous input of energy from an agent external to the process (Comrie 1976, p.49). Dynamic situations also typically represent situations which do not occupy every moment within a given timeframe. For instance, in example (15a) below, we will generally interpret, due to the fact that the average human body is only capable of so much exercise, that Jack was not swimming for the entirety of last year, but for a period (or periods) of time which occurred within the timeframe of last year. Static situations, in contrast, represent unchanging situations which do not require an input of energy from an external agent to persist. Moreover, unlike dynamic

situations, static situations generally hold over time. This capacity to ‘hold’ over time is shown in example (15b) Although not a typical construction, we can infer that there was not a single moment in that year that Jack did not love Mary.

- (15) a. Jack swam last year.  
b. Jack loved Mary last year.

In addition to the feature of dynamism, there are also the features of durativity and telicity. Durativity differentiates ‘durative’ situations from ‘punctual’ situations. Durative situations are situations which exist through time and are the result of many distinguishable phases, as in the example below (16a). Punctual situations form the other side of the spectrum of duration, as they are instantaneous single-phase events, that only require a moment to actualise, like example (16b). Although for punctual situations we use the term instantaneous, the instantaneity of these situations is highly idealised, as many of the situations that are punctual, at the very least, still take several milliseconds to transpire. Nevertheless, it is their ability to transpire in an expeditious singular phase that is indicative of their punctual nature.

- (16) a. Jack climbed the wall.  
b. Jack won the race.

Lastly, the feature of telicity classifies ‘telic’ situations and ‘atelic’ situations. Telic situations express an inherent endpoint, from which the situation cannot continue beyond. If we look at the situation expressed in example (17a) below, we can deduce that once Jack has walked to work, he can no longer be ‘walking to work’; the event has terminated, i.e., reached its internal endpoint. Rothstein (2004, p.26) proposes that such situations carry a “scalar implicature”, where the measure of the eventuality in question is equal to the distance/time/temperature etc. stated. Atelic situations, alternatively, do not express such a scalar implicature, and will continue indefinitely, unless arbitrarily interrupted. For instance, if we do not assign an arbitrary endpoint to the situation presented in example (17b), Jack will continue to run, and will still be running long after this thesis has ended.

- (17) a. Jack walked to work.  
b. Jack is running.



When combined, these three lexical aspect features (dynamism, durativity, and telicity) are employed to distinguish five distinct ‘Situation Types’ (Smith 1997, p.19): ‘States’; ‘Activities’, ‘Accomplishments’, ‘Achievements’ and ‘Semelfactives’.

### 2.4.3. Situation Types

The concept of situation type, or ‘Aspectual Class’, was proposed by Vendler (1957; 1967) who discerned four ‘Time Schemata’ (situation types), which each encompass a divergent combination of the abovementioned lexical aspect features. The four time schemata differentiated by Vendler (1957; 1967) are: States; Activities; Accomplishments; and Achievements. In addition to these four, Smith (1991) proposed a fifth category of Semelfactive, to account for dynamic punctual events that are atelic (this situation type will be greater explained later in this section). These five situations types are divided into two categories, based on the situation’s dynamism (Smith 1997): state situation types (States), and event situation types (Activities, Accomplishments, Achievements and Semelfactives).

States constitute the only static situation type as they require the presence of an external agent to induce change. They are also durative, as they typically hold for extended periods of time, although this has been debated (cf. Murphy 2010, p.210). They do not, however, engage with the lexical aspect feature of Telicity. Smith (1997, p.32) claims that States do not take time, they simply hold for a period of time. In this sense, the initial and final endpoints of a state do not form part of the state itself, they instead represent distinct situations, and indicate changes of state (Smith 1997, p.32). Despite this understanding, States have still been conceptualised as continuous by authors such as Declerck (2006, p.71), due to their capacity to hold over time. As a result, there will be instances later in this thesis (when discussing nominal State instances) where we talk about States expressing ‘continuity’. A typical State situation type is presented in (18a) below. Activities, alternatively, comprise dynamic events which typically involve animate beings and “movement, activity and volition” over a sustained duration (Smith 1997, p.24). Moreover, Activities are atelic, as the description of an Activity event does not imply an endpoint. They express the potential to continue indefinitely, as shown in example (18b) below. Like Activities, Accomplishments are also dynamic and can take place over a duration of time. Accomplishments differ from Activities as they express an inherent endpoint; they are telic. By adding the complement *to the park* to example (18c), we can aspectually coerce the Activity into an Accomplishment, by providing an endpoint to Jack’s cycle, as shown in (18c).

Achievements, in contrast to Activities and Accomplishments constitute highly dynamic events whose occurrence is instantaneous. They represent events that are dynamic, punctual, and telic, as shown in example (18d). Achievements, in particular, often pose challenges to lexical aspect analysts, as certain Achievement verbs, e.g., *shoot*, can appear in the progressive, e.g., *shooting*, contradictory to their supposed punctual nature. These situations are called ‘Durative Hypersituations’, i.e., situations “consisting of a number of repeated punctual sub-situations” (Declerck 2006, p.36). For instance, *Jack is shooting the birds*, although made up of multiple punctual actions, represents a durative situation which will take more than a moment to actualise. They are durative events which originate from typical telic and punctual situations (Heyvaert et al. 2019, p.42). Following the recent work of Heyvaert et al. (2019), such instances should not be analysed as Achievements due to their durative nature, but as either Activities or Accomplishments, dependent on the situation’s telicity. Lastly there are Semelfactives. Semelfactives are dynamic, punctual, and atelic events. They represent single-stage events which do not result in an outcome other than the expression of the event itself (Smith 1997, pp.29–30), as in example (18e) below.

- (18) a. Jack loves Mary.  
 b. Jack cycled.  
 c. Jack cycled to the park.  
 d. Jack died.  
 e. Jack sneezed.

These distinctions between the five situation types are outlined in Table 2-2.

Table 2-2. *Lexical aspect features of situation types (Smith 1991).*

Situation Type	Dynamic/Stative	Durative/Punctual	Telic/Atelic
States	Stative	Durative	N/A
Activities	Dynamic	Durative	Atelic
Accomplishments	Dynamic	Durative	Telic
Achievements	Dynamic	Punctual	Telic
Semelfactives	Dynamic	Punctual	Atelic

Interestingly, work has begun to challenge the boundaries of the five main situation types (Dowty 1979; Radden and Dirven 2007; Rothstein 2008b; Balvet et al. 2011). Radden and Dirven (2007) suggest that somewhat intermediary lexical aspect classes can occur through the influence of the situation's grammatical aspect. Most notably, they propose the classes of 'Accomplishing-Activities' and 'Culminating-Activities'. Accomplishing-Activities relate to situations, such as (19a) below, where the viewpoint imposed by the progressive aspect is able to provide an Activity focus on this typical Accomplishment situation. We know from our real-world knowledge that we cannot change a nappy indefinitely. However, in this example, Radden and Dirven (2007, p.184) note that the endpoint of the event, i.e., when the nappy has been completely changed, is not present in the viewpoint structured by the progressive aspect. The endpoint is not reached in the construction, leading to its supposed interpretation as an Accomplishing-Activity. Culminating-Activities, alternatively, such as (19b), are able to express extended singular situations due to their progressive aspect, which, when reaching their internal endpoint, will culminate at an Achievement reading (Rothstein 2004, p.37; Radden and Dirven 2007, p.188). Due to the viewpoint structured by the progressive aspect, an Activity focus is imposed on this typical Achievement situation, as the endpoint is not presented in the viewpoint of the situation. Smith (1991, p.237), although not specifically subcategorising these situations into a separate lexical aspect class, shares a similar view about the behaviour of these situations, claiming that these progressive Achievements focus on "detachable" preliminary stages of the Achievement.

- (19) a. Jack is changing the nappy.  
b. Jack is dying.

#### **2.4.4. The Inclusion of Nominals**

As well as the acknowledgement of intermediary lexical aspect classes, recent literature on lexical aspect has begun to realise the potential for lexical aspect to extend into different lexical classes, as noted in Section 2.4.1. Huyghe (2017, p.118) suggests that lexical aspect is not a verb specific concept, and is principally dependent on the denotation of eventualities, regardless of lexical class. In particular, there has been a recent increase in the investigation of temporal semantics in the nominal domain (Balvet et al. 2011; Fábregas and Marín 2012; Huyghe et al. 2017; Heyvaert et al. 2019). Unlike verbs, however, nominals predominantly

denote ‘Objects’ which generally do not express temporal semantics, so do not belong to the lexical aspect domain (Fábregas and Marín 2012, p.36), shown below in example (20a). Nevertheless, ‘States’ and ‘Events’ can still be distinguished in nominal form. State nominals, like their verbal counterparts, are static entities, which typically transpire over time, such as example (20b). On the other hand, event nominals, which encompass the lexical aspect feature combinations of Activities, Accomplishments, Achievements and Semelfactives, are dynamic nominals which occur or take place, such as example (20c).

- (20) a. The table.  
b. Jack’s anger at Mary.  
c. Jack’s run.

Despite the large collection of research into lexical aspect, it is surprising that relatively little research has worked with corpus data, analysing language in its actual use. Instead, many scholars have relied upon the use of manufactured examples to substantiate their claims. One particular application of lexical aspect within both the verbal and nominal domain that makes use of corpus data is Heyvaert et al’s. (2019) study into the lexical aspect features of nominal and verbal gerunds in present-day English. In their study, Heyvaert et al. (2019) analysed 500 nominal gerunds and 500 verbal gerunds from the British National Corpus (2018) and Corpus of Contemporary American English (2018) for the lexical aspect features of dynamism, durativity and telicity (although how this was achieved is not explained explicitly). In doing so, they found that, as well as designating a high percentage of Activities (58.4%), both nominal and verbal gerunds were also able to express Accomplishment (18.3%), Achievement and Semelfactive (20.8%) and also State (2.5%) situation types. Such a finding foregrounds the capacity of both nominal and verbal gerunds to express five different situation types, somewhat contrary to Brinton’s (1998) assertion that the *-ing* suffix predominantly transforms situations into Activities.

To a lesser extent, underived nominals (nominals which have not morphologically derived from a different lexical class) have also been subject to investigation for their ability to express different lexical aspect features. Huyghe et al. (2017) explore the temporal semantics of underived nominals in French using lexical aspect. From a sample size of 3489 underived nominals, collected from the *Lexique3* lexicon ([www.lexique.org](http://www.lexique.org)), they identified 283 underived nominals in French that were able to express at least one lexical aspect event feature

(e.g., either dynamism, duration or telicity) after analysing each underived nominal in a series of diagnostic syntactic tests. For instance, they claim that *riot*, which is a morphologically underived nominal in French, expresses atelicity, as it does not express an inherent temporal endpoint.

However, while Huyghe et al. provide clarity on the lexical aspect features certain French underived nominals can express, their paper does not explore the potential correlations between situation type and particular underived nominals or even between situation type and underived nominals as a word formation class. Furthermore, as Huyghe et al.'s data was collected from a lexicon and not a corpus. As a result, the nominals were analysed in isolation, without considering the influence of the rest of the clause. Accordingly, the underived nouns were not analysed for their exact meaning in use. As a result, some of the lexical aspect features exhibited by certain nominals, while syntactically acceptable, may have been largely uncharacteristic. Nevertheless, both the above studies indicate the potential lexical aspect poses as an analytical resource in the analysis of the variable semantic behaviour displayed by nominals.

## **2.5. Summary and Aims**

From the literature covered over the last chapter, it is apparent that prototype categories are largely beneficial, as they provide a flexible alternative to more rigid classification through offering a space for lexical items which were previously rejected from categorical membership (Aarts 2006, p.363). Despite the increase in inclusivity, there appears to be an excessive focus on the prototypical instances within these categories. This excessive focus, in turn, has led to less attention being provided to more peripheral instances, which have frequently only been ambiguously referred to as 'less prototypical'. Language use is inherently dynamic. The inherent dynamicity of semantic behaviour naturally causes a tension with the very idea of lexical classes, which, at large, remain static entities. By working from the standpoint of a continuum of similarity, we are restricting ourselves in only examining peripheral instances based on their similarity with prototype instances. Accordingly, the exact nature and degree of nominality in English currently lacks clarity when we only look through the lens of Prototype theory.

However, as shown in Section 2.3.2, meaning potential represents a helpful heuristic to look beyond similarity and explore the multidimensional semantics of nominals at a level that cannot be attained from only looking at classifications of lexical class. Nevertheless, while certainly intuitive, it does not provide us with empirical data on the semantic behaviour of nominals in use. The concept of lexical aspect has shown that it can potentially provide a beneficial analytical tool for analysing the semantic behaviour of nominals, as it can enable the analyst access to the dynamic semantic behaviour that nominals can potentially express. It is therefore rather surprising that only very few studies have explicitly explored the dynamic semantic behaviour of nominals within a large sample, using lexical aspect. Moreover, while studies using lexical aspect in the nominal domain have largely focused on the identification of temporal semantics in different nominal forms, significantly less work has been devoted to the examination of the specific character of these temporal semantics across different nominal forms, and the syntactic contexts which support them. For example, the temporal semantics expressed by adjectival nominalisations may typically express States more than deverbal nominalisations. Furthermore, such a study has yet to be composed which also explicitly focuses on the temporal semantics of underived nominals in their actual use. Therefore, through using lexical aspect, the central aim of this study will be to carry out a detailed examination into the behaviour of nominal semantics and how different temporal meanings are expressed. To achieve this aim, three sub-aims will be considered: (1) to determine how event and state meaning come to be expressed in nominal forms; (2) to examine the potential relationships between the syntactic and semantic properties of underived event nominals (UENs); and (3) to determine how the nature and degree of nominality can be evaluated in English.

## **Chapter 3: Methodology for Investigating Category Feature Influence on Nominal Semantics**

Over the last chapter, we have seen that, while notions of prototypicality have significantly advanced our understanding of lexical class categories (namely noun and verb), the dynamic nature of language use has proven the semantic classification of noun and verb to be less clear cut than once conceptualised. Specifically, it was argued that, to gain a clearer understanding of lexical class, we must not always categorise nominals on their supposed prototypicality alone. If we only classify nominals on a continuum of prototypicality, we restrict ourselves by only categorising more peripheral nominals on their similarity with prototypical nominals. In doing so, more peripheral nominals are often ambiguously regarded as ‘less prototypical’, and consequently regularly lack the clarity and attention given to supposed prototypical nominals. Furthermore, a better understanding of how category features function in the establishment of prototypicality is needed. Accordingly, if we are to accurately understand the nature and degree of nominality, we must adopt a finer approach to investigating lexical class that extensively explores word meaning.

For this reason, Hanks’ (2013) concept of meaning potential was introduced in Section 2.3.2. Through reviewing the notion of meaning potential, we were able to gain a greater understanding of the dynamic nature of word meaning and how different meanings come to be expressed by lexical items. However, despite its intuitive nature, the concept does not provide empirical access to the dynamic semantic behaviour of nominals. The concept of lexical aspect was therefore brought forward in Section 2.4, as it was argued that analysing nominals for their expression of lexical aspect could bridge this gap, by providing empirical access for the analyst to the semantic behaviour of nominals. Accordingly, through analysing nominals in accordance with the framework of lexical aspect, the three sub-aims of this thesis are as follows (as stated at the end of Chapter 2):

1. To determine how event and state meaning come to be expressed in nominal forms.
2. To examine the potential relationships between the syntactic and semantic properties of underived event nominals (UENs).
3. To determine how the nature and degree of nominality can be evaluated in English.

To address these aims, this thesis will conduct two separate studies. In Section 2.4, I noted that while studies using lexical aspect in the nominal domain have largely focused on the identification of temporal semantics in different nominal forms, significantly less work has been devoted to the examination of the specific character of these temporal semantics across different nominal forms. Accordingly, to cover this unaddressed territory in the literature, the first study will focus predominantly upon the nature of nominal semantics by assessing the extent to which category features such as abstract/concrete status, count/mass status, word formation type, and also contextual changes such as genre, influence nominal semantic behaviour. Chapters 3, 4 and 5 will cover the first study. This chapter will cover the each of the methodological decisions made, whilst Chapter 4 will note the results of the statistical analyses (for information on the statistical analyses used, see Section 3.6). Chapter 5 will then present a discussion of how the results of the study provide insight into our understanding of the nature and degree of nominality. The second study (described and discussed in Chapters 6 and 7) will also focus on nominal semantics, but by examining how the semantics of four selected underived event nominals are influenced by their surrounding syntax.

In the following paragraphs, I will now describe the corpus and the sample that were used in the first study of the thesis. Afterwards, I will outline the annotation scheme used, detailing how nominal instances were coded for their ‘experiential’ meaning their expression of lexical aspect, their abstract/concrete status, and their count/mass status. Section 3.3 will then move to discuss the reductions made to the sample after this annotation and Section 3.4 will provide information on how each nominal instance was additionally classified in relation to its word formation type. Penultimately, Section 3.5 will cover how the data was prepared for the statistical analyses, with specific reference to how repetition instances (nominal instances that occurred more than once) were handled within the sample. The chapter then concludes with Section 3.6, which outlines the hypotheses of the study and the statistical analyses that were conducted using the frequencies from each of the nominal instance annotations.

### **3.1. Corpus and Sample**

The corpus used for the current study was the British National Corpus (BNC 2020). The BNC, originally developed by Oxford University Press, is a relatively large corpus of British English, consisting of a wide range of texts and 96,134,547 words. The texts within the corpus were



taken between 1960-1993. The corpus predominantly consists of written English (which constitutes 90% of the corpus) and spoken English, constituting 10%. The texts within the corpus exhibit a large variety of genres, including 47 distinct written genres and 24 distinct spoken genres. These written genres range from newspaper writing, fiction writing, email writing etc. to a number of academic writing genres, e.g., academic natural science writing.

The BNC was chosen for several factors. Firstly, it presents a readily accessible and easy to navigate corpus, using the corpus manager and text analysis software Sketch Engine (Kilgarriff et al. 2004). Moreover, the BNC presents a large corpus, so provides a plentiful amount of data to identify possible relations between nominal form, semantic behaviour, count/mass status and concrete/abstract status. Furthermore, the large number of genres within the corpus facilitates the possibility for a cross-genre comparison of the nominals, to explore the potential influence that genre conventions may have on nominal semantic behaviour.

The genres chosen in the current study were ‘Academic Natural Science Writing’; ‘Broadsheet Newspaper Home and Foreign News Writing’; ‘Print Advertisement Writing’; and ‘(Auto)Biographical Writing’. Each of these genres was chosen on the assumption that they would provide interesting data in relation to the semantic behaviour of nominals – specifically the expression of event meaning. For instance, scientific writing is claimed to exhibit an orientation towards the use of nominalisation, as nominalisations can express clausal meaning in the form of a nominal (Halliday and Martin 1993). Thompson (2014b, p.239) claims that these nominalisations regularly take the form of what he labels ‘specialised lexis’, that allows writers to efficiently indicate the results and conclusions of events. e.g., *extrapolation*. On the other hand, ‘Broadsheet Newspaper Home and Foreign News Writing’, especially that which concerns ‘hard news narratives’, regularly focuses upon “ever-unfolding social events” (Richardson 2007, p.71). In the discussion of these social events, due to the information density of the texts, ‘nominalising language’ is typically employed, where verbs and adjectives are expressed syntactically as nouns (Fowler 1991, p.79).

The genre of ‘Print Advertisement Writing’ was also chosen due to its supposed orientation to event meaning. Danesi (2015, p.38) claims that adverts regularly aim to create associations between the products they are attempting to sell with other aspects of life and society, of which, significant life events are prevalent (e.g., romance). Moreover, Bloor and Bloor (2007, p.145) show that advertisement writers will often attempt to create a sense of urgency for the

consumer, to increase the likelihood of the consumer purchasing the product in question. In doing so, writers may orient to event nominals that express ‘selling events’, that only last a certain amount of time. For instance, ‘the sofa *sale* ends this Friday’.

Lastly, the genre of (Auto)Biographical Writing was chosen to provide a more ‘neutral’ genre, that did not specifically orient to nominalisations and other event nominals. As the broad purpose of (auto)biographies is to recount the stories of an individual’s life, it was anticipated that (auto)biographical writing may share similar conventions with narrative writing, where the use of nominalisation is somewhat counteractive to the essence of narrative. This view is taken as nominalisations are considered to “de-narrativise” processes (Toolan 2001, p.224), by excluding clausal elements such as finiteness and, commonly, participants. As a result, the narrative elements are stripped back, and the clause becomes background to the product or thing, i.e., the nominalisation (Toolan 2001, p.224).

Moving back now to detail the corpus and sample used, the BNC (2020) is lemmatised and includes linguistic annotation for lexical class information (using automatic part-of-speech (POS) tagging). Using POS information, millions of nouns were identified. The random sample generator and download function included in SketchEngine were then used to create and download a sample of 5000 nominal instances, comprised of the four genres detailed above. The nominal token distribution of this random sample across the four genres is present in Table 3-1 below. In Table 3-1, there is a somewhat disproportionate distribution of genre within the random sample. However, as we can see in Table 3-1, this distribution displays an approximate reflection of each genre’s prevalence in the BNC with regards to token nominal instances. A total of 5000 nominal instances was chosen to provide a large enough number to attain reliable results. The decision to extract 5000 nominal instances was also based on the expectation that many of these instances would have to be excluded for various reasons, e.g., if they did not function in the ‘syntactic head’ position of their linguistic context. Further detail on exclusions is given in Section 3.3, which addresses reductions made to the sample.

*Table 3-1. Distribution of nominal instances collected in relation to their genre.*

Genre	Number of Instances	Number of Instances in Entire BNC
Academic Natural Science Writing	1002	349507
Broadsheet Newspaper Home and Foreign News Writing	621	222449
Print Advertisement Writing	609	206092
(Auto)Biographical Writing	2768	966465

Once collected, the 5000 instances were placed into an annotation scheme created to analyse the nominal instances for their expression of lexical aspect. Before the lexical aspect annotation was conducted however, the lexicogrammar of each nominal was analysed, following the framework of Systemic Functional Linguistics (SFL) (Halliday 1985). In the next section, I will briefly review the SFL analysis undertaken. Afterwards I will move onto the description of the lexical aspect annotation.

## **3.2. The Nominal Annotation Scheme**

### **3.2.1. SFL Annotation of Nominal Instances**

While SFL was introduced in Section 2.3.1, in this section I will briefly review the framework again, but this time with a focus on how SFL conceptualises the clause, to explain why SFL was used in the current study to understand nominal grammatical behaviour within its context of use.

SFL is a heuristic theory that treats language as a semiotic system that speakers employ to create meaning within particular sociocultural contexts (Bloor and Bloor 2018, p.151). This semiotic system is comprised of options, and meaning is understood to arise through the patterned arrangement of these options in use, that are expressed through units of the given language's grammar (Hao 2018, p.527). One specific unit of language, the clause, is conceived within SFL as multifunctional (Halliday and Matthiessen 2014). In SFL, the clause exhibits three distinct strands of structure that each simultaneously convey separate meanings (Thompson 2014b, p.30). These distinct strands are known as the 'Experiential', 'Interpersonal' and 'Textual' metafunctions.

The experiential metafunction focuses on the use of language to talk about one's experience of the world, including the worlds within one's mind, to describe the events, states and entities that are present within them (Thompson 2014b, p.24). The interpersonal metafunction, on the other hand, concerns how language is used to convey one's relationship with other individuals, through expressing emotion, attitude and judgement (Bloor and Bloor 2013, p.13). Lastly, the textual metafunction relates what is said or written "to the rest of the text and other linguistic events" (Bloor and Bloor 2013, p.13). In the nominal annotation scheme, only the experiential metafunction was examined. This decision was made on the basis mentioned above that the experiential metafunction principally concerns the expression of events, states and entities (Thompson 2014b, p.24). The experiential metafunction therefore provides a useful resource for uncovering the dynamic semantic behaviour expressed by nominal instances. Alternatively, the interpersonal and textual metafunctions were not examined, as the research conducted in this thesis principally relates to the expression of events, states, and objects.

The examination of the experiential metafunction allowed for the functional analysis of the 5000 nominal instances in their social environment (Hasan 2009, p.37). Through analysing how the nominals functioned within their context of use, I was able to attain an understanding of how each nominal was construed (see example below). As a result, each applicable nominal instance (see Section 3.3 for information on reductions to the sample) was annotated for its expression of lexical aspect. The need for this functional analysis is demonstrated well by an example such as *party*, which can express a notably different meaning, dependent on its surrounding syntax. Given its surrounding syntax, the nominal *party* can express both object and event meaning. This difference in construal is evidenced in examples (21a) and (21b) below. While the *party* in (21a) expresses no kind of temporal meaning, typical of object meaning, the *party* in (21b) can 'take place' over time, typical of event meaning.

- (21) a. The political party suffered a heavy defeat.  
b. The party took place.

Within the nominal annotation scheme, three distinct analyses of experiential meaning were applied to each nominal instance. First, the experiential function of the nominal instance was analysed. This first analysis was undertaken to obtain information on how each nominal instance was experientially expressed within the nominal group. The central purpose of this analysis was to distinguish nominals which denoted a 'thing' from nominals which functioned

as divergent elements within the nominal group that could not be accurately analysed for lexical aspect due to their functional role, such as classifiers (Balvet et al. 2011, p.1). For a definition of classifiers and information on their exclusion from the dataset, see Section 3.3.

The second analysis of experiential meaning concerned how the nominal group, in which the nominal instance was situated, was represented within the clause. Fontaine (2013, pp.48; 93) claims that the nominal group can function in a variation of semantic roles, based on its representation within the clause. This analysis was undertaken to monitor the influence of these different clausal semantic roles on the semantics of nominal instance in question. After all, I show throughout this thesis that the construal of lexical items can be largely impacted by their surrounding linguistic context (e.g., the clause).

Lastly, the third analysis of experiential meaning set out to uncover the ‘next higher unit’ of the nominal group in which the noun was functioning. For example, a main clause would be the ‘next higher unit’ of a nominal group that was situated within a main clause, e.g., *Jack went to the house*. Thompson (2014b, p.22) notes that different groups within the clause typically express a range of functionalities. Accordingly, this analysis was designed to monitor whether the functionality of higher rank groups within the clause influenced the nominal instance’s semantic behaviour. Figure 3-1 below provides a visual representation of these three analyses, provided with the example nominal instance *worms*. After the SFL annotation was completed, the nominal instances were then annotated for their expression of lexical aspect features.

Context	Exp Function of Noun	Exp Function of Nominal Group	Next Higher Unit of the Nominal Group
The adults are slender reddish-brown <u>worms</u> up to 1cm long.	Thing	Participant in Clause	Clause = Main

Figure 3-1. Three types of experiential metafunction analysis, provided with context and example (Noun 83, All Data without Repetition Dataset, Appendix A).

### 3.2.2. Lexical Aspect Annotation of Nominal Instances

The most typical method of analysing the expression of lexical aspect is through the use of diagnostic syntactic tests (DSTs) (Kenny 1963; Vendler 1967; Dowty 1979; Smith 1991; Rothstein 2004; Declerck 2006; Balvet et al. 2011). DSTs signal whether a situation can

satisfactorily function in particular syntactic constructions that indicate the presence of certain lexical aspect features. They provide a method of attaining empirical evidence on the lexical aspect features of situations. Due to this ability to gain empirical data, DSTs were employed within the current study to reliably analyse the semantic behaviour of the 5000 nominal instances. Furthermore, as noted in Section 2.4.4, nominals can also display atemporal structure, through the expression of object meaning. Accordingly, DSTs were also used to uncover atemporal semantic structures. Taking both temporal and atemporal semantic structures into account, the DSTs aim to capture the complete range of situation types discussed in Section 2.4 on lexical aspect. However, it should be noted in advance that the analysis (described in Section 3.6) will only report on the situation types that occurred within the lexical aspect annotation of the nominal instances.

Eight DSTs were used to annotate the lexical aspect features of the 5000 nominal instances. These DSTs are shown in Table 3-2 below. Two of the DSTs were taken from Dowty (1979) and a further six were taken from Balvet et al. (2011). Overall, these eight DSTs were chosen as, together, they provide a typology of nominals, distinguishing event, state, and object nominals from one another.

The six DSTs from Balvet et al. (2011) were chosen as they provide a succinct collection of DSTs for nominals that are appropriate for English and convey information on the lexical aspect features of dynamism and durativity (Balvet et al. 2011). For example, if a nominal instance can satisfactorily function in the construction posed in Test1, then it indicates that the situation expressed by the nominal actively ‘takes place’ over a given time frame; a key property of dynamism (Murphy 2010, p.204). Alternatively, if a nominal instance functions adequately in either construction posed by Test2 and Test5, then the situation expressed by the nominal displays stativity, as opposed to dynamism. This is because Test2 and Test5 assess whether a situation can ‘hold’ (instead of actively take place) for all moments of a given time frame, a key tenet of stative meaning (Murphy 2010, p.204). Moreover, Test4 examines if a nominal instance expresses durativity, as it signals whether the situation expressed by the nominal instance can transpire over a period of time, i.e., over a duration. This semantic behaviour deviates from punctual nominal instances, which only transpire in expeditious instantiations (Smith 1997, p.19). The structure of the six DSTs is specifically designed as a rough control mechanism to diminish the chance of error in the annotation process. For instance, if a nominal can adequately function in Test6, it will generally always entail adequate

functionality in Test1. Annotations which did not fit these expected patterns were easy to spot and were given additional attention.

Balvet et al’s. (2011) collection of DSTs alone do not test for telicity. To account for this gap, an adapted version of Dowty’s (1979, p.60) entailment DST for verbs was employed in Test7 and Test8. Following the work of Barque et al. (2009, p.171), Dowty’s (1979, p.60) entailment DST was adapted for the analysis of nominal telicity through the addition of the verbal phrase ‘was interrupted’. By including ‘was interrupted’, we can examine whether nominal instances express an inherent endpoint, or whether they can continue indefinitely. For instance, (22a) below does not entail that ‘the house was built’. (22a) expresses an inherent endpoint. Alternatively, (22b) does entail that ‘they demonstrated’. Accordingly, (22b) does not express an inherent endpoint. Lastly, object meaning was observable through the non-compliance of nominal instances with all the DSTs. If a nominal instance did not function adequately in any of the DSTs, then the nominal instance displayed no indication of temporal semantic structure.

- (22) a. The building of the house was interrupted.
- b. Their demonstration was interrupted.

*Table 3-2. DSTs for lexical aspect annotation and their literature source.*

Syntactic test	Literature source
Test1: X took place:	Balvet et al. (2011)
Test2: Felt X / experience X:	Balvet et al. (2011)
Test3: X lasted time:	Balvet et al. (2011)
Test4: Perform / Carry out X:	Balvet et al. (2011)
Test5: In a state of X:	Balvet et al. (2011)
Test6: X unfolded:	Balvet et al. (2011)
Test7: X is v-ing entails X has v-ed + was interrupted’:	Dowty (1979); Barque et al. (2009)
Test8: X is v-ing entails X has v-ed + ‘was interrupted’:	Dowty (1979); Barque et al. (2009)

### **3.2.3. Count/Mass and Concrete/Abstract Status Annotation**

In addition to the lexical aspect annotation, the nominal instances were annotated for three other variables: (1) count/mass status; (2) abstract/concrete status; and (3) word formation type (see Section 3.4). In this section, we will focus on count/mass status and abstract/concrete

status. To start, I will briefly explore these terms, before then laying out the DSTs used to analyse each variable. I will begin with count/mass status.

Count/mass status relates to a syntactic property of nominals that determines whether a nominal can be pluralised, occur with an indefinite determiner, allow numeral modifiers and occur with the quantifiers (Kiss et al. 2017, p.189). To go over a previous example from Section 2.1.2, we can see that while *house* is a count noun, as it can be pluralised (e.g., two houses), *salt* is a mass noun as, generally speaking, it is not pluralised (e.g., \*two salts). The decision to include a count/mass variable in the nominal annotation scheme was made because of the supposed relationship between the count/mass status and telicity (Mourelatos 1978; Brinton 1995; Bauer et al. 2013).

Bauer et al. (2013) suggest that nominals which derive from verbs expressing either an Accomplishment, Achievement or a Semelfactive will more likely be count nominals, while Mourelatos (1978) and Brinton claim (1995) that nominals which derive from verbs expressing either an Activity or State will more likely be mass nominals. This analogy reflects certain nominalisation patterns. As Accomplishments, Achievements and Semelfactives all express endpoints within their structure, it is argued that, when nominalised, these endpoints are typically reflected through the precise limits imposed by count status, i.e., countability. Alternatively, as Activities and States do not express endpoints in their structure, it is suggested that their nominalised forms will typically not express precise limits, and subsequently express mass status. Within cognitive grammar, this idea of nominals expressing limits and boundaries concerns the notion of ‘bounding’. Bounding, broadly speaking, relates to whether a thing or process is expressed as discrete from other entities of the same type through the existence of a boundary (Bennett 2014, p.38). Unfortunately, we do not have space here to explore the notion of bounding in greater detail. However, we will revisit it in Section 5.2, where I will discuss its relationship with the concepts of homogeneity (see also Section 5.2) and telicity in the nominal domain. Moreover, Langacker (1987b; 1991; 2008a) provides a series of comprehensive overviews on bounding for those interested.

Abstract/concrete status, on the other hand, concerns the degree to which a lexical item expresses a perceptible entity (Brysbaert et al. 2014). Nominals such as *table* and *ocean* are concrete due to their material existence. Nominals such as *love* and *desire* are abstract, as they are imperceptible through the five human senses. The decision to include this distinction within



the annotation scheme was based on the assumption that abstract nominals will be more likely to express event and stative meaning than concrete nominals. Conceptually, according to Langacker (1987b, p.90) and Radden and Dirven (2007, p.84), abstract nominals are to be treated as reifications (see section 2.2.3 for a discussion on reification in relation to concrete, abstract nominals) that enable us to comprehend relational concepts and situations as things. An example of this reification is shown by the nominalisation *removal* in example (23) below.

- (23) Jack removed the mouldy cheese from the fridge. The removal was difficult because of the bad smell.

Recent developments on abstract/concrete status have shown that the distinction between abstractness and concreteness might not necessarily be binary, but scalar, where nominals are placed on a continuum of concreteness (Anderson et al. 2014; Bolognesi et al. 2020). In some cases, these continuums are established in relation to each human sense (Brysbaert et al. 2014). However, while it is acknowledged that scalar approaches to abstract/concrete status may be the route for research going forward, a binary approach was adopted within the current study, so that the variable could efficiently operate in the Classification and Regression Tree (CRT) algorithm that was used in the statistical analysis. As I will discuss later in Section 3.6, CRT analysis produces an algorithm that presents statistical data on the degree to which predictor variables are significantly present within the data (Deshors and Götz 2020, p.13). In the following paragraphs, I will now detail the DSTs used to analyse both the count/mass status and the abstract/concrete status of the 5000 nominal instances.

Three DSTs were applied in the annotation scheme to accurately analyse 5000 nominal instances for both count/mass status and abstract/concrete status. These DSTs are presented in Table 3-3 below. Test9 and Test10 were again taken from Balvet et al. (2011), as they provide a reliable test for count/mass status. Test9 functions to identify count nominals, as only count nominals can be pluralised and satisfactorily function in the ‘several X’ grammatical construction. Test10 functions in the opposite manner. While mass nominals can function adequately with ‘measure’ terms such as ‘a bit’ and ‘some’, count nominals do not express this capacity. Lastly, Test11 was chosen to uncover abstract/concrete status of the nominal instances. Originally, Test11 was taken from Balvet et al. (2011), where it was employed to capture instances of object meaning. However, this DST is not an accurate test for object meaning, as certain nominals may express object meaning abstractly, meaning they cannot be

‘located’, e.g., *reason*. Nevertheless, when partnered with the definition of abstract/concrete status given by Brysbaert et al. (2014), Test11 functions adequately to identify abstract/concrete status in nominals. If a nominal could ‘be perceived’, by any of the five human senses, it was categorised as concrete. If it could not, it was categorised as abstract.

*Table 3-3. DSTs for count/mass status and abstract/concrete Status and their literature source.*

<b>Syntactic test</b>	<b>Literature source</b>
Test9: Several X:	Balvet et al. (2011)
Test10: A bit/some X:	Balvet et al. (2011)
Test11: Can X be perceived (by any human sense)?:	Balvet et al. (2011); Brysbaert et al. (2014)

### **3.3. Reductions and Repetition in the Sample**

In this section, we will explore the different types of instances discounted from the sample and the reasons why they were discounted. After all the reductions have been established, the section will then focus on the remaining nominal instances, looking specifically at how repetitive instances were handled in the sample.

#### **3.3.1. Reductions to the Sample**

During the SFL and lexical aspect annotation, there were many instances within the sample that were incompatible with the lexical aspect annotation. Overall, 2300 incompatible instances were identified. The breakdown of these types of incompatible instances is presented in Table 3-4 below. The reasons why these instances were not included in the final sample size are numerous and are explained in the following paragraphs.

Table 3-4. Breakdown of instances incompatible with lexical aspect annotation.

Incompatible Type	Number of Instances
Classifiers and quantifiers	392
Context unclear	64
Initialisms	103
Genitive Phrases and Possessives	51
Multiword expressions	76
Numbers	32
POS errors	134
Proper Nouns	1398
Unrecognisable words	50

Firstly, many nominal instances functioned as ‘classifiers’ in the nominal group (n=347), meaning that their primary function was to classify the head noun in the nominal group. For instance, the nominal *train* in the nominal group *the train station* functions to classify the station as one that is used for trains, and not busses etc. These nominal instances were discounted, as nominals that did not function as the syntactic head of the nominal group were incompatible with the DSTs used in this study (Balvet et al. 2011, p.1). This position was also adopted to discount ‘quantifiers’ (the element in the nominal group that indicates quantity (Fontaine 2013, p.53)) (n=45), that function in the determiner role of the nominal group, not as the syntactic head, e.g., the *five* cats.

In addition to classifiers and quantifiers, there were also several nominal instances whose surrounding linguistic context was unclear or insufficient (n=64). To accurately annotate nominal instances for lexical aspect, it is integral that the experiential information of each nominal instance is obtained. By obtaining this information, it enables an understanding of how the nominal instance functions within its linguistic context. Nominals were discounted if this experiential information could not be reliably obtained in the SFL analysis.

Multiword expressions (n=76) and initialisms (n=103) were also present within the sample. Hanks (2013, p.61) claims that there is no way to precisely predict the meaning of multiword expressions through the analysis of their component parts. This lack of semantic precision in multiword expressions and initialisms resulted in a certain amount of ambiguity when applying the DSTs to these instances. To avoid this ambiguity, instances that were made up of multiple

nominals that expressed a meaning different to the meaning of their component parts were discounted from the sample. For instance, *cream* in *ice cream*.

By far the largest incompatible type recorded was proper nouns (n=1398). Proper nouns, unlike lexical nouns, principally function as references to individual names of people, places institutions etc. (Bloor and Bloor 2013, p.194). Their meaning is generally derived from reference to something known, e.g., a person, and as a result, proper nouns do not express any lexical content of their own (Fontaine 2013, p.28; c.f. Hanks 2013). Proper nouns were therefore discounted from the sample, as they could not be satisfactorily annotated for lexical aspect using the DSTs. As well as proper nouns, genitive phrases and possessives (n=51) were also discounted from the sample, due to their structural position in the nominal group. Genitive phrases in particular are a determiner element within the nominal group which contains both a possessor in the nominal form, e.g., *table*, and a genitive marker, e.g., -'s (Fontaine 2013, p.53). Together, they function as possessive determiners, not as head nouns within the nominal group structure, as shown in example (24) below. Furthermore, instances not recognised as lexical items were discounted from the sample. With these instances, it was impossible to accurately determine their semantic meaning as no real-world information could be retrieved about the instance. Because of this ambiguity, their lexical aspect annotation would have been unreliable.

(24) The table's legs are strong.

Lastly, it was apparent when annotating the sample that the extraction process using POS information had been slightly inaccurate, as several lexical items that were not nominals, e.g., verbs, adjectives, prepositions etc. (n=134), and also numbers (n=32), were located in the sample. These instances were discounted, as the current study is centrally focused on lexical nominals alone.

After the reductions were made, 2700 nominal instances were left in the sample. This figure was comprised of 1662 nominal instances which occurred only once and 1038 repetition nominal instances (instances that occurred more than once). We can discern from these two figures that repetitive instances clearly constitute a considerable proportion of the remaining sample. The distinction between singular and repetition instances will now be discussed below (although, the specific reason for the inclusion of repetitions is later addressed in Section 3.5).

### 3.3.2. Repetition in the Sample

As stated in the previous section, repetition instances were prevalent in the sample, accounting for approximately 38% of the remaining instances. Fundamentally, repetition instances are reoccurrences of already occurred nominal instances. Looking back at our discussion of ‘meaning potential’ in Section 2.3.2 however, we know that the nature of repetition goes far beyond the reoccurrence of the same orthographical unit alone. For instance, we know that examination in *the examination on the table* is not a repetition of examination in *the examination took three hours* (Grimm and McNally 2013). They are semantically discrete. With this in mind, the structure of the nominal annotation scheme (outlined in Section 3.2) was created to differentiate singular instances from repetition instances.

Initially, general ‘encyclopaedic’ knowledge was attained on each nominal instance through one’s real-world knowledge. This provided information on the nominal instances’ ontological profile. Then, by undertaking the SFL experiential analysis on the nominal instances (discussed in Section 3.2.1), information was obtained in relation to the nominal instances’ grammatical profile, through examining how each nominal instance grammatically functioned within its context of use. Lastly, by placing each nominal instance into a series of DSTs, information was attained on the nominal instances’ temporal semantic profile. The information expressed at each of these three levels (ontological, grammatical, and temporal semantic) was used to establish whether the nominal instance was a repetition or not. If an instance was ontologically, grammatically, and temporal-semantically the same as an instance already recorded, then it was labelled as a repetition. If the instance displayed a differentiation on either of these levels from all previous nominal instances, then it was not deemed as an instance of repetition.

Now that the working definition of repetition instances has been established, we will move on to the last annotation of the nominal instances. However, we will revisit the presence of repetition in the sample again in Section 3.5, where we will discuss the motivation behind leaving repetition instances in the sample for the analysis. In the meantime, I will now move our focus to the annotation of the nominal instances for their word formation properties. This annotation involved researching the nominal instances’ etymology information, contained within the Oxford English Dictionary (OED) (2022). Within this section, I first outline the distinction between inflectional and derivational morphology, before then moving on to detail this word formation annotation. In doing so, I will describe the word formation categories used

to classify the nominals. I will then briefly touch on how hybridity was handled within the data in relation to these word formation categories.

### **3.4. Word Formation Annotation**

After the incompatible instances were removed from the sample, the remaining nominal instances were subjected to a word formation annotation, using the OED (2022). This analysis quite simply involved taking each nominal instance and researching its etymology in the OED. After the etymology of the nominal instance was noted, the nominal instance was then assigned to one (or potentially more than one) of eight categories based on its word formation properties. The categories used within the current study to classify these word formation properties were: (1) Borrowing; (2) Compound; (3) Morphologically Derived from Adjective (MDA); (4) Morphologically Derived from Noun; (5) Morphologically Derived from Verb; (6) Other; (7) Transcategorisation (TC); and (8) Underived. We will explore each of these categories in greater detail in Section 3.4.2. First however, we will briefly look at the terms ‘inflection’ and ‘derivation’, and examine the morphological status of *-ing* suffixation in nominals. By understanding the general distinctions between inflectional morphology and derivational morphology, we will be better equipped to engage with the eight word formation categories later in this section.

#### **3.4.1. Inflection vs Derivation**

Inflection and derivation are concepts in the field of morphology - a subdiscipline of linguistics which focuses on the internal structures of lexical items (Booij 2006, p.654). Inflectional morphology principally concerns the relationship between the different grammatical forms of an individual lexeme (Aronoff and Fudeman 2011, p.47). In English, this includes the presence of grammatical forms such as past, present, future, singular, plural, etc. For example, while we have the lexical item *jump*, through the addition of inflectional morphology, we can also have *jumps*, *jumped* and *jumping*. Derivational morphology, on the other hand, concerns the relationship between different lexemes (Carstairs-McCarthy 2002, p.45). Derivational morphemes function to produce different lexemes altogether. For instance, the nominal *teacher* is derived from the verbal base *teach*, through the addition of the derivational morpheme *-er*. In fact, the derivational morpheme *-er* is one of the most generally used morphemes for forming nouns out of verbs (Carstairs-McCarthy 2002, p.51). Fundamentally then, the distinction between inflection and derivation is a functional one within the morphological process (Booij

2006, p.654) (this position is disputed by Haspelmath (1996), see below). Many other distinctions between inflectional and derivational morphology have been proposed, such as ‘obligatoriness’, ‘syntactic relevance’, ‘productivity’, ‘paradigms’, etc. Unfortunately, I do not have the space to cover each of these distinctions right now. Detailed discussions about these distinctions can be found in Booij (2006) and Lieber (2010).

The morphological status of specific morphemes can regularly prove ambiguous, however. This is certainly the case for the morphological status of *-ing* suffixation, which has proven a highly contested topic of discussion. Within current literature in the field of morphology, there is debate as to whether *-ing* suffixation constitutes an inflectional morpheme or a derivational morpheme. Generally, *-ing* is treated as an inflectional morpheme in the verbal domain (Bauer et al. 2013, p.537). Nevertheless, in the case of nominals, Bauer et al. (2013, p.537) acknowledge that the morphological status of *-ing* suffixation is ambiguous. Scholars such as Carstairs-McCarthy (2002), Blevins (2006) and Fabregas and Scalise (2012) appear to suggest that in the case of nominalisation, *-ing* can be considered a derivational morpheme, as it expresses the capacity to alter a lexeme’s lexical class. For instance, the nominal *building* is formed through the combination of the verb form *build* and *-ing* suffixation (OED 2020c). Furthermore, scholars have argued that the semantic meaning expressed by *-ing* suffixation in gerundive nominals can be compared to that of other derivational suffixes (Quirk et al. 1985; Brinton 1991; Brinton 1995; Brinton 1998; Alexiadou et al. 2013).

However, Haspelmath (1996) argues that *-ing* should not be viewed as derivational, but as a ‘category-changing inflection’. Haspelmath (1996, p.47) proposes that the inflectional/derivational morphology distinction is not binary, and should instead be viewed on a continuum between clear inflection and clear derivation. To state once again, this debate on the morphological status of *-ing* suffixation is highly contested. Bauer et al. (2013, p.55) concede that the treatment of these forms has always been an intricate matter, and suggest that each interpretation of the morphological status of *-ing* suffixation should fall in line with the theoretical framework in play. Accordingly, as this thesis is primarily focused on the semantics expressed by nominals, *-ing* suffixation will be interpreted as a derivational morpheme in the case of nominalisation, due to its supposed relationship with derivation-based semantics and its lexical class changing capabilities. In the next section below, I will now detail the eight different word formation categories used in this thesis.

### 3.4.2. Word formation Categories

As stated at the beginning of Section 3.4, the word formation annotation involved classifying each of the remaining nominal instances into eight word formation categories, based on their etymology information within the OED. These eight distinct word formation categories are: (1) Borrowing; (2) Compound; (3) Morphologically Derived from Adjective (MDA); (4) Morphologically Derived from Noun; (5) Morphologically Derived from Verb; (6) Other; (7) Transcategorisation (TC); and (8) Underived. The eight categories represented the most prevalent word formation types within the sample. Definitions of each of the eight word formation categories are provided below.

- 1. Borrowing:** Nominal instances assigned to the Borrowing category constitute nominals that have been ‘borrowed’ into the English language from a separate language (Yule 2014:52). For instance, the nominal *pocket* originally entered the English language from French (OED 2020h).
- 2. Compound:** Nominal instances assigned to the Compound category constitute nominals that have been formed through the combination of two or more lexical roots (Carstairs-McCarthy 2002, p.59). For instance, the nominal *aircraft* was formed through the combination of the two distinct nominals *air* and *craft* (OED 2020b).
- 3. Morphologically Derived from Adjective (MDA):** Nominal instances assigned to the MDA category constitute nominals which are derived from an adjectival base through derivational morphology (Roy 2010, p.129). For instance, the nominal *happiness* derives from the adjective *happy*. Through the addition of the derivational morpheme *-ness*, its lexical class is altered to a noun (OED 2020e)
- 4. Morphologically Derived from Noun:** Nominal instances assigned to the MDN category constitute nominals which combine with derivational morphology to yield a different nominal (Carstairs-McCarthy 2002, p.49). For instance, the nominal *journalist* is produced through the combination of the nominal *journal* and the derivational suffix *-ist* (OED 2020g).



5. **Morphologically Derived from Verb:** Nominal instances assigned to the MDV category constitute nominals that were originally verbs, but have been reclassified as nominals through the addition of derivational morphology (Heyvaert 2015, p.41). For instance, the nominal *producer* derives from the verbal form *produce*. Through the addition of the derivational morpheme *-er*, its lexical class is altered to a noun (OED 2020i).
6. **Other:** Nominal instances assigned to the Other category constitute nominals that do not neatly fit into any of the other word formation categories in this study. The word formation types of these instances are highly various, such as ‘abbreviations’, ‘alterations on lexical items’, ‘clippings’, ‘unknown origins’ etc. The individual quantities of each of these word formation types was low in relation to the overall sample. As it was uneconomic to create individual word formation categories for every type of word formation, the Other category was created to capture all of these low quantity word formation types.
7. **Transcategorisation (TC):** Nominal instances assigned to the TC category constitute nominals that have been converted, i.e., ‘transcategorised’, from a different lexical class into a noun, without any explicit morphological markers indicating an alteration in grammatical function (Davies 2004, p.2). For instance, the nominal *rush* does not display any distinguishing morphological markers from its verbal counterpart (OED 2020k). However, it grammatically functions as a nominal. The term ‘transcategorisation’ differs from the more commonly employed term ‘Conversion’. The choice to use the term transcategorisation is a motivated decision, following Ježek and Ramat (2009), as it provides a more transparent term that explicitly foregrounds the process of altering lexical categories despite the lack of derivational morphology (Ježek and Ramat 2009, p.394).
8. **Underived:** Nominal instances assigned to the Underived category constitute what Huyghe et al. (2017) call ‘morphologically simple’ nominals. These are nominals that do not derive from any other lexeme through any kind of morphological process (Huyghe et al. 2017, p.119). They have also been present within the English language since the ‘Old English’ language period. For instance, the nominal *house* is classified as an Underived nominal as it has always functioned as a nominal in the English

language, it expresses no derivational morphology and it is inherited from Germanic (OED 2020f).

### 3.4.2.1. Hybridity in the Sample

As we might expect with any type of classification, certain nominal instances did not neatly fit into just one of the eight categories, as they expressed word formation properties of more than one word formation type. These instances were known as ‘Hybrids’. To combat the presence of hybrids in the dataset, each hybrid nominal was assigned to as many of the word formation types that their etymology expressed. To illustrate this point, we can view the OED etymology information on the nominal *advance*. According to the OED (2020a), *advance* is partly a borrowing from French and also partly a transcategorisation in English from its verbal counterpart. As *advance* expresses word formation properties of both Borrowing and Transcategorisation, it was consequently assigned to the Borrowing category and the Transcategorisation category.

After each of the nominal instances had been successfully categorised for its word formation properties, the entire annotation process was complete. To recap, this data was compiled of frequency information on the lexical aspect, count/mass status, abstract/concrete status, and word formation classification of each nominal instance, as well as the genre in which each nominal instance appeared. In the following section, I will now focus on how the annotated data was prepared for analysis. Section 3.6 will then detail the analyses undertaken, providing justification as to why each form of analysis was chosen for the current study.

## 3.5. Dataset Preparations for Analyses

An important point to address, before moving on to the different forms of statistical analysis conducted in this study, is to make clear how repetition was handled when these analyses were undertaken. Repetition instances posed a particularly tricky issue in relation to preserving the reliability of the data. From the beginning of the study, it was apparent that if repetition instances were included in the data, then the significance of relationships between certain word formation types and distributions of situation type could become inflated. For example, if an Underived nominal instance that expressed object meaning occurred 40 times within the sample, this could inaccurately inflate the relationship between Underived word formation type and Objects. On the contrary, the removal of repetitions could deny potentially interesting data

on genre information. As a result, there was a tension between the preservation of both word formation type information and genre information. With this tension in mind, the dataset was split into two different datasets: a dataset that did not include repetition instances (the central dataset), and a dataset which did include repetition instances. The same analyses were conducted on both datasets to monitor the influence of repetition on the data. Furthermore, to capture exactly how the nominal instances behaved in relation to just event and state semantics, an additional dataset was created that only contained nominal instances that expressed temporal semantics (all Object instances were removed). As the expression of object meaning has been traditionally associated with archetypal nominal semantics (Langacker 1991, p.298; Givón 2001, p.51; Halliday 2009, p.117), the removal of Objects from the dataset poses an interesting area of exploration.

Overall, three separate datasets, distinguished by repetition and Object expression were used in the current study. These three different datasets and the number of nominal instances contained within them are recorded in Table 3-5. Interestingly, when repetition instances were not included in the data, 25.99% of all instances expressed some type of temporal meaning.

*Table 3-5. The three different datasets used in the study accompanied with the number of instances they each contained.*

Dataset Name	Number of Nominal Instances
All Data without Repetition	1662
All Data with Repetition	2700
No Object Data without Repetition	432

For the analysis of this data, the Fisher-Exact test (accompanied with post-hoc Fisher-Exact tests), the Cramer's V effect size test and Classification and Regression Tree (CRT) analysis were used. In the next section, I will describe each of these analyses, their practicalities and the motivation which led to their employment in the study. I will first cover the Fisher-Exact test and Cramer's V test, and afterwards, the CRT analysis.

## **3.6. Statistical Analyses**

### **3.6.1. Fisher-Exact Tests and Cramer's V Test**

In this section, I will outline both the Fisher-Exact test and Cramer's V test and describe the motivation behind the use of each statistical test. The statistical analyses were conducted using the software 'R' (R Core Team 2018). R is a free quantitative software environment that includes a range of functions that contribute to enhancing data descriptions. Specifically, RStudio, an integrated development environment for R, was used to conduct statistical analyses and create visualisations, using programming language (RStudio Team 2020). I will begin first with the Fisher-Exact test.

The Fisher-Exact test is a non-parametric statistical test for independence that measures whether the frequency distribution between two or more categorical variables is statistically significant (Hess and Hess 2017, p.878). For instance, if two different word formation types expressed a largely different distribution of situation type, the Fisher-Exact test would indicate a significant (dependent) relationship between word formation type and situation type. The statistical significance of this relationship is computed through the comparison of two or more 'actual frequencies' (the actual frequencies recorded in the data) with their 'expected frequencies' (the frequencies that one would expect if there were no differences between the frequencies in the data (Levshina 2015, p.211)). Essentially then, the Fisher-Exact test measures the extent to which actual data frequencies correspond with their expected data frequencies and the extent to which differences in data frequencies have occurred by chance.

There were several reasons for selecting the Fisher-Exact test, of which, the first is that it provides a robust statistical test for the use of frequency data (Hess and Hess 2017, p.878). Another specific benefit is that it does not assume a normal distribution of the data; it is a non-parametric test. This non-parametric nature was advantageous in the current study, as the data expressed an unbalanced distribution. Furthermore, the Fisher-Exact test also provides a robust test for variable dependence, which effectively handles low frequency data (Levshina 2015, p.211). This advantage was a key feature that motivated the use of the Fisher-Exact test in this study, as other tests of variable dependence, such as the Chi-Squared Test for Independence, are not as reliable when working with lower frequency data (notably, data with expected frequencies of under 5) (Hinton 1995, p.242).

In the current study, the Fisher-Exact test was used to identify whether a significant relationship existed between word formation type and situation type, count/mass status and situation type, abstract/concrete status and situation type, and genre and situation type. The results of the Fisher-Exact tests each generated a p-value that indicated the level of dependence between the above stated variables. If this p-value was lower than the set alpha level of  $p=0.05^3$ , then a significant result was recorded. The variables were identified as dependent on one another. If the results from the Fisher-Exact tests indicated a significant (dependent) relationship, then further post-hoc Fisher-Exact tests were used to investigate where these significant frequency differences lied in the data. When carrying out the post-hoc Fisher-Exact tests, a ‘Bonferroni correction’ was applied to the alpha level, which adjusted the p-value in relation to how many pairwise tests were conducted in the post-hoc test. This correction level was applied in order to avoid incorrectly disregarding the null hypothesis ( $H_0$ ), as the traditional alpha level of  $p=0.05$  is an unreliable measure of significance when multiple tests are undertaken in a pairwise fashion (Sedgwick 2014).

As noted above, the Fisher-Exact test provides a statistical output on whether the relationship between variables is dependent and additionally the amount of confidence we can have that a difference in the data is reliable. If the returned statistic is under the set alpha level (e.g.,  $p=0.05$ ), then we can be confident that this difference is significant, and thus not due to chance. Despite this capacity, the Fisher-Exact test does not provide information on the strength of the assessed relationship, or the extent of the difference (Levshina 2015, p.129). For this information, an effect size test is required. Effect size fundamentally measures the strength of a relationship between variables, ranging from ‘weak’ effect size to ‘strong’ effect size:  $0.1 \leq \phi < 0.3$  is a weak effect size;  $0.3 \leq \phi < 0.5$  is a moderate effect size; and  $\phi \geq 0.5$  is a strong effect size (Sheskin 2011, p.535). It is understood that the Phi-coefficient provides the most accurate test for effect size when analysing the strength of a relationship presented in a two-by-two table (Levshina 2015, p.129). However, in the current study, the data was far larger than a two-by-two table, as the dependent variable of situation type + object meaning constitutes six distinct categories (i.e., Accomplishment; Achievement; Activity; Object; State; and Semelfactive). To accommodate this data structure, the Cramer’s V effect size test was conducted, as Cramer’s V test presents a valuable alternative to the Phi-coefficient because it

---

<sup>3</sup> Within the field of linguistics, alpha levels are typically set at 0.05 or lower. This measure allows for 95% percent confidence in the result (Rasinger 2013, p.173).

is a “generalised version of Phi” (Jenset 2008, p.12) that can accommodate matrices larger than a two-by-two tables.

The use of the Fisher-Exact test, post-hoc Fisher-Exact tests, and the Cramer’s V test, will allow this study to address the following research question (RQ):

**RQ1:** To what extent are word formation type, count/mass status, abstract/concrete status, and genre influential in the distribution of nominal semantic behaviour?

To investigate the above RQ, I propose the following hypotheses for the analysis of each of the three datasets. These hypotheses are outlined below and presented in relation to the explorative variable in question:

**Word Formation Type:**

H0 - Word formation type does not share a dependent/significant relationship with situation type.

H1 - Word formation type does share a dependent/significant relationship with situation type.

**Count/Mass Status:**

H0 - Count/mass status does not share a dependent/significant relationship with situation type.

H1 - Count/mass status does share a dependent/significant relationship with situation type.

**Abstract/Concrete Status:**

H0 - Abstract/concrete status does not share a dependent/significant relationship with situation type.

H1 - Abstract/concrete status does share a dependent/significant relationship with situation type.

**Genre:**

H0 - Genre does not share a dependent/significant relationship with situation type.

H1 - Genre does share a dependent/significant relationship with situation type.

### 3.6.2. Classification and Regression Tree Analysis

In addition to the statistical measures noted above, a Classification and Regression Tree (CRT) analysis was used to model the simultaneous influence of word formation type, count/mass status, abstract/concrete status, and genre on the semantic behaviour expressed by nominals. Essentially, the approach aims to provide a clear visualisation of how the combined effects of above-mentioned variables distinguish different distributions of semantic behaviour in the nominal domain. This analysis was again carried out on the quantitative software environment R (R Core Team 2018), using the *cree* function in the *Party* R package (Hothorn et al. 2006). At its most basic description, CRT analysis, otherwise known as ‘conditional inference tree’ analysis (Levshina 2015) or ‘binary decision tree’ analysis (Roberts et al. 2015), uses predictor variables to separate a dependent variable (in this case, situation type expression) into sub-sets (Roberts et al. 2015, p.7). Essentially, CRT analysis identifies variables that give rise to outcomes or predictions (Glynn 2014). In doing so, CRT analysis produces an algorithm that presents statistical data on the degree to which these variables are significantly present within the data (Deshors and Götz 2020, p.13). This statistical data is displayed in a hierarchical tree-like structure. The given significance level of these variables indicates the variable’s influence on the distribution of the dependent variable. Once again, the alpha level was set at  $p=0.05$ , meaning that any variable association lower than  $p=0.05$  was considered significant.

When a significant variable is identified, the data is split in a binary fashion. The first variable to be split in the tree structure represents the central node, i.e., the most influential variable in the tree structure. After the central node is established, the splitting process is repeated with other variables until there are no longer any variables associated with an outcome at a significant level (Levshina 2015, p.291). The result of this splitting process is the formation of various ‘branches’ (i.e., paths), that each result in a ‘leaf’, otherwise known as a ‘terminal node’, i.e., an outcome. A visualisation of the decision tree structure created in this study is present in Section 4.2. During this section, information will be provided to help the reader interpret the information displayed in the decision tree.

CRT analysis poses many advantages that are specifically beneficial to this study. Like the Fisher-Exact test discussed in the previous section, CRT analysis is also non-parametric, meaning it can be used with data that exhibits an abnormal distribution. In addition to this quality, CRT analysis is considered robust in the presence of outliers and also highly efficient

in its handling of categorical variables (Levshina 2015, p.292). These qualities are particularly desirable for the current study, as CRT analysis can attain highly reliable results on associations between multiple categorical variables. Moreover, CRT analysis, when used in relation to nominal semantics, can help uncover how different types of meanings are generated. For instance, CRT analysis can uncover non-linear relationships and interactions between variables that influence nominal semantic behaviour (Tagliamonte and Baayen 2012). CRT analysis also provides information on the amount of influence a variable expresses in relation to predicting an overall outcome (Deshors and Götz 2020, p.13). In taking the influence of a variable into account, we can begin to understand what nominal features may be ‘weighted’ differently with regards to supposed prototypical nominal semantics. Furthermore, the use of CRT analysis provides a relatively novel approach to analysing nominal semantics in a large-scale study. Currently, it appears that only a very recent study by Lieber and Plag (2022) provides a similar approach, using CRT analysis to explore how interactions between count/mass status and lexical aspect features can influence morphological form.

However, I must concede that the structure of CRT analysis is sensitive to the selection of predictor variables and also the specific sample of data (Strobl et al. 2009; Tagliamonte and Baayen 2012; Roberts et al. 2015). For instance, the selection of the central (most important) node could only be based on a slight trend in the data, that could then affect the subsequent branches in the tree structure. To combat this flaw, following Levshina (2015, p.294), the CRT analysis was run several times within R, using various different number seeds<sup>4</sup>, to assess whether there were any notable differences within the structure of the tree. Furthermore, the reliability of the CRT analysis was also assessed against the results of the already undertaken Fisher-Exact tests, that also examined the dependency of many of the categorical variable associations present within the CRT analysis. Accordingly, through the application of the CRT analysis, this study will address the following two research questions:

**RQ2:** How do the predictor variables differ in relation to their influence on the distribution of semantic behaviour expressed by nominals?

---

<sup>4</sup> The use of different number seeds refers to the *set.seed()* function in R. This function allows an analyst to obtain a reproducible random result when they specify the same number that was originally put into the *set.seed()* function (Levshina 2015, p.294), e.g., *set.seed(115)*. By changing this number, and running the analysis several times, it allowed for the inspection of many random results, to check the reliability of the results overall.



**RQ3:** To what extent do predictor variable interactions influence the distribution of semantic behaviour by nominals?

In the next chapter, the results of the Fisher-Exact tests, Cramer's V tests and the CRT analysis are presented in relation to the three above-stated RQs. Section 4.1 will examine the individual influence of word formation type, count/mass status, abstract/concrete status, and genre on situation type, while Section 4.2 will examine the simultaneous influence of the variables on situation type. Afterwards, Chapter 5 will offer a discussion of the results noted in Chapter 4. This discussion will focus specifically on each of the variables that registered significant data splits in the CRT analysis and how a greater acknowledgement of their hierarchical and interactive behaviour can provide insight into our current understanding of the nature and degree of nominality.

## **Chapter 4: Investigating Category Feature Influence on Nominal Semantics - Results**

Following from the previous chapter, this chapter will present the results of the analyses undertaken in the first study of the thesis. The results of the analyses are given in relation to their relevant RQs. First, the chapter will address RQ1 (see below), detailing the results from Fisher-Exact tests, post-hoc Fisher-Exact tests and Cramer's V tests. In presenting these results, the chapter will display if and where significant differences in situation type distribution lie between different word formation types, count/mass status, abstract/concrete status, and genre. Moreover, the chapter will consider the influence of repetition and the influence of Object expression on these results through the additional analysis of the two other datasets: (1) All Data with Repetition dataset; and (2) No Object Data without Repetition dataset. The presentation of the results for RQ1 will be divided with regards to the variable in question. For instance, first we will explore the relationship between word formation type and situation type. Then we will progress to the relationship between count/mass status and situation type etc. After the results from the Cramer's V test and Fisher-Exact tests have been established, the chapter will move to the results of the Classification and Regression Tree (CRT) analysis to address RQs 2 and 3 (see Section 4.2). Unlike the Fisher-Exact test, the CRT analysis (see Section 3.6.2) will provide information on the simultaneous influence of the above variables on the distribution of situation type in nominals. Once again, the influence of repetition and the influence of Object expression on these results will be additionally considered.

### **4.1. Individual Influences on Nominal Semantic Behaviour**

Below we will now explore the results of the Cramer's V tests, Fisher-Exact tests and post-hoc Fisher-Exact tests on each of the following variables' association with situation type: word formation type; count/mass status; abstract/concrete status; and genre. In doing so, I will address the following research question (previously stated in section: 3.6.1):

RQ1: To what extent are word formation type, count/mass status, abstract/concrete status and genre influential in the distribution of semantic behaviour in nominals?

#### **4.1.1. The Relationship Between Word Formation Type and Situation Type**

Overall, the nominal instances expressed five distinct situation types and object meaning: Accomplishments (ACC); Achievements (ACH); Activities (ACT); Objects (OBJ); Semelfactives (SEM); and States (STA). A visualisation of the distribution of these situations and object meaning in relation to the different word formation types (stated in Section 3.4.2) is present in Figure 4-1 below. Each visualisation shown in Section 4.1 was produced using the *ggplot()* function in the *ggplot2* R package (Wickham 2016). In the visualisation below, every bar represents a meaning distribution expressed by a word formation type. From this overview, we see that events, states and objects were expressed by the different word formation types. Rather expectedly, Object was the predominant meaning expressed. Specifically, 74.9% of all Borrowings were Objects, 90.5% of Compounds, 48.8% of MDAs, 76.9% of MDNs, 59.1% of MDVs, 81.3% of Others, 40.9% of TCs and 84.4% of Underiveds. Rather interestingly, this predominant orientation to Objects was not expressed by MDAs, which expressed an equally large number of States (48.8%). Most notably, MDV and TC exhibited the greatest variation in situation type, expressing Accomplishments at a frequency of 16.1% and 18.3%, Achievements 3.2% and 10.8%, Activities 14% and 25.8%, Objects 59.1% and 40.9%, Semelfactives 0% and 1.1% and States 7.5% and 3.2% respectively. For the interested reader, the complete representation of these percentages is available in Table AB-1, Appendix B.

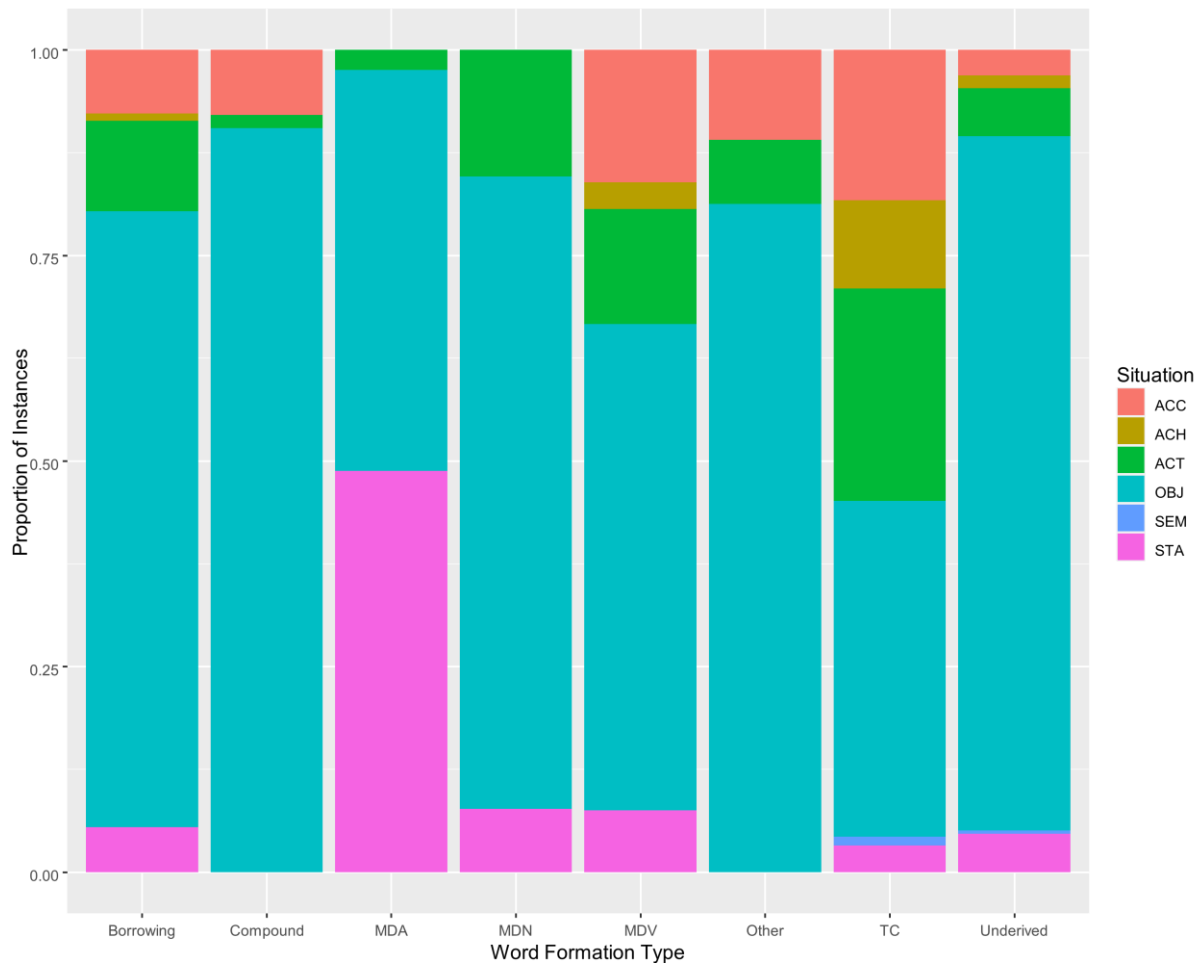


Figure 4-1. Barplot of situation type proportion expressed by word formation type.

As stated in the previous chapter, a Cramer’s V test was used to measure the strength of the relationship between the two variables. In addition, a Fisher-Exact test was calculated to compare the frequencies of Accomplishments; Achievements; Activities; Objects; Semelfactives; and States over the word formation types and to assess whether word formation type shares a dependent relationship with situation type expression. Overall, a weak effect size was returned (0.189) and a significant dependence was found  $p < 0.001$ , indicating that while the relationship between word formation type and situation type is not very strong, the situation type distributions expressed by the word formation types were significantly different. As a result, the null hypothesis was rejected.

To assess specifically where the significant differences between situation type distributions lie, post-hoc Fisher-Exact tests were conducted on each word formation type comparison, e.g., Borrowing vs Compound. The results are detailed below in Table 4-1. Due to the implementation of the Bonferroni correction to the alpha level, the result of each post-hoc

Fisher-Exact test must register a value under  $p=0.0018$  to indicate a significant difference in situation type distribution. In Table 4-1, 16 out of 28 word formation type comparisons returned a significant difference<sup>5</sup>. Such a high proportion of significant comparisons further suggests the significant influence word formation type holds over situation type distribution. These differences were generally apparent when the word formation types MDA and TC were considered. In particular, the most different situation type distribution was expressed by MDA, which was involved in seven significant situation type distribution comparisons. MDA expressed a situation type distribution that was significantly different to every other word formation type. The second most variable situation type distribution was expressed by TC, which was involved in six significant distribution comparison differences. The only situation type distribution comparable to that expressed by TCs was the situation type distribution expressed by MDVs. MDVs returned three significant situation type distribution comparisons. Because MDVs are morphologically derived from verbs, we might have expected the situation type distribution of MDV to be more distinct. However, MDVs did express a considerable number of Objects, mostly driven by agent creating morphology, such as *-er* in the nominal *hunter*. This semantic behaviour will be discussed in greater detail in Section 5.3.2.1. An example of an MDA (25a), TC (25b) and MDV (25c) is shown below.

- (25) a. Like many people branded with the attributes of reliability, honesty, and politeness, the band were attracted to the George Best myth... (Noun 746, All Data without Repetition Dataset, Appendix A)
- b. Now all this hinted, I'm sure you'll agree, that the split had been brewing, indeed planned for some time. Both Marr and Morrissey were instantly involved... (Noun 1919, All Data without Repetition Dataset, Appendix A)
- c. In another sense, it's the consumer or the taxpayer. (Noun 1094, All Data without Repetition Dataset, Appendix A)

---

<sup>5</sup> Significant results are indicated by the presence of an asterisk and bold font.

Table 4-1. Post-hoc Fisher-Exact test results on comparisons of semantic behaviour expressed by different word formation types.

Number	Comparison	p.value
1	MDV-TC	p=0.019
2	MDV-Borrowing	p=0.0035
3	MDV-MDA	<b>*p&lt;0.001</b>
4	MDV-MDN	p=0.0355
5	MDV-Compound	<b>*p&lt;0.001</b>
6	MDV-Underived	<b>*p&lt;0.001</b>
7	MDV-Other	p=0.014
8	TC-Borrowing	<b>*p&lt;0.001</b>
9	TC-MDA	<b>*p&lt;0.001</b>
10	TC-MDN	<b>*p&lt;0.001</b>
11	TC-Compound	<b>*p&lt;0.001</b>
12	TC-Underived	<b>*p&lt;0.001</b>
13	TC-Other	<b>*p&lt;0.001</b>
14	Borrowing-MDA	<b>*p&lt;0.001</b>
15	Borrowing-MDN	p=0.265
16	Borrowing-Compound	p=0.01
17	Borrowing-Underived	<b>*p=0.001</b>
18	Borrowing-Other	p=0.194
19	MDA-MDN	<b>*p&lt;0.001</b>
20	MDA-Compound	<b>*p&lt;0.001</b>
21	MDA-Underived	<b>*p&lt;0.001</b>
22	MDA-Other	<b>*p&lt;0.001</b>
23	MDN-Compound	<b>*p=0.001</b>
24	MDN-Underived	p=0.225
25	MDN-Other	p=0.0095
26	Compound-Underived	p=0.16
27	Compound-Other	p=0.229
28	Underived-Other	p=0.0435

#### 4.1.1.1. The Influence of Repetition on Situation Type Distribution across Word Formation Types

To remind the reader, this thesis views an instance of repetition as an instance that shares the same ontological, grammatical, and temporal semantic profile as another instance already recorded in the dataset (for a full description, see section 3.3.2). A separate Cramer's V test, Fisher-Exact test and post-hoc Fisher-Exact tests were undertaken to consider the influence of

repetition on situation type distribution across word formation type, count/mass status, abstract/concrete status, and genre. While the main results for each variable will be presented throughout the chapter, the results are also presented in full in Appendix B. In this section, I will focus particularly on the influence of repetition on situation type distribution across word formation type. Overall, the addition of repetition instances in the dataset did not greatly affect either the effect size (0.161) or the significance of the relationship between word formation type and situation type ( $p < 0.001$ ). From the post-hoc Fisher-Exact tests, it is apparent that the presence of repetition in the dataset has a small influence on significance levels. Specifically, two word formation type comparisons (MDV-Borrowing; MDV-Other) indicate new significant relationships from the addition of repetition instances. Accordingly, with the addition of repetition instances, the situation type distribution expressed by MDV becomes more significantly different, recording five significantly different situation type distribution comparisons, compared to just three when repetition instances were not included.

#### **4.1.1.2. The Influence of Object expression on Situation Type Distribution across Word Formation Types**

In addition to the analyses above, a Cramer's V test, Fisher-Exact test and post-hoc Fisher-Exact tests were carried out on a dataset that also took into consideration the influence of Object expression on situation type distribution across word formation type, count/mass status, abstract/concrete status, and genre. In a similar vein to the discussion on the influence of repetition above, the main results for each variable will be presented throughout the chapter, while the complete set of results is detailed in Appendix B. To assess the influence of Object expression, every instance that expressed Object was removed, leaving only nominal instances that expressed some sort of temporality. As the expression of object meaning has been traditionally associated with archetypal nominal semantics (Langacker 1991, p.298; Givón 2001, p.51; Halliday 2009, p.117), the removal of Objects from the dataset poses an interesting area of exploration. By removing Objects, it allows us to additionally cover how different types of nominals vary based on their temporal semantics alone, i.e., their expression of event and state meanings.

Overall, despite the removal of Object instances from the dataset, the effect size of the relationship between word formation type and situation type remained weak (0.261) and the relationship itself remained statistically significant  $p < 0.001$ . Such a result implies that Object expression holds little influence over the relationship between word formation type and

situation type. However, from the individual word formation type comparisons, carried out in the post-hoc Fisher-Exact tests, Object expression does appear to be influential. Specifically, when Objects are removed from the dataset, only eight significant word formation type comparisons (MDV-MDA; TC-Borrowing; TC-MDA; Borrowing-MDA; MDA-MDN; MDA-Compound; MDA-Underived; MDA-Other) remain. TCs record notably less significant word formation type comparisons when Objects are removed, recording only two significant comparisons. This large reduction in the number of significant comparisons is especially intriguing, as it suggests that the distribution of Objects is highly influential in creating significant differences in semantic distributions across word formation types. MDAs, on the other hand, return the highest number of significant comparisons (7), despite the removal of Objects. The high number of significant differences suggest that MDAs express the most significantly different situation type distribution of the word formation types irrespective of the presence of Objects in the dataset.

#### **4.1.2. The Relationship Between Count/Mass Status and Situation Type**

In addition to the relationship between word formation type and situation type, the relationship between count/mass status and situation type was also assessed. Count/mass status, as discussed in Section 3.2.3, is a syntactic property of nominals that determines whether a nominal can be pluralised, occur with an indefinite determiner, allow numeral modifiers and occur with the quantifiers (Kiss et al. 2017, p.189). An overview of the distribution of situation type in relation to count/mass status is shown in Figure 4-2 below. Interestingly, mass nominal instances displayed a notably weaker orientation to Object expression. When we explore the frequency at which events, states and objects were expressed, we see that while count nominal instances expressed Accomplishments at a frequency of 8.6%, Achievements 2.1%, Activities 8.5%, Objects 80.1%, Semelfactives 0.1% and States 0.7%, mass nominal instances expressed Accomplishments at a frequency of 5.8%, Achievements 0%, Activities 16.5%, Objects 52.5%, Semelfactives 0.3% and States 25%. From these percentages, we see that although both count and mass nominals predominantly expressed Objects, mass nominals also displayed a strong tendency to express States. Below, three examples are presented. Example (26a) is a count Object nominal, while example (26b) is a mass Object nominal. Example (26c), on the other hand, is a mass State nominal.



- (26) a. Acanthodian fishes generally resembling this one are known from quite complete specimens from Devonian to Permian rocks. (Noun 1158, All Data without Repetition Dataset, Appendix A).
- b. The pollen stratigraphy indicates that Salix-dominated vegetation, with a wide variety of herbs was prominent until about 6000 B.P. (Noun 1196, All Data without Repetition Dataset, Appendix A)
- c. ... I have since acquired an insatiable thirst and desire for champagne. (Noun 55, All Data without Repetition Dataset, Appendix A).

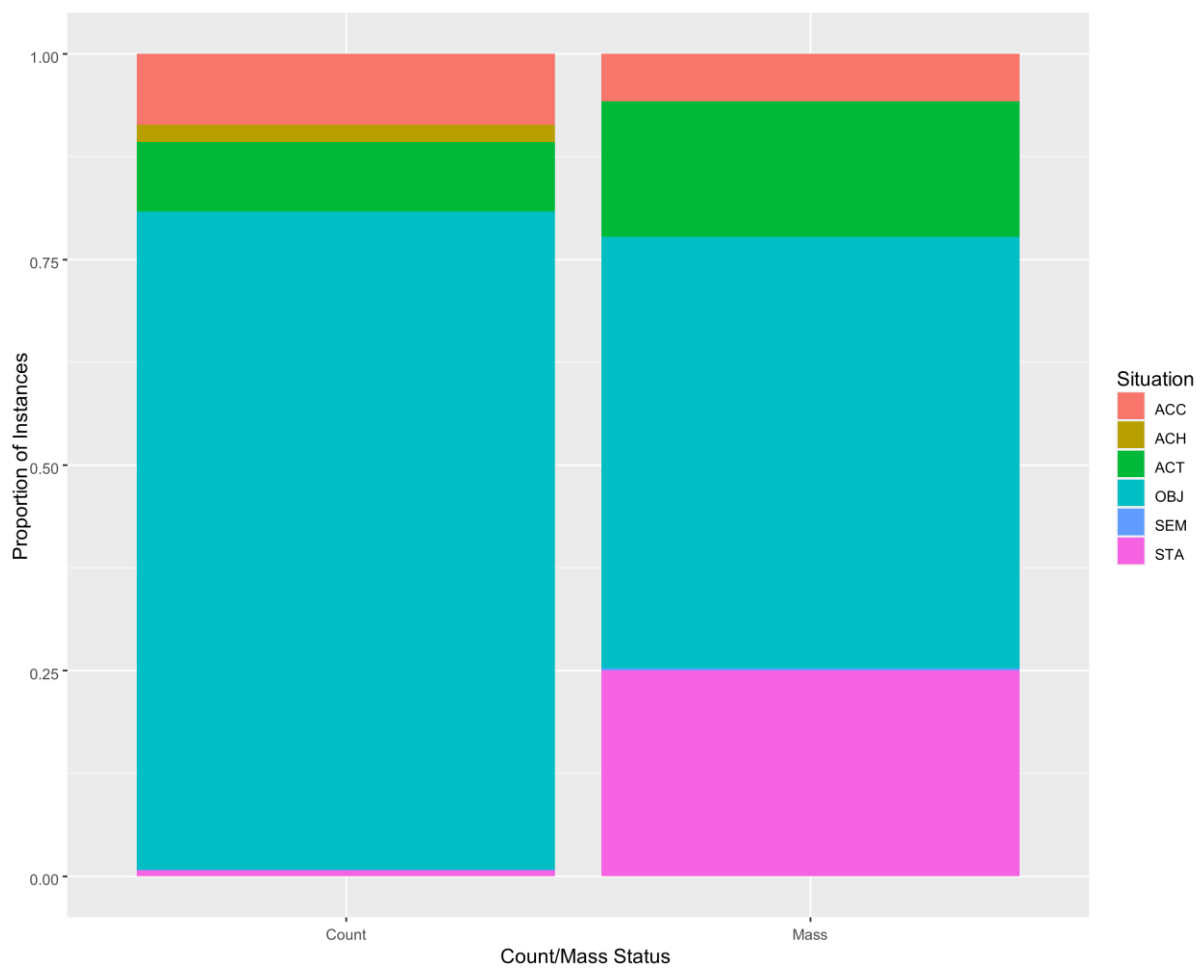


Figure 4-2. Barplot of situation type proportion expressed by count and mass nominal instances.

Taking the situation type frequencies of the count and mass nominal instances, Cramer’s V test was carried out to measure the strength of the relationship between count/mass status and situation type expression. Moreover, a Fisher-Exact test was calculated to compare the frequencies of Accomplishments; Achievements; Activities; Objects; Semelfactives; and States between count nominal instances and mass nominal instances. Overall, a moderate effect size was recorded (0.467) and a significant dependence was identified  $p < 0.001$ , which indicates that count/mass status shares a significant relationship with situation type in nominals. Therefore, the null hypothesis was rejected.

#### **4.1.2.1. The Influence of Repetition on Situation Type Distribution across Count and Mass Nominal Instances**

Once again, the influence of repetition was examined, but this time focusing on how repetition may influence situation type distribution across count and mass nominal instances. Interestingly, when repetition was considered, only minimal changes in the situation type distribution of count and mass nominals were detected. For instance, Objects are expressed 2.8% more by count nominals when repetition is considered in the dataset (an increase from 80.1% to 82.9%). Accordingly, the result from the Cramer's V test was similarly moderate (0.441) and the Fisher-Exact test statistic remained significant  $p < 0.001$ . This lack of change in the situation type distribution across count and mass nominals suggests that repetition holds little influence over the situation type distribution expressed by count and mass nominals.

#### **4.1.2.2. The Influence of Object expression on Situation Type Distribution across Count and Mass Nominal Instances**

Interestingly, when Objects were removed from the data, a strong effect size was recorded (0.647). Moreover, the Fisher-Exact test signalled that the relationship between count/mass status and situation type expression remained statistically significant  $p < 0.001$ , despite the lack of Objects. However, the omission of Objects from the data did lead to highly different situation type distributions from those shown in Figure 4-2 above. Most notably, when Objects are not considered, the expression of States presents the greatest difference in situation type expression between count and mass nominals. While mass nominals expressed States at a rate of 52.6%, count nominals only expressed States at a rate of 3.5%. In addition to States, notable differences were also recorded in relation to the expression of Accomplishments. Count nominals expressed a greater percentage of Accomplishments (43.2%) than mass nominals (12.1%). Accordingly, these results suggest that count and mass nominal instances still express a significant relationship with situation type, even when state and event semantics alone are considered. Nevertheless, this significant relationship is evidently different to the significant relationship indicated earlier in Section 4.1.2, when Objects were included in the dataset. The central difference between the distributions of count and mass nominals does not concern the expression of Objects - it is largely the result of variation in Accomplishment and State expression.

### 4.1.3. The Relationship Between Abstract/Concrete Status and Situation Type

Progressing to the relationship between abstract/concrete status and situation type, we will first remind ourselves of the definition of abstract/concrete status adopted in this study. The distinction between abstract/concrete status, according to Brysbaert et al. (2014), concerns the degree to which a lexical item expresses a perceptible entity. While nominals such as *table* and *ocean* are concrete due to their material existence, nominals such as *love* and *desire* are abstract, as they are imperceptible through the five human senses. An overview of the distribution of situation type in relation to abstract/concrete status is displayed in Figure 4-3 below, where Objects are again the principle meaning expressed in both abstract and concrete nominal instances. However, like the count/mass status distinction noted in the previous section, there is a clear difference in the variation of situation type abstract and concrete nominal instances express. Abstract nominal instances expressed Accomplishments at a frequency of 15.2%, Achievements 2.9%, Activities 18.8%, Objects 49.2%, Semelfactives 0% and States 13.9%. Concrete nominal instances, alternatively, largely orient to Object expression, and do not express as many events and states. Concrete nominals had a frequency distribution as follows: Accomplishments 2.9%, Achievements 0.7%, Activities 4.1%, Objects 91.7%, Semelfactives 0.2% and States 0.4%. In examples (27a-b) below, both an abstract Object nominal (27a) and a concrete Object nominal (27b) are shown. In addition, example (27c) displays an abstract Activity nominal, as this was the most frequently expressed type of event across both abstract and concrete nominals.

- (27) a. ...so far as it made me known & procured me employment in Zoological drawing - answered my expectations... (Noun 190, All Data without Repetition Dataset, Appendix A)
- b. The scents are possibly modified to pheromones in mating, for bees carrying scent apparently attract other males... (Noun 132, All Data without Repetition Dataset, Appendix A)
- c. ...get on and we will stop the war (Noun 60, All Data without Repetition Dataset, Appendix A).

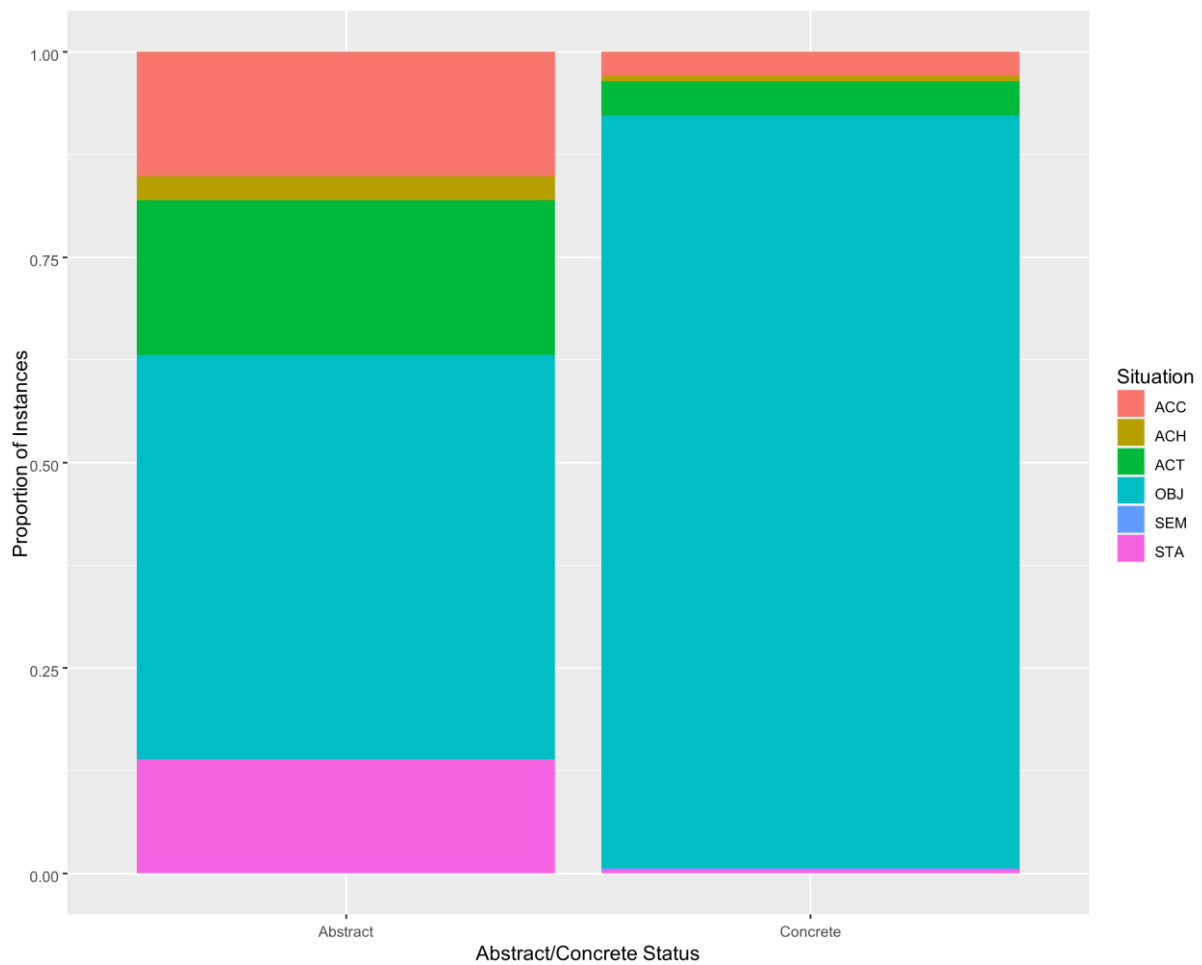


Figure 4-3. Barplot of situation type proportion expressed by abstract and concrete nominal instances.

A Cramer’s V test was again carried out to assess the strength of the relationship between abstract/concrete status and situation type expression. In addition, a Fisher-Exact test was computed to compare the frequencies of Accomplishments; Achievements; Activities; Objects; Semelfactives; and States between abstract nominal instances and concrete nominal instances. Overall, a moderate effect size was signalled (0.489) and a significant dependence was found  $p < 0.001$ , indicating that abstract/concrete status shares a significant relationship with situation type in nominals. Accordingly, the null hypothesis was rejected. The situation type distribution expressed by nominals is largely influenced by the status of the nominal as either abstract or concrete.

#### **4.1.3.1. The Influence of Repetition on Situation Type Distribution across Abstract and Concrete Nominal Instances**

When looking at the influence of repetition on situation type distribution across abstract and concrete nominals, the Cramer's V test returned a moderate effect size (0.451), and the Fisher-Exact test indicated a significant relationship between abstract/concrete status and situation type  $p < 0.001$ . The significant nature of this relationship remains unchanged even when repetition instances are included in the dataset. Looking specifically at the situation type distributions, when repetition is considered, only minimal changes in distribution between abstract and concrete nominals arise. The largest of these changes is in the Object expression of abstract nominals. Specifically, the number of Object occurrences expressed by abstract nominals increases 7.1% (an increase from 49.2% to 56.3%) when repetition is considered. Despite this small alteration however, it is clear from the general the lack of distribution change that repetition holds little influence over the situation type distribution of abstract and concrete nominals.

#### **4.1.3.2. The Influence of Object expression on Situation Type Distribution across Abstract and Concrete Nominal Instances**

Turning to the influence of Object expression on situation type distribution across abstract and concrete nominals, a weak effect size (0.27) was recorded when Object instances were omitted from the data, which suggests that the strength of the relationship between abstract/concrete status and situation type is highly related to the expression of Objects. However, the relationship between abstract/concrete status and situation type remained significant  $p < 0.001$ , showing that there is still a statistically significant difference between the situation type distributions of abstract and concrete nominals when Objects are not considered. Intriguingly, while still significantly different, the situation type distributions of abstract and concrete nominals are not quite as divergent when only based on event and state semantics. For instance, both abstract and concrete instances express a similar number of Accomplishments (29.9% and 34.6% respectively). They also express a similar number of Achievements (5.7% and 8.6% respectively).

It is only when we compare the frequencies of Activity and State expression across abstract and concrete nominals that we see the primary drivers of this statistically significant difference. Abstract nominals express considerably more States than concrete nominals, at a rate of 27.4% compared to 4.9%. Alternatively, concrete nominals express more Activities, at a rate of 49.4%

compared to 37%. Together, these results show that abstract and concrete nominal instances still express a significant relationship with situation type, even when only event and state meaning is examined. This significant relationship is therefore largely different to the significant relationship indicated earlier in Section 4.1.3. The main difference in situation type distribution between abstract and concrete nominals is no longer their expression of Objects, but their expression of Activities and States.

#### **4.1.4. The Relationship Between Genre and Situation Type**

The relationship between genre and situation type is the last relationship we will explore in this section. An overview of the situation type distribution in relation to the different genres is shown in Figure 4-4 below. Each bar represents a situation type distribution expressed by a genre. The genre in question is displayed under each bar. Mirroring the results seen in this section so far, Figure 4-4 shows that nominal instances in each genre predominantly expressed Objects. Despite this general orientation towards Object expression, there are apparent differences in the distribution of situation type across the different genres. Nominal instances in the genre of Academic Natural Science Writing had the following frequencies: Accomplishments 6.2%, Achievements 2.4%, Activities 7.3%, Objects 81.4%, Semelfactives 0% and States 2.7%. This distribution differed from the genre of Print Advertisement Writing, where nominal instances were found to have the following distribution: Accomplishments 4.9%, Achievements 1.6%, Activities 7%, Objects 83.2%, Semelfactives 0% and States 3.2%. Nominal instances in the genre of (Auto)Biographical Writing displayed a situation type distribution of Accomplishments at a frequency of 8.4%, Achievements 1.1%, Activities 11.5%, Objects 70.6%, Semelfactives 0.2% and States 8.1%. Lastly, nominal instances in the genre of Broadsheet Newspaper Home and Foreign News Writing interestingly presented the weakest orientation to Object expression across the genres, with Accomplishments occurring at a frequency of 12.2%, Achievements 2.4%, Activities 12.7%, Objects 67.3%, Semelfactives 0% and States 5.4%.

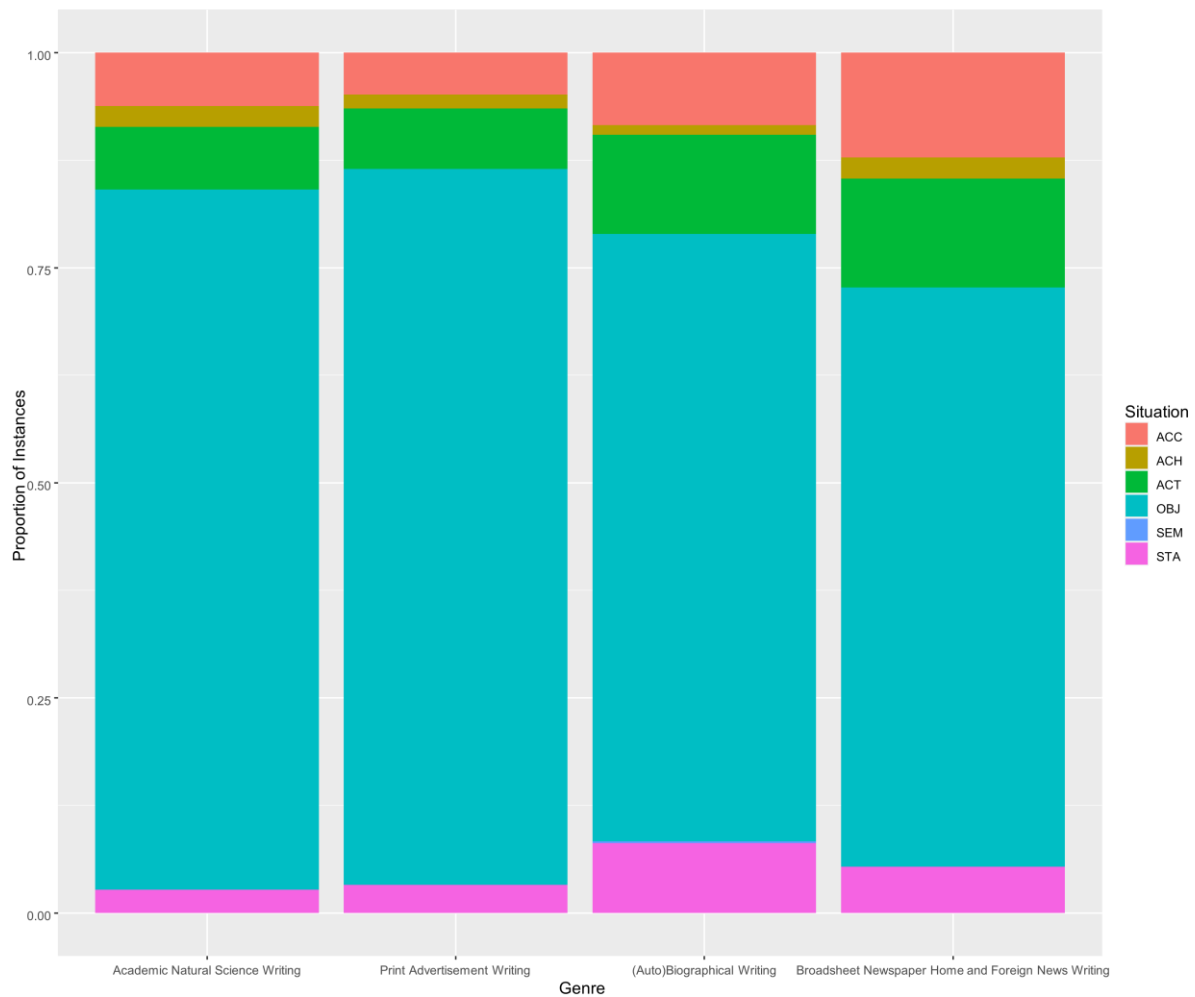


Figure 4-4. Barplot of situation type proportion expressed by genre.

A Cramer's V test was first computed to assess the strength of the relationship between genre and situation type expression. Afterwards, a Fisher-Exact test was calculated to compare the frequencies of Accomplishments; Achievements; Activities; Objects; Semelfactives; and States between each genre. Overall, a very weak effect size (0.095) and a significant dependence was identified  $p < 0.001$ , meaning that while the relationship between genre and situation type is not strong, a statistically significant difference was still reported between the genres. Subsequently, the null hypothesis was rejected. Due to this significant result, post-hoc Fisher-Exact tests were conducted on each genre comparison to identify where the significant differences in situation type distribution lied between genres. The results from these post-hoc Fisher-Exact tests are displayed in Table 4-2. Due to the implementation of the Bonferroni correction to the alpha level, the numeric result had to return a value under  $p = 0.0083$  to be considered a significant difference in situation type distribution.



From the results shown in Table 4-2, three of the six comparisons recorded a significant difference in distribution. This result further consolidates the finding that genre holds an influence over the distribution of situation type expressed by nominals. Most interestingly, both Academic Natural Science Writing and (Auto)Biographical Writing are involved in two significant genre comparisons. From Figure 4-4 above, I can suggest that the two significant differences involving Academic Natural Science Writing are the result of Academic Natural Science Writing expressing a notably higher distribution of Objects than (Auto)Biographical Writing and Broadsheet Newspaper Home and Foreign News Writing, and a notably lower distribution of States. (Auto)Biographical writing, alternatively, recorded significant comparisons because a lesser proportion of Objects were expressed in the genre, in addition to the highest proportion of States.

*Table 4-2. Post-hoc Fisher-Exact test results on comparisons of semantic behaviour expressed by different genres.*

Number	Comparison	p.value
1	Academic Natural Science Writing - Print Advertisement Writing	p=0.94
2	Academic Natural Science Writing - (Auto)Biographical Writing	<b>*p&lt;0.001</b>
3	Academic Natural Science Writing - Broadsheet Newspaper Home and Foreign News Writing	<b>*p=0.0045</b>
4	Print Advertisement Writing - (Auto)Biographical Writing	<b>*p=0.007</b>
5	Print Advertisement Writing - Broadsheet Newspaper Home and Foreign News Writing	p=0.0095
6	(Auto)Biographical Writing - Broadsheet Newspaper Home and Foreign News Writing	p=0.211

#### **4.1.4.1. The Influence of Repetition on Situation Type Distribution across Genre**

From the statistical analyses, the result from the Cramer’s V test signals an even weaker effect size (0.071) when repetition was considered. However, the Fisher-Exact test still returned a significant result of  $p < 0.001$ . Overall, the results show that the addition of repetitions in the dataset did not considerably influence the relationship between genre and situation type shown in Section 4.1.4. Looking specifically at the post-hoc Fisher-Exact tests, a further genre distribution comparison (Print Advertisement Writing - Broadsheet Newspaper Home and Foreign News Writing) expressed a significant difference when repetition was considered. Due to where these significant results lie, the results indicate a divide in the data. Academic Natural

Science Writing and Print Advertisement Writing share a similar situation type distribution. On the other hand, (Auto)Biographical Writing and Broadsheet Newspaper Home and Foreign News Writing also indicate a general similarity between their situation type distributions. Once again, the significant differences across these genres are largely attributable to the expression of Objects. While Academic Natural Science Writing and Print Advertisement Writing distribute Objects at a rate of 82.2% and 84.4% respectively, (Auto)Biographical Writing and Broadsheet Newspaper Home and Foreign News Writing only distribute Objects at rates of 75.1% and 71.2%. Through the addition of repetition instances, I therefore suggest that repetition acts to further establish the significant difference in Object expression between each side of this division.

#### **4.1.4.2. The Influence of Object expression on Situation Type Distribution across Genre**

In terms of the influence of Object expression on situation type distribution, its influence in relation to genre displays an especially interesting case. While the Cramer's V test did return a slightly stronger (although still weak) effect size (0.115), the Fisher-Exact test returned the first non-significant overall result of any relationship covered in Section 4.1. The Fisher-Exact result indicates that when Objects are removed from the data, the relationship between genre and situation type is non-significant  $p=0.1264$ . This result suggests that significant differences in situation type distribution across genre are reliant on the distribution of Objects. The distributions of event and state meanings alone by nominal instances across genres are not significantly different from one another.

In the next section, I will now consider the simultaneous influence of the variables covered in this section (word formation type; count/mass status; abstract/concrete status; and genre) on situation type distribution. In doing so, I will specifically examine which variables may be more influential on situation type distribution, and the extent to which these variables interact with one another to influence situation type distribution.

## **4.2. Simultaneous Influences on Nominal Semantic Behaviour**

In the previous section, I identified that word formation type, count/mass status, abstract/concrete status and genre each shared a dependent relationship with situation type, when each situation type (i.e., Accomplishment; Achievement; Activity; Object, Semelfactive and State) was included in the data. These relationships were all examined individually, e.g.,

word formation type and situation type, count/mass status and situation type etc. In the following section, we will again focus on these variables, but this time, I will assess their simultaneous influence on situation type distribution using a CRT analysis. The resulting decision tree visualisations present an especially interesting avenue of exploration, as they allow us to discover which variables are more influential in the distribution of situation type, and how these variables are able to interact with one another to influence situation type distribution in nominals.

The structure of this section will specifically address RQs 2 and 3 (presented below). I will first detail relevant information about the decision tree visualisation. Namely its prediction accuracy (covered in Section 4.2.1), the number of distinct classifications made by the tree and the variables that form the tree structure. During the presentation of the decision tree visualisation, I will guide the reader around the tree, detailing the results of tree including the most relevant hierarchical structures and variable interactions to the current study.

RQ2: How do the predictor variables differ in relation to their influence on the distribution of semantic behaviour expressed by nominals?

RQ3: To what extent do predictor variable interactions influence the distribution of semantic behaviour in nominals?

#### **4.2.1. CRT Analysis:**

The results of the CRT analysis convey an interesting depiction of the simultaneous influence of word formation type, count/mass status, abstract/concrete status and genre on situation type distribution. Figure 4-5 below presents a decision tree illustrating this simultaneous influence as a structure. The accuracy of the CRT analysis is measured by the number of instances that are correctly predicted by the analysis, which is called the prediction accuracy measure. This value provides an assessment of the model's predictive capability. Overall, the decision tree model returned a fairly high prediction accuracy of 75.4%, which is considerably higher than the random baseline prediction accuracy of 16.7% (a random prediction of 1 out of 6 dependent variable categories). Moreover, similar prediction accuracy percentages have been reported in other linguistic studies using CRT analysis, such as Roberts et al. (2015). Figure 4-5 also shows the nine distinct classifications that were made by the CRT analysis. Most notably, only three

predictor variables significantly influence the distribution of situation type in nominals: word formation type, count/mass status and abstract/concrete status. Genre, although present in the CRT analysis, does not record any meaningful splits in the decision tree structure.

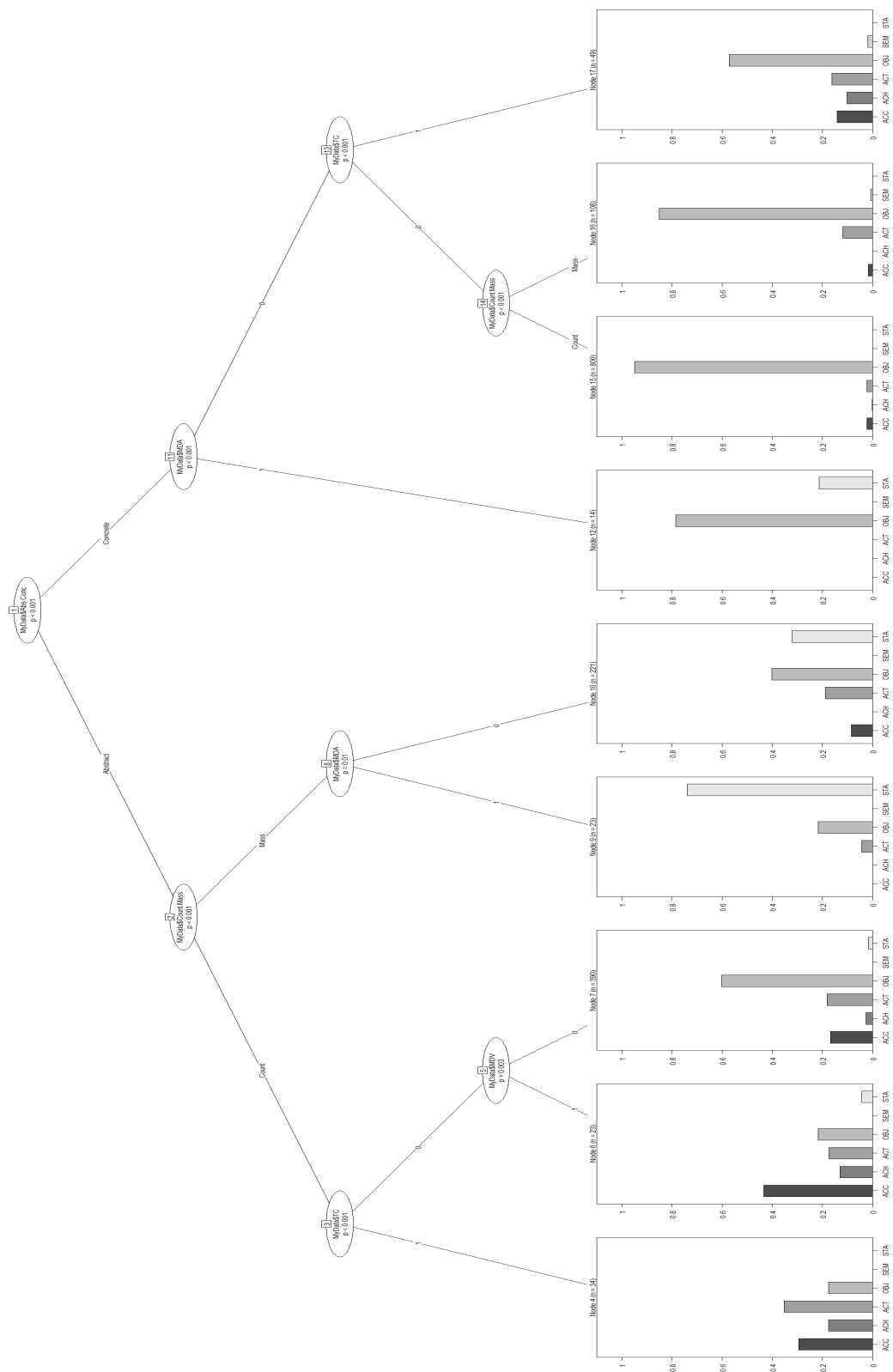


Figure 4-5. All Data without Repetition decision tree visualising situation type ~ word formation type + count/mass status + abstract/concrete status + genre.

To briefly recap on our understanding of decision tree presentation, Figure 4-5 should be read and interpreted as follows:

- The starting point of the decision tree is at the top of the decision tree structure (node 1). Node 1, i.e., the central node, represents the most influential split in the data.
- From node 1, we can either go right or left to the next node. In this decision tree, that is either node 2 or node 11. The direction taken influences the situation type distribution we will arrive at, as each situation type distribution encompasses the values of the variables in its path. For instance, if we travel from node 1 to node 11, the situation type distribution that we arrive at will be the product of two features: (1) concrete status and (2) MDA word formation type.
- The splits in the data are repeated until there are no more significant splits (under alpha level of 0.05) that can be made on a given path. The end of each path is often called the ‘terminal node’ (Deshors and Götz 2020, p.15), where the path terminates. This node provides information on the situation type distribution of a given path (in relation to the six situation types we have seen throughout this chapter), and the number of nominal instances that express a given combination of features. For instance, terminal node 12 shows that 14 nominal instances expressed that feature combination (n=14). This means that 14 nominal instances (a notably low number of observations) indicated both concrete status and MDA word formation type, resulting in the situation type distribution in terminal node 12.

Figure 4-5 displays several trends in the data. The first is that Objects are expressed notably less by abstract nominal instances than concrete nominal instances. So much so, that the distinction between abstract and concrete status operates as the central node within the decision tree structure. Evidently then, abstract/concrete status is the most influential variable in predicting situation type distribution. Such a finding may not be too surprising, as it is already well documented that abstract nominals have a greater tendency to express events and states than concrete nominals, given that many events and states represent reifications of relational concepts and not tangible things (Radden and Dirven 2007, p.78), e.g., *marriage*.

In addition to abstract/concrete status, the predictor variables of count/mass status and word formation type, while less influential than abstract/concrete status, also appear at varying levels of the decision tree structure, forming multiple predictor variable interactions. Such a high amount of predictor variable interactions in the tree suggests that variable interactions play a large part in the distribution of situation type in nominals. For instance, one notable interaction comes through the combination of node 1 (abstract/concrete status), node 2 (count/mass status) and node 8 (MDA word formation type). From this interaction we see in terminal nodes 9 (n=23) and 10 (n=221) that States are far more likely to be expressed by abstract nominals with mass status than any other feature combination in the decision tree. Specifically, the interaction of abstract status and mass status alone predicts State expression at a percentage of 32.1%. An example of this interaction is shown in (28a) below. Moreover, when MDA is also considered in this interaction, State expression increases to 73.9%. This interaction, including MDA, is shown in (28b).

(28) a. Maturation of larvae arrested due to host immunity. (Noun 82, All Data without Repetition Dataset, Appendix A)

b. but perhaps what was needed was a slightly sharper, more bristly Dora with a hidden vulnerability, which came more naturally when I gave her a North London accent (Noun 606, All Data without Repetition Dataset, Appendix A)

Turning more specifically to the influence of word formation type in the decision tree, only three (MDA, MDV, TC) word formation types cause significant data splits. This result is particularly compelling, given that there were eight word formation types involved in the analysis. The influence on situation type distribution indicated by these three word formation types generally reflects the post-hoc Fisher-Exact test results shown in Section 4.1.1. In Section 4.1.1, MDA, MDV and TC were involved in the most significant situation type distribution comparisons among the word formation types. In the above decision tree, this influence regularly interacts with the influence of abstract/concrete status and count/mass status. For instance, as we have already seen above, MDA word formation type interacts with abstract status and count status to produce the significantly different situation type distribution shown in terminal node 9. In addition to this interaction, MDA is also influential on the other side of the tree through its interaction with concrete status. Unlike the situation type distribution shown

in terminal node 9, the interaction of MDA and concrete status shown in terminal node 12 (n=14) produces a predominant orientation towards Object expression, at a rate of 78.6%.

TC word formation, like MDA, also appears on both sides of the decision tree. Interestingly, on the left side of the tree (that concerns abstract nominals), TC interacts with count status (n=34) to form the distinct situation type distribution shown in terminal node 4. Abstract, count TCs express the lowest number of Objects of all the distributions, at a rate of just 17.6%, and an increased expression of event semantics. Notably, Accomplishments (29.4%), Achievements (17.6%) and especially Activities (35.3%) are all expressed at distinctively high rates. In examples (29a-b) below, an abstract count Object TC (29a) is presented, in addition to an abstract count Activity TC (29b).

- (29) a. ...and we shall be under no liability for any extra costs incurred by such a person as a result of our doing so. (Noun 1139, All Data without Repetition Dataset, Appendix A)
- b. A number of the more left-leaning activists from the first strike... (Noun 805, All Data without Repetition Dataset, Appendix A).

The presence of TC on the (right) concrete side of the tree also makes for interesting reading. On this side of the tree, TC, like MDA, does not interact with count/mass status. It only interacts with concrete status (n=49) (terminal node 17). In this situation type distribution, Objects are expressed at their lowest rate by concrete nominals, at a percentage of 57.1%. Moreover, the highest frequency of event meaning by concrete nominals is shown in this distribution. Specifically, Accomplishments are expressed at a rate of 14.3%, while Achievements and Activities are distributed at rates of 10.2% and 16.3% respectively. Therefore, Activity once again had the highest prediction rate of all situations that expressed event meaning. Accordingly, examples (30a-b) below display a concrete Object TC (30a) and a concrete Activity TC (30b).

- (30) a. I wandered among turbanned headstones in the Muslim cemetery, mingled with the crowds on the Galata Bridge... (Noun 470, All Data without Repetition Dataset, Appendix A)



b. ... the lifelong bachelor whose main interests are astronomy, cricket and military marches... (Noun 4042, All Data without Repetition Dataset, Appendix A)

Furthermore, in a similar vein to the two previous word formation types discussed, MDV word formation is also influential on the left (abstract) side of the decision tree. The combination of abstract status, count status and MDV in terminal node 6 (n=23) creates a distinct situation type distribution, which returns a low amount of Object expression, at a rate of 21.7%. In addition to this low expression of Objects, abstract count MDVs, much like abstract count TCs, display a notably high amount of event meaning. Abstract count MDVs predict Accomplishments at a rate of 43.5%, Achievements 13%, and Activities 17.4%. We can see from these prediction rates that there is a clear orientation from abstract count MDVs to Accomplishment situation type. An abstract count Accomplishment MDV is shown below in example (31). On the concrete side of the tree, however, MDV does not express the same influence, and is therefore not responsible for any significant splits in the concrete nominal data.

(31) ...because no consistent increase in aggression followed the removal of either a relatively dominant or a low-ranking female. (Noun 2579, All Data without Repetition Dataset, Appendix A)

Lastly, as I mentioned earlier in this section, the results indicate that Object meaning occurs considerably more in concrete nominals than mass nominals. This is visualised in Figure 4-5, where the right side of the tree displays far higher percentages of Object expression than the left side. It is therefore of no surprise that on the right side we find the situation type distribution which expresses the highest number of Objects. This result occurs when concrete status and count status interact, and the influence of MDA and TC is not considered (see terminal node 15). Accordingly, concrete, count Borrowings, Compounds, MDNs, MDVs, Others, and Underiveds (n=800) together express Objects at a combined rate of 94.9%. Due to the well-documented association of nominals with object meaning (Langacker 1991, p.15; Halliday 2009, p.117), it could be suggested that nominal instances expressing this feature combination present the most 'prototypical' type of nominal. An example of this type of nominal is shown in example (32).

(32) ...government research was published showing that 32 out of 102 ovens failed to heat food (Noun 1299, All Data without Repetition Dataset, Appendix A)

#### **4.2.1.1. The Influence of Repetition on Situation Type Distribution across All Predictor Variables.**

Like the additional analyses undertaken throughout Section 4.1, a supplemental CRT analysis was carried out to examine the influence of repetition on situation type across all predictor variables. A similar analysis was also carried out on the influence of Object expression on situation type across all predictor variables, which will follow after this section.

When repetition instances were included in the dataset, a larger decision tree structure was formed, consisting of twelve different situation type distributions, i.e., terminal nodes (shown below in Figure 4-6). The prediction accuracy of this model is 78%, an increase of 2.6% prediction accuracy compared to the model shown in Figure 4-5. Like Figure 4-5, only three predictor variables significantly influence the distribution of situation type in nominals: word formation type, count/mass status and abstract/concrete status. Once again, genre did not record any meaningful splits in the decision tree structure, despite its inclusion in the CRT analysis. In addition to this structural similarity, abstract/concrete status remains the most influential predictor of situation type in nominals, once again occupying the central node position.



In terms of the significant splits in the decision tree, the structure of the tree has somewhat altered due to the inclusion of repetition instances, increasing the number of terminal nodes from 9 to 12. Firstly, when repetition instances were included in the data, one new data split arose on the abstract side of the tree in node 8. This split concerns the interaction of abstract status, count status and Underived word formation type (n=165). The product of this interaction, shown in terminal node 12, is a distribution that expresses the highest number of Objects for abstract nominals, at a rate of 80.6%. This expression of Objects is notably higher than any distribution on the abstract side of the tree. Nevertheless, I suggest that the result is the product of significance inflation given that the significant split did not arise when repetition instances were not included in the data.

On the concrete side of the tree, another new significant split is made in node 18. With the addition of repetition instances in the dataset, the influence of MDV interacts with concrete status and mass status to generate the situation type distribution shown in node 22 (n=7). Interestingly, concrete mass MDVs do not share the typically high expression of Objects shown by other concrete nominals, only expressing Objects 42.9%. However, this distribution is based on a limited number of instances (n=7). Accordingly, the validity of this distribution is highly questionable. The final additional data split that arose when repetition instances were considered focuses on the situation type distribution of concrete mass Compounds (node 19). Again, as this split is only based on a limited number of instances (n=7), I will not offer any further discussion about these situation type distributions.

#### **4.2.1.2. The Influence of Object expression on Situation Type Distribution across All Predictor Variables.**

When Objects were removed from the dataset, a far smaller decision tree structure was formed, consisting of five terminal nodes (see Figure 4-7 below). The prediction accuracy of this model is 50.2%, a considerable decrease of 25.4% compared to the model shown in Figure 4-5. This decrease in accuracy is likely related to the decrease in data used to create the model. Moreover, this decrease suggests that situation type might be easier to predict when Objects are included in the dataset. Despite this decrease, the accuracy of the model remains notably higher than the random baseline prediction accuracy of 20% (a random prediction of 1 out of 5 dependent variable categories). Like the previous decision tree structures discussed above, three of the four predictor variables still significantly influence the distribution of situation type in the tree

structure. Accordingly, we should note that genre does not arise as significantly influential in any of the decision trees discussed throughout Section 4.2.1. From the size of this decision tree, it is apparent that the presence of Objects is highly influential in the creation of significantly different situation type distributions.

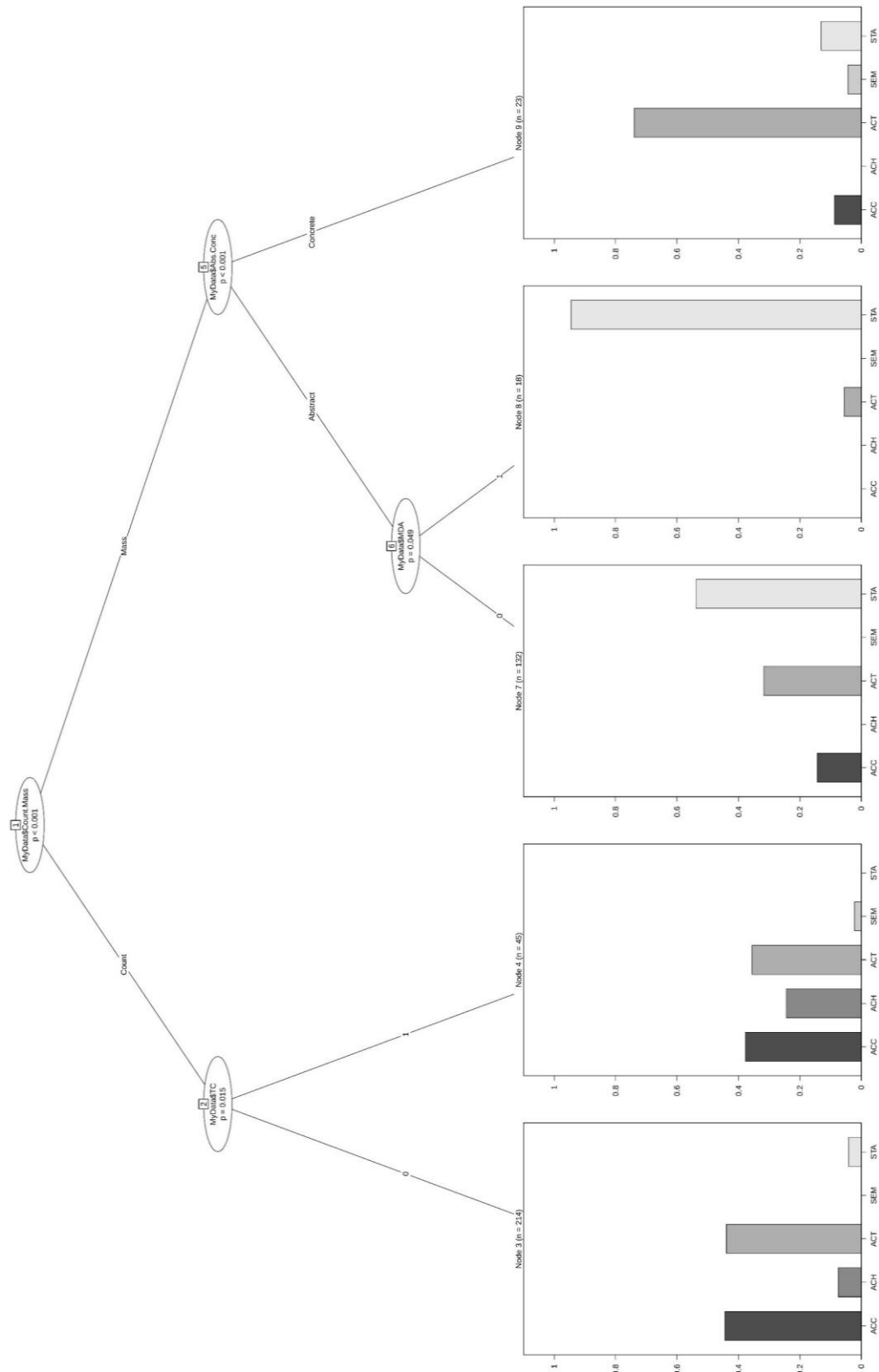


Figure 4-7. No Object Data without Repetition decision tree visualising situation type ~ word formation type + count/mass status + abstract/concrete status + genre.

Interestingly, when Objects are removed from the dataset, count/mass status operates as the central node in the tree structure and is therefore the most important variable in predicting situation type distribution. In Section 4.2.1, I discussed how it was not too surprising that abstract/concrete was the central node in Figure 4-5, given that many events and states represent reifications of relational concepts and not tangible things (Radden and Dirven 2007, p.78), e.g., *marriage*. Taking this information, it is perhaps expected that when tangible Objects are removed from the dataset, the semantic distinction expressed by abstract and concrete nominals would diminish.

This alteration in structure marks a compelling shift, as it allows us to view patterns in the data that were not as easily detected beforehand in Figure 4-5. For instance, in the terminal nodes on the right side of the tree, Activities and States are generally the most prevalent type of situation type expressed. On the left side of the tree, the pattern is different, as, while Activities are still prevalent, the expression of Accomplishments and Achievements increases. Such a result is particularly interesting as it links to previous suggestions by Mourelatos (1978), Brinton (1995), and Bauer et al. (2013) about a supposed relationship between the concepts of telicity and count/mass status. This point will be expanded upon further in Chapter 5 (Section 5.2.2), where I will directly discuss the relationship between count/mass status, telicity and the notions of boundedness and homogeneity in the nominal domain.

Lastly, while I have pointed out many changes in the structure of the tree in Figure 4-7, certain feature interactions remain similar, despite the removal of Object instances. For example, the interaction of abstract status and mass status (n=132) still provides fertile ground for the expression of States. This is apparent in terminal node 7, where States are expressed at a rate of 53.8%. In addition to this interaction, terminal node 8 (n=18) also presents a familiar outcome to that shown in Figure 4-5, as when MDA interacts with abstract status and mass status, the expression of States rises. In terminal node 8, a very high percentage of States is predicted (94.4%).

### **4.3. Concluding Comments**

In this section, we will now remind ourselves of the main results from the analyses, before moving onto the following discussion chapter. Firstly, the Fisher-Exact test analysis signalled that word formation type, count/mass status, abstract/concrete status and genre each shared a

dependent relationship with situation type, when each situation type and object meaning (i.e., Accomplishment; Achievement; Activity; Object; Semelfactive; and State) was included in the data. Looking specifically at the relationship between word formation type and situation type, the post-hoc Fisher-Exact tests showed that the significant nature of this relationship was largely the result of the divergent situation type distributions expressed by MDAs, MDVs and TCs. In addition to word formation type, post-hoc Fisher-Exact tests also assessed the relationship between genre and situation type. These tests revealed that nominals in the genre of Academic Natural Science Writing expressed the most divergent situation type distribution.

From the CRT analysis, it was identified that abstract/concrete status held the most influence over expression of situation type. However, this influence did decrease greatly when Objects were removed from the dataset. Along with abstract/concrete status, count/mass status and word formation type were also identified as influential variables in the CRT analysis. In the decision tree structure, both count/mass status and word formation type interacted with abstract/concrete status and one another to form significantly different situation type distributions. Following a similar trend shown in the post-hoc Fisher-Exact test results in Section 4.1.1, the influence of word formation type in the tree structure was largely driven by three word formation types: MDA; MDV; and TC. On the other hand, genre, despite registering a significant relationship with situation type in Section 4.1.4, was not responsible for any significantly different situation type distributions in the CRT analysis. Such a result suggests that while genre is influential on the situation type expression of nominals, it is markedly less influential than the other three predictor variables assessed in this study.

The results outlined in this chapter will now be discussed in the following chapter, with focus dedicated to each variable's influence over the situation type distribution by nominals. The chapter will centre on each of the predictor variables in the CRT analysis that were responsible for significant splits in the data: abstract/concrete status; count/mass status; and word formation type. When discussing the influence of word formation type specifically, the following chapter will only explore the semantic behaviour of MDA, MDV and TC, as they expressed the most significantly different situation type distributions, and regularly interacted with abstract/concrete status and count/mass status in the CRT analysis. Throughout all the discussions in the next chapter, I will address the relevance of these findings in relation to our current understanding of the nature and degree of nominality in addition to how event and state meaning come to be expressed in nominal forms.



## **Chapter 5: Investigating Category Feature Influence on Nominal Semantics – Discussion**

This chapter provides a comprehensive discussion of the results presented in Chapter 4. In doing so, the chapter will be split into three sections that will each focus on one of the three predictor variables that caused significant data splits in the Classification and Regression Tree (CRT) analysis: abstract/concrete status; count/mass status; and word formation type (see Section 4.2). These sections will principally explore the nature of the relationships these variables share with nominal semantic expression and will draw upon interesting examples from the data to support this discussion. Considering the results of the previous chapter, this chapter will discuss these findings in relation to two of this thesis' sub-aims:

1. To determine how event and state meaning come to be expressed in nominal forms.
2. To determine how the nature and degree of nominality can be evaluated in English.

I will first explore the influence of abstract/concrete status on situation type expression, providing justification as to why this influence was the strongest of all the variables examined in the study. Afterwards, the chapter will then move to count/mass status, placing specific focus on the interaction of mass status with abstract status and its resulting semantic expression. This section will then critically examine the relationship between 'bounding', 'homogeneity' (for complete descriptions of both these terms, see Section 5.2.1) and the lexical aspect feature telicity in the nominal domain. While examining this relationship, the chapter will consider the validity of the relationship suggested by Mourelatos (1978); Brinton (1995); and Bauer et al. (2013) between both lexical aspect and count/mass status and whether this association still applies when situation types are expressed in the nominal domain. The chapter will lastly cover the influence of word formation type on situation type expression. This section will pay specific attention to the influence of MDA; MDV; and TC – the most semantically divergent word formation types in the study - on nominal semantics.

### **5.1. Abstract/Concrete Status: The Most Influential Variable**

It was apparent from the results shown in Chapter 4 that abstract/concrete status held the strongest influence on situation type expression of all the predictor variables (see Section

4.2.1). In this section, I offer an explanation for why abstract/concrete status holds such an influence over the temporal semantics expressed by nominals. Before we engage with the crux of this issue however, it is worth pointing out that the definition of abstract/concrete status adopted in this thesis is available in section 3.2.3.

### **5.1.1. Abstract/Concrete Status and Situation Type Distribution**

To efficiently dissect the influence of abstract/concrete status on situation type expression by nominals, this section will bring back Figures 4-3 (reproduced here as Figure 5-1) and 4-5 (reproduced here as Figure 5-2) as reference points for the following discussion. Figure 5-1 displays information on the proportion of situation types expressed by abstract and concrete nominals. Figure 5-2, alternatively, provides a decision tree structure of information on the simultaneous influence of word formation, count/mass status, abstract/concrete status and genre on the situation type expression of nominals. What is also interesting to observe in Figure 5-2 is how abstract/concrete status interacts with these other variables to create significantly different situation type distributions. Both figures are presented below.

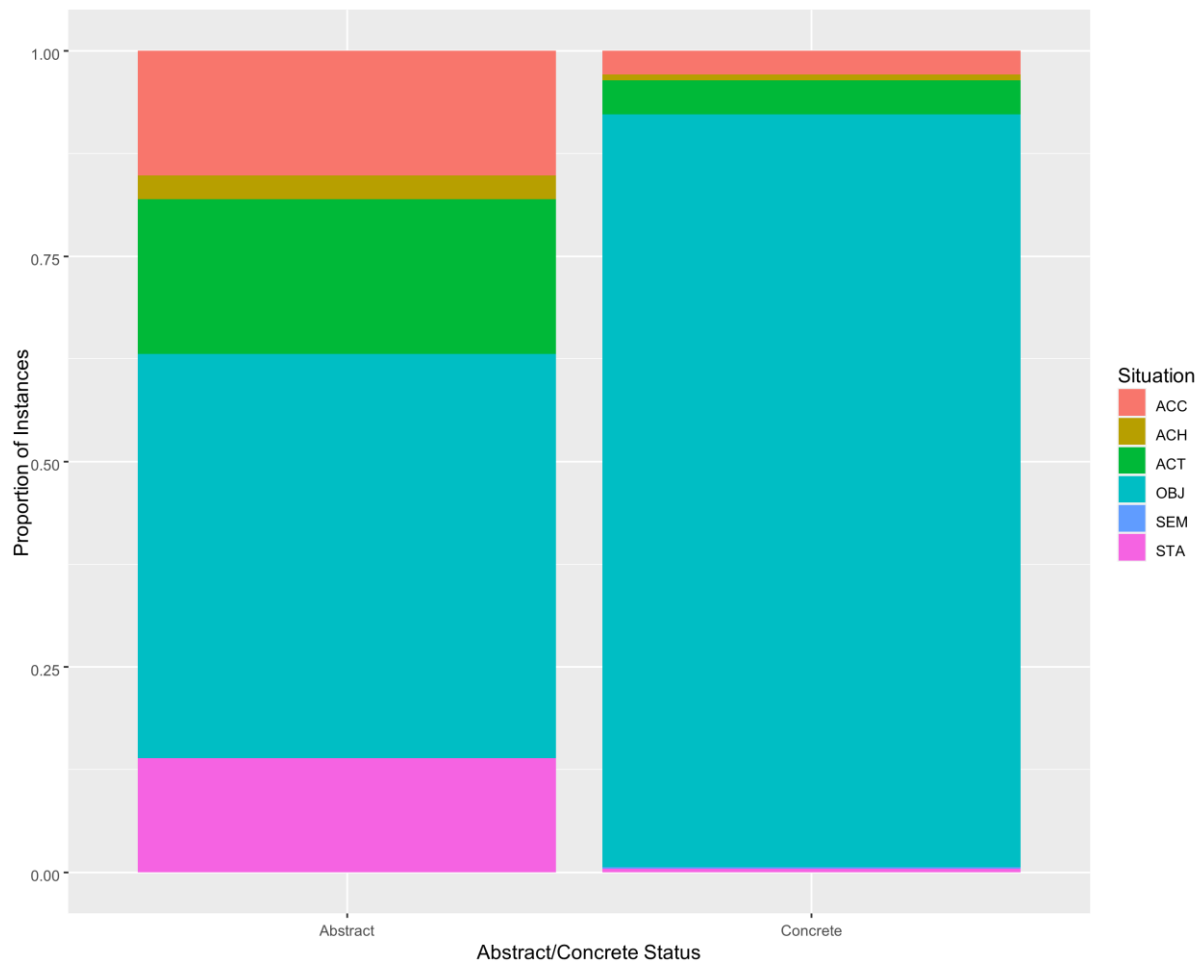


Figure 5-1. Barplot of situation type and object meaning proportion expressed by abstract and concrete nominal instances.

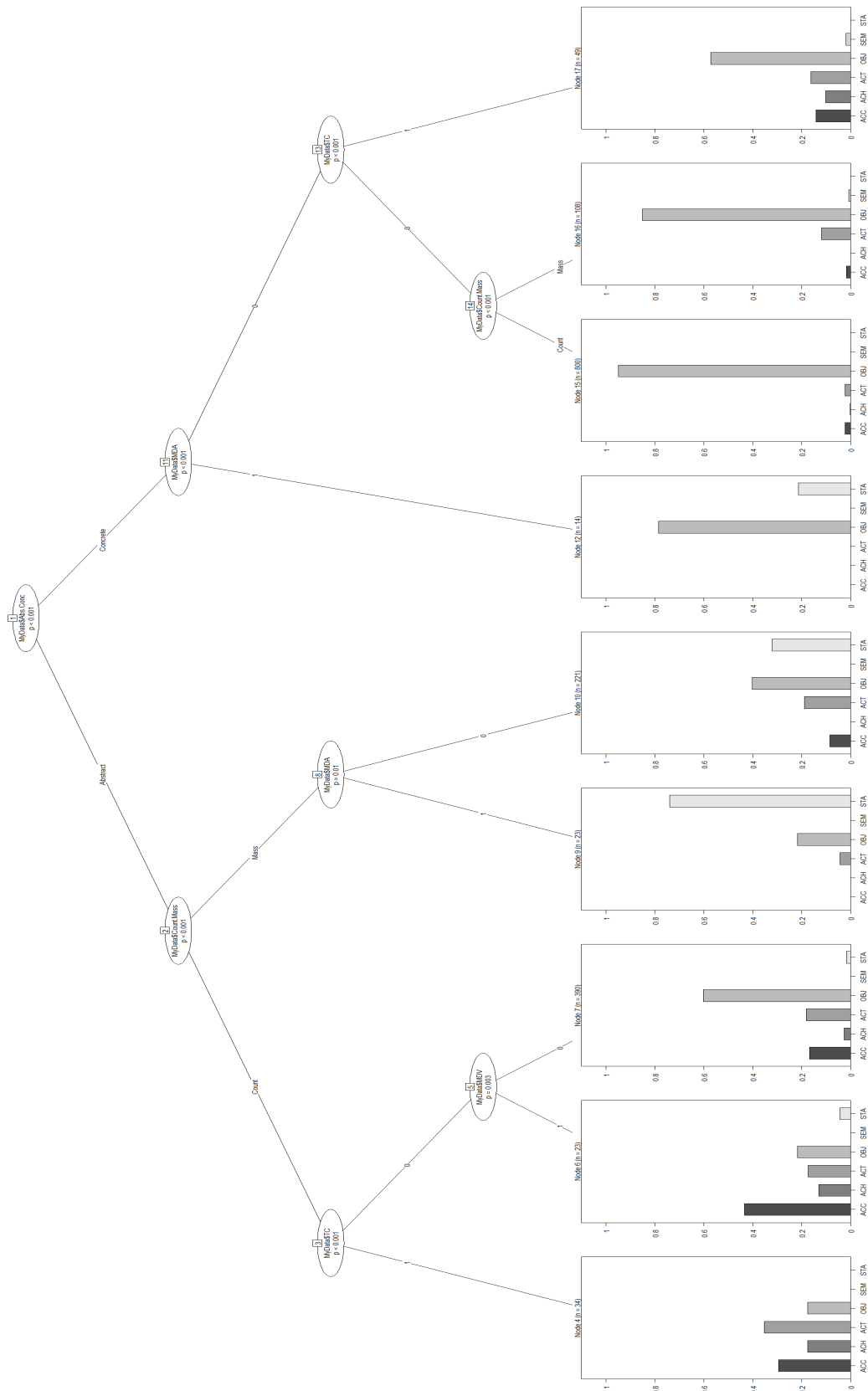


Figure 5-2. Decision tree visualising situation type ~ word formation type + count/mass status + abstract/concrete status + genre.

In Figure 5-1, the main difference between the semantics expressed by abstract and concrete nominals concerns the expression of Objects. While abstract nominals expressed Objects at a rate of 49.2%, concrete nominals expressed Objects at a greater rate of 91.7%. As a result, a far larger variation of situation type was expressed by abstract nominals, with abstract nominals expressing considerably higher values of event and state meaning – specifically, Accomplishments, Activities and States. For the exact percentages of all the situation types in these distributions, see Section 4.1.3.

In Figure 5-2, abstract/concrete status obtains the central node position in the decision tree structure (i.e., the most influential variable on situation type expression). From the structure, it is evident again that Objects are expressed considerably more by concrete nominals than abstract nominals, as Object expression is far higher in concrete nominal situation type distributions. Taking this result into account, it is fair to claim that the substantial influence abstract/concrete status holds over situation type distribution in nominals is the result of the highly distinguished expression of Objects by abstract and concrete nominals. Such a result acts to support current literature on the semantics of abstract and concrete nominals; that concrete nominals are more likely to be Objects, while abstract nominals have a greater tendency to express more relational concepts and situations (Radden and Dirven 2007, p.78). However, it is still worth exploring why this distinction in Object expression is apparent between abstract and concrete nominals. In the following paragraphs, I will explore this relationship, using data examples, whilst also making reference to Langacker's (1987a; 1991) theory of Cognitive Grammar. A background into the foundations of Cognitive Grammar is available in this thesis, beginning in Section 2.2.1 and following through until Section 2.2.4.

As stated above, the greatest semantic distinction between abstract and concrete nominals was related to the expression of Objects. For concrete nominals, a high expression of Objects is not surprising. Concrete entities, according to Fontaine (2017c, p.180), allow people to identify with the lexical class noun more easily, given that they represent tangible entities which people can interact with and observe, such as such as *dog* and *ball*. For an entity to exist in the physical domain of space, it must represent some kind of physical entity and, as a result, be perceptible by at least one of the five human senses. But with the concept of physicality comes the potential of certain temporal limitations. For instance, concrete count nominals, as shown in examples

(33a-b) below, are typically spatially compact, i.e., they represent a restricted amount of substance which all shares a contiguous space (Givón 2001, p.51). This high degree of spatial compactness influences the temporal expansiveness of the entity, as entities that interact with time tend to unfold over a larger amount of space and, as a result, are not spatially compact (Givón 2001, p.52). Consequently, entities which are more spatially compact, due to their limited spatial expanse, typically express static entities, with little interaction with time (Langacker 1991, p.14).

(33) a. ...the emperor seems to have sat at an inconvenient height at his desk in a small chair, covered in green plush... (Noun 1023, All Data without Repetition Dataset, Appendix A)

b. ...it was the first time I wore a rather special red coat with matching red hat. (Noun 2091, All Data without Repetition Dataset, Appendix A)

That is not to say that entities which are not as spatially compact as those in example (33) cannot also be static. For instance, nominals like the concrete mass nominal *glue*, shown in example (34), differ to the nominals we have seen in example (33), as unlike *chair* and *coat*, our conception of *glue* is not limited, i.e., bounded (for an in-depth discussion about the concept of ‘bounding’, see Section 5.2.1), within the domain of space. *Glue* is not a discrete object. It is instead more spatially expansive, made up of many scattered instantiations (Langacker 1990, p.99). However, *glue* represents a physical substance. Because of this physicality, the principal domain in which *glue* is instantiated remains the domain of space, despite its spatial expansiveness. Accordingly, as the physical composition of many concrete count and mass entities leads to their conceptualisation being principally realised within the domain of space, their engagement with temporal semantics is restricted. As a result of this temporal restriction, and as shown in Figures 5-1 and 5-2, many of these concrete nominals express an atemporal semantic structure, i.e., ontological objects. Such a capacity also relates to why the influence of abstract/concrete status on situation type expression decreased when Objects were omitted from the dataset; the distinction between the semantics expressed by concrete nominals and abstract nominals decreased when the enhanced capacity of concrete nominals to express atemporal semantic structures was removed.

- (34) David Hockney’s latest work arrived in his home city of Bradford last night in a welter of spray-on glue and sheets of fax paper... (Noun 1469, All Data without Repetition Dataset, Appendix A)

Now let’s visit the other side of the coin. As established above, it is regularly the physical nature of concrete nominals and their primary instantiation in the domain of space that impedes certain nominals from expressing temporal semantics. For abstract nominals, this semantic restriction is not relevant, as abstract nominals express more intangible referents and relational concepts, such as ideas, qualities, states, particular events etc. When realised in nominal form, relational concepts are viewed as abstract ‘things’ through the process of ‘reification’ (Langacker 2008a, p.105). Reification (as discussed earlier throughout Section 2.2) in abstract nominals allows one to view relational concepts as things, and imposes a sort of ontological existence onto abstract concepts (Radden and Dirven 2007, p.78). This conceptual reification enables us to view something as a unitary entity for higher-order cognitive purposes (Langacker 2009, p.288). In example (35) below, the nominal *exploration* is expressed as a thing. Consequently, we can predicate many of the same attributes to an *exploration* as we can concrete count nominals. For instance, both a *table* and an *exploration* can be described as *large, robust etc.*

Where many of these abstract nominals differ semantically from concrete nominals is their increased potential for the expression of temporal meaning. During the reification process of relational concepts such as events and states, Langacker (1987b, p.90) claims that the serial stative relations of events or states group together to form the constitutive entities of a region, i.e., a thing. This reification process does not substantially alter the conceptual content of the original relational concept. Typically, abstract reifications of events and states preserve elements of their “relational and often processual character” when expressed as abstract nominals (Radden and Dirven 2007, p.79). As a result of this preservation, certain abstract nominals can satisfactorily enter event syntactic contexts like ‘*did NOUN take place*’ and express temporality. For instance, if we visit example (35) again below, an *exploration* can fit into event syntactic contexts as an *exploration* can *take place*, and it can also *be observed over a period time*. Thus, it is the enhanced capacity of abstract nominals to largely preserve the original relational/processual meaning potential of concepts during the reification process (where concepts are viewed and expressed as things) that typically results in a greater number of events/states in abstract nominals than concrete nominals. This capacity of abstract nominals

is evident in the results shown in Figure 5-1 above, where abstract nominals express temporal semantics at a high rate of 50.8%.

- (35) Exploration in Northern Ireland was facilitated by the passage of the Mineral Development Act (Northern Ireland) 1969. (Noun 3724, All Data without Repetition Dataset, Appendix A)

Now that I have argued reasons why abstract/concrete status held the strongest influence over the expression of situation type, I will now focus on the influence of count/mass status. In doing so, I will principally look at the interaction of abstract/concrete status and count/mass status, and how count/mass status seemingly holds a greater influence over the semantics of abstract nominals than concrete nominals. Moreover, the section will assess the relationship between the notions of bounding, homogeneity, and telicity in the nominal domain, using examples such as count Activities and count States.

## **5.2. Count/Mass Status and Semantic Behaviour**

This section will first provide a recap on the definition of count/mass status used in this thesis. To help establish this definition, I will introduce the concept of ‘bounding’, along with the terms ‘homogeneity’, ‘contractability’ and ‘replicability’. Afterwards, the section will explore the finding from Section 4.2 that count/mass status holds greater influence over the situation type expression of abstract nominals than concrete nominals. The rest of Section 5.2 will then primarily focus on the supposed relationship between lexical aspect and count/mass status in the nominal domain - that Accomplishment, Achievement and Semelfactive situation type correlate with count nominals (Bauer et al. 2013) and Activity and State situation type correlate with mass nominals (Mourelatos 1978; Brinton 1995). I will assess the validity of this relationship against the data analysed and identify whether this relationship holds when situation type is expressed in the nominal domain. During this discussion, attention will be placed on the count Activity and State instances found in the data, which do not subscribe to the relationship outlined above.

### **5.2.1. Recapping Concepts: Count vs Mass**

In Section 3.2.3, I stated that count/mass status relates to a syntactic property of nominals that determines whether a nominal can be pluralised, occur with an indefinite determiner, allow



numeral modifiers and occur with the quantifiers (Kiss et al. 2017, p.189). The examples I provided to explicate this brief definition were *house* and *salt*. While a *house* can be pluralised and thus counted (e.g., two *houses*), *salt* tends not to express this syntactic quality, as generally speaking, it is not pluralised (e.g., \*two *salts*). Following on from this brief definition, I will now define count/mass status in greater detail. To do so accurately, I will first outline the notion of ‘bounding’.

Bounding, as discussed in Section 3.2.3, fundamentally relates to whether a thing or process is expressed as discrete from other entities of the same type, through the existence of a boundary (Bennett 2014, p.38). Bounding denotes the cognitive operation where lexical items are either construed as within the ‘Immediate Scope’ (IS) in their domain, i.e., ‘bounded’, or construed as greater than the IS in their domain, i.e., unbounded (Langacker 2013, p.132). The IS in this definition relates to the part of the entity which is the relevant focus of attention; a subsection of the ‘Maximal Scope’ (MS), which relates to the full extent of the entity’s coverage (Bennett 2014, p.32). A visual representation, taken from Langacker (2008a, p.133), is used below in Figure 5-3 to graphically depict the notion of bounding in relation to count and mass nominals. Count nominals generally express bounded entities, as they represent a limited region within a domain (Langacker 1987b, p.58). For instance, if we take *a book* as an example, we understand that, when examining the constitutive entities of *a book*, such as particles of paper, we will naturally reach a point where these particles are no longer extended. At this point, we are met with a point of contrast between the book and its surrounding environment. The region of *the book* is limited. Mass nominals, alternatively, express unbounded entities. They represent a region of indefinite extension. For example, *sand* represents a typical construal of a mass substance, as *sand* does not represent a bounded region in the domain of three-dimensional space. *Sand* consequently lacks the ability to undergo pluralisation, as we will never reach the point at which a single instance of *sand* is indefinitely exhausted, e.g., \*two sands.

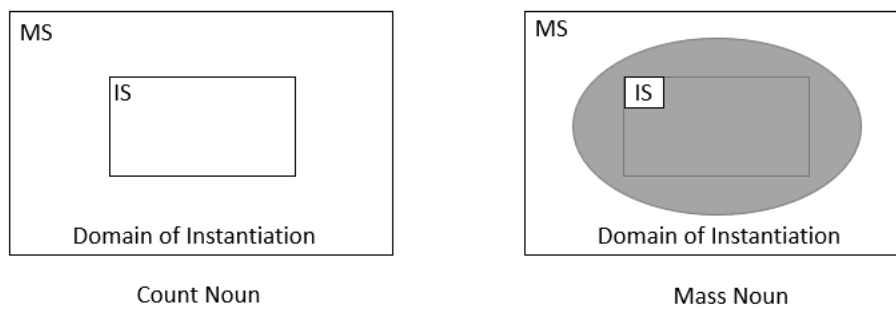


Figure 5-3. Distinction of boundedness between count and mass nominals (Langacker 2008, p.133).

It should be pointed out here that bounding specifically relates to how an entity is construed rather than how it is objectively discerned (Drożdż 2017, p.64). For instance, no mass is conclusively the same throughout when examined at a fine enough detail. While objectively we can separate *sand* into its individual grains, a specific level of schematicity is generally invoked in our conception of *sand* which neutralises the minor differences between the constituents of the entity (i.e., *the sand*). As Langacker (2008a, p.140) claims with the similar example *water*, we use the lexical item *water* every day without feeling obliged to draw upon its chemical analysis.

Bounding, however, is not just determined on our conception of IS alone. The internal configuration of an entity is also a factor that can help determine whether we view an entity as bounded. Taking an example from Langacker (2008a, p.137), a *bicycle* is an entity composed of a particular set of parts that are placed together in certain way for the *bicycle* to adequately function. To recognise an instance of this bounded type and to interpret an entity as a *bicycle*, we must observe all these parts in their appropriate formation. Otherwise, the entity's status as a bounded instance of a *bicycle* remains unclear. Furthermore, the functionality of an entity also influences how that entity is bounded. For some entities, it is not enough to just observe the region the entity obtains, as it is unclear where some entities end and others begin (Drożdż 2017, p.63). Once again using an example from Langacker (2008a, p.137), we can instantly view the complexities of bounding when confronted with an entity such as *the handle of a baseball bat*. Questions arise as there is no obvious boundary between the *handle of the bat* and *its barrel*. Where our conception of boundedness for this entity arises is in its function: *a handle* is used for holding the bat, while *the barrel* is generally used for striking a ball. The

point at which the handle ends is the point where it is no longer appropriate for the batter to hold the bat to achieve the best strike.

In addition to the concept of bounding, we also have the terms ‘homogeneity’ and ‘heterogeneity’, ‘contractability’ and ‘replicability’. The notions of homogeneity and heterogeneity relate to whether particular quantities of space and time are viewed as either continuous or discrete (Evans and Green 2006, p.534). Mass nominals are typically viewed as homogenous, as they are understood as continuous entities, where every substance of the entity is qualitatively the same. Count nominals, on the other hand, are typically considered as heterogeneous, as they are countable and therefore express discrete boundaries. Count nominals express an inherent diversification in the character of the entities they express, whereas mass nominals do not. This distinction between homogeneity and heterogeneity relates to ‘contractability’. Mass nominals, due to their typically homogenous nature, are generally contractable, in that a portion of mass is still able to provide a valid instance of a type (Langacker 2008a, p.141), e.g., a portion of *sand* is still *sand*. This quality is generally not present in count nominals, as a portion of their composition will not qualify as a valid instance of a type, due their discrete nature. For instance, *a boot of a car* is not a *car*, only *part of a car*. Lastly, we can also distinguish count and mass nominals through their ‘replicability’ (Bennett 2014, p.36). As count nominals are discrete, they are inherently replicable. Adding more of the same entity will result in separate entities. Mass nominals, alternatively, are not replicable, as due to their continuous nature, the replication of a mass entity will only result in more of the same entity, and not a separate instantiation (Bielak and Pawlak 2013, p.34).

Like nouns, verbs are also subject to the notion of boundedness (as discussed in Section 2.4.1). Perfective verbs, in particular, represent bounded and replicable events, as perfective grammatical aspect expresses a situation in its entirety (Declerck 2007, p.52). For example, the verb *baked* in *Jack baked a cake* describes a perfective event that is bounded by reaching its endpoint – once the cake is baked, the event is terminated. Imperfective verbs, on the contrary, express unbounded, internally homogenous states and processes that are unsusceptible to change (Williams 2002; Niemeier and Reif 2008). For example, the verb *running* in *Jack is running* relates to a continuous process that is construed as still unfolding. From this discussion, you might think that boundedness is largely similar to the notion of telicity. However, the difference between these notions is that boundedness concerns whether a situation is described as reaching a temporal boundary whereas telicity concerns whether the situation has a temporal

boundary (Declerck 1989, p.277). To convey this distinction, we can visit Depraetere's (1995) example verb of *sunbathing*. Although the situation of sunbathing is over (bounded) when an individual decides to stop, this endpoint is not inherent to the situation of sunbathing. So, although the situation can be construed as bounded, it still expresses a lack of telicity. Unlike telicity, the notion of homogeneity in the verb domain largely shares a relationship with the concept of boundedness, as homogenous situations are typically unbounded, e.g., *he is running* (Huyghe 2011, p.113). There is a general one to one relationship between these two notions. As we will see in Section 5.2.3, this is not always the case in the nominal domain.

From these examples of how verbs are conceptualised in relation to bounding, it would appear that there is a somewhat similar relationship between count/mass status in nouns and the perfective/imperfective distinction expressed by verbs. This sentiment is echoed by Mourelatos (1978); Talmy (1978); Bach (1986); Jackendoff (1990; 1991); Brinton (1998), Kearns (2000); and many others. As a result of this apparent relationship, it is suggested that there is an additional correspondence between telic situations such as Accomplishments, Achievements and Semelfactives and count status (Bauer et al. 2013), as well as a correspondence between atelic situations, such as Activities and States, and mass status (Mourelatos 1978; Brinton 1995). However, as we will explore later in Section 5.2.3, these correspondences may be an overgeneralisation. In the next section, I will now discuss the results (shown in sections 4.1.2 and 4.2) related to the influence of count/mass status on the expression of situation type in nominals.

### **5.2.2. Count/Mass Status: An Unequal Influence**

To discuss the results on the influence of count/mass status on the expression of situation type by nominals, I will use two figures: Figure 5-4 and Figure 5-5 (also shown in Figure 5-2). Figure 5-4 presents information on the proportion of situation types expressed by count and mass nominals. On the other hand, Figure 5-5 presents a familiar figure (shown in the previous section) that provides a decision tree structure of information on the simultaneous influence of word formation type, count/mass status, abstract/concrete status and genre on the situation type expression of nominals. These figures are presented directly below.

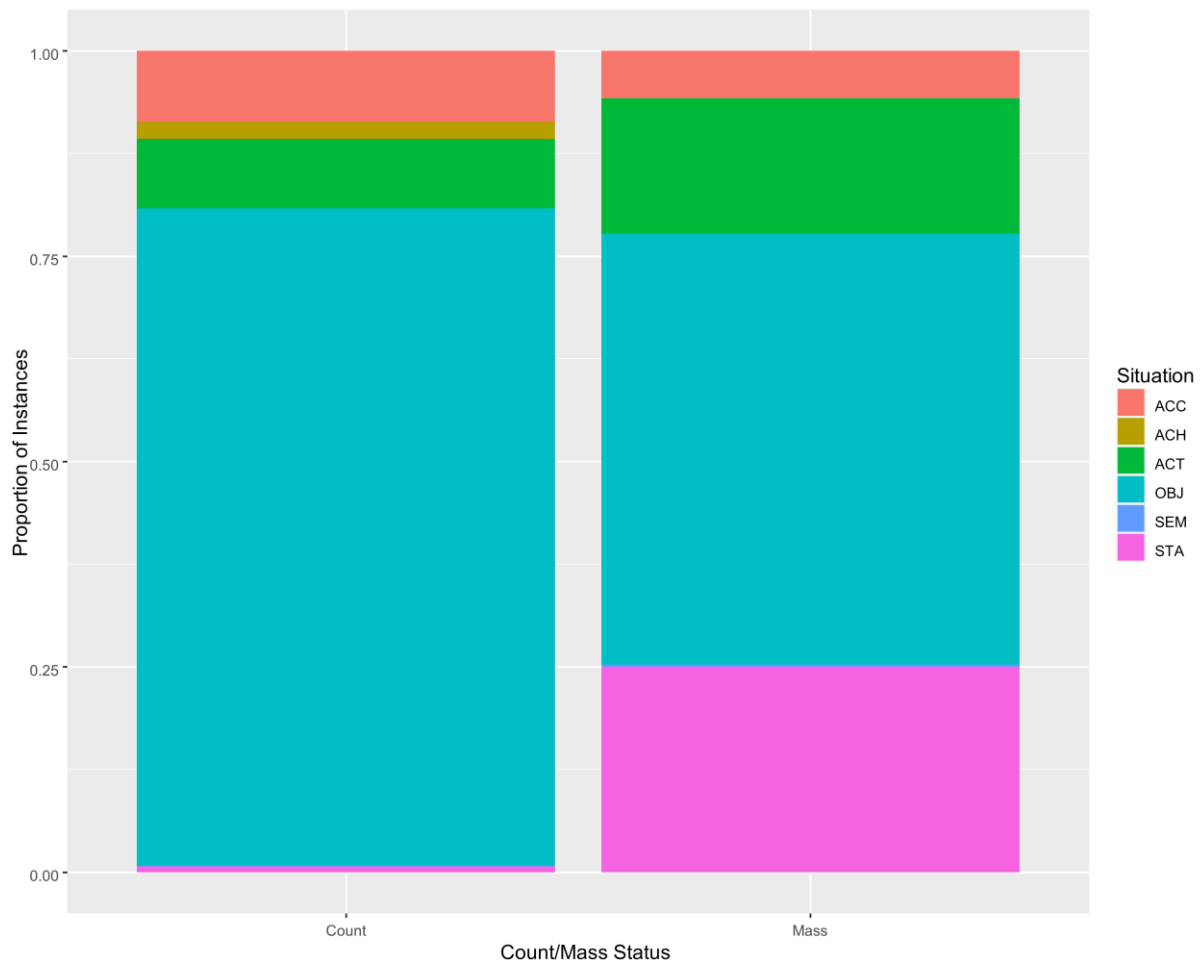


Figure 5-4. Barplot of situation type proportion expressed by count and mass nominal instances.

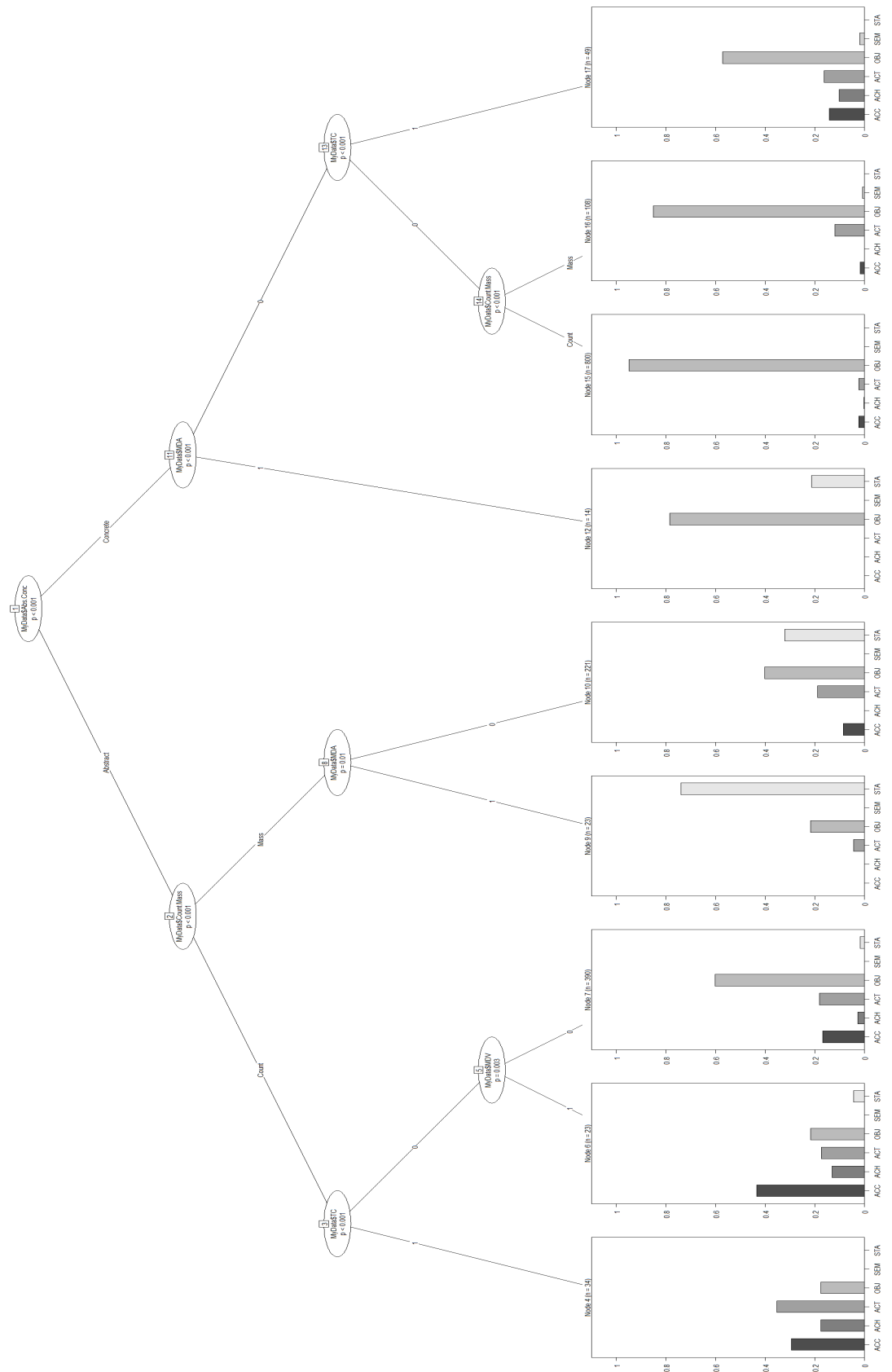


Figure 5-5. Decision tree visualising situation type ~ word formation type + count/mass status + abstract/concrete status + genre.

From Figure 5-4 above, the main semantic distinction between the count and mass nominals is the expression of Object and State situation type. While count nominals expressed Objects and States at a rate of 80.1% and 0.7% respectively, mass nominals expressed Objects and States at a rate of 52.5% and 25%. Below, examples (36a-b) provide instances of count Object and State nominals, while examples (36c-d) provide instances of mass Object and State nominals.

- (36) a. I have in my possession a letter from one of the RAF's most revered names... (Noun 67, All Data without Repetition Dataset, Appendix A)
- b. ...this can begin to do justice to the life and work of Paul Nizan, inextricably enmeshed as it is in the illusions, deceits, hopes, aspirations, successes and failures of its time. (Noun 4844, All Data without Repetition Dataset, Appendix A)
- c. ...it also smothered the various Instructors who were resting whilst the pupils were carrying out a detail in the air. (Noun 52, All Data without Repetition Dataset, Appendix A)
- d. ...I have since acquired an insatiable thirst and desire for champagne. (Noun 55, All Data without Repetition Dataset, Appendix A)

In addition to these more obvious distinctions, telic situation types, such as Accomplishments and Achievements, were expressed more by count nominals than mass nominals. The high expression of State situation type by mass nominals and the greater expression of Accomplishments and Achievements by count nominals relates to previous observations by Mourelatos (1978); Brinton (1995); and Bauer et al. (2013) about the relationship between count nominals and telic situations and between mass nominals and atelic situations.

Activities were also expressed more by mass nominals than count nominals, at a rate of 16.5% compared to 8.5%. However, unlike the situation types expressed above, this difference generates questions in relation to the validity of the above relationship, as the difference in Activity expression by count and mass nominals is not as notable. As stated above, Activities would be expected to be expressed largely in mass form, due to their atelic nature. An example of a count Activity is shown below in example (37). Furthermore, in a similar fashion, the 0.7%

of States that were expressed by count nominals also poses food for thought, as States, like Activities, do not express a temporal endpoint. With these results, there is evidently a lot to unpick. Below I will first cover why State situation types are expressed more prevalently by mass nominals. Afterwards, using the results recorded on the relationship between count/mass status and situation type expression, Section 4.2.3 will investigate the behaviour of the relationship between count/mass status, boundedness, homogeneity and telicity in the nominal domain.

- (37) But their feelings for each other had been irreparably damaged by the brawls and fights and the recriminations afterwards. (Noun 998, All Data without Repetition Dataset, Appendix A)

As noted above, Figure 5-4 shows a large distinction in the number of States expressed by count and mass nominals. While count nominals only expressed States at a rate of 0.7%, mass nominals expressed States at a rate of 25%. From this distinction, mass nominals appear to share a much stronger association with state meaning than count nominals. This result is then made even more interesting when we look at Figure 5-5, where the influence of count/mass status on the situation type expression of nominals differs according to whether the nominals in question are abstract or concrete. On the left side of tree which concerns abstract nominals, count/mass status provides the first feature interaction with abstract status, causing the first data split on that side of the tree. On the concrete side of the tree, the influence of count/mass status is not as strong, as two word formation types (MDA; TC) hold a greater influence over the situation type expression of concrete nominals. But why is this difference apparent? I suggest that this difference is due to a greater expression of state meaning by abstract nominals.

In Section 5.1.2, I discussed how concrete nominals largely expressed Objects as a result of their physical nature and their primary instantiation within the domain of space. I also discussed how abstract nominals were more likely to express temporal meaning, as abstract nominals are often able to retain and express some of the relational and processual elements associated with the abstract phenomena they represent (Radden and Dirven 2007, p.79). This potential of abstract nominals was shown in the data, as abstract nominals expressed a far greater amount of event and state meaning than concrete nominals. Focusing on States in particular, the influence of count/mass status becomes more important in determining the situation type expression of nominals, as States are largely abstract entities which are homogeneous and



unbounded; typical qualities of mass status (Rothstein 2010, p.348). For instance, in example (38) below, the State of *courage* expresses no boundaries as it holds over time and does not express a temporal endpoint. Furthermore, in accordance with mass status, States can also be efficiently contracted due to their homogeneous nature. A portion of a State still remains a valid instance of the State as a whole (Langacker 2008a, p.141). For instance, a portion of *courage* is still valid instance of *courage*.

- (38) a. William Hill was a great bookmaker, who had the courage to take enormous bets.  
(Noun 3156, All Data without Repetition Dataset, Appendix A)

The strong influence of count/mass status on situation type expression by abstract nominals comes because abstract mass nominals express highly distinct situation type distributions through their increased orientation towards the expression of States. In fact, within this study, there was not a variable interaction that shared a stronger association with a given situation type (i.e., States) than the interaction between abstract status and mass status. Inevitably, this relationship caused the first data split within the decision tree structure in Figure 5-5. This finding is particularly interesting as it strongly supports the suggestion that mass status shares a strong relationship with State situation type due to its unbounded atelic nature. It is also interesting that this relationship appears to largely hold even when situation type is expressed in the nominal domain (however, as we will see in the following section through the presence of ‘count States’ in the data, this position may be an overgeneralisation). As the strength of this interaction and, consequently, the expression of States, is not replicated to the same degree by concrete nominals, we find that count/mass status does not hold the same influence on the right side of the tree. The next section will investigate notable instances within the data that appear to contradict the supposed relationship between Activity and State situation type and mass nominals; namely ‘count Activities’ and ‘count States’.

### **5.2.3. Count Activities and Count States**

In this chapter, I have mentioned the sentiments of Mourelatos (1978); Talmy (1978); Bach (1986); Jackendoff (1990; 1991); Brinton (1991; 1998), Kearns (2000); and many others that note the apparent relationship between count/mass status in nouns with the perfective/imperfective distinction expressed by verbs. As a result of this apparent relationship, it is suggested that there is a further correspondence between telic situations like

Accomplishments, Achievements and Semelfactives and count status (Bauer et al. 2013), as well as a correspondence between atelic situations, such as Activities and States, and mass status (Mourelatos 1978; Brinton 1995). In this section we will see that while the relationship between certain situation types and count/mass status proposed generally stands when situation types are expressed by nominals, it should not be interpreted as definite, as in one way or another, probably every noun has the potential to be construed as either count or mass (Langacker 2008a, p.142). First, I will examine the presence of count Activities within the data and how they are still able to express properties such as homogeneity and atelicity, despite their count status. In doing so, I will consider the distinction laid out by Barque et al. (2009) between ‘occurrence Activities’ and ‘routine (habit) Activities’. Afterwards, I will then focus on count States, looking specifically at the how count States are also able to express the properties of homogeneity and atelicity, irrespective of their count status. When discussing the behaviour of these States, I will contemplate the difference proposed by Grimm (2014, p.197) between ‘experiencer-state’ States and ‘stimulus’ States.

The presence of countable Activity nominals was an interesting finding, as the notion of count status is typically assumed to negate certain fundamentals of Activity meaning such as homogeneity, unboundedness and atelicity (Lieber 2004, p.136). Count status, when accompanied with temporal semantics, typically allows for the production of individualised instances of an event (Barque et al. 2009, p.170). These events are discrete and delimited, so are therefore generally bounded (Barque et al. 2009, p.170). This is shown in example (39) below. An *explosion* is a bounded entity that will only last for a period of time. If there were to be another *explosion*, this would constitute a separate *explosion* event, distinct from the previous *explosion* due to its bounded nature. In the verbal domain, bounded situations largely imply additional properties such as heterogeneity and frequently telicity (Huyghe 2011, p.113). However, for many of the count Activities located within this study, not all these properties were expressed.

- (39) The concentrated explosions were throwing lumps of dirt and large pieces of wood in all directions. (Noun 1614, All Data without Repetition Dataset, Appendix A)

In examples (40a-c) below, instances of count Activities are presented. These instances are fascinating as while count status does impose an effect of making the Activities discrete, i.e., bounded, in that one *adventure*, *attack* and *battle* may be different to another *adventure*, *attack*

and *battle*, these Activities remain atelic, as they do not express a temporal endpoint. An individual might start an *attack*, and then begin another. The second *attack* does not necessarily entail the end of the first. The presence of these count Activities foregrounds the behaviour of boundedness and telicity in the nominal domain. A count Activity nominal may be bounded, but this boundedness does not always presuppose that the nominal will also express a temporal endpoint, i.e., count nominals do not always describe a culmination (Huyghe 2011, p.113).

- (40) a. ...Miss Howard, formerly of Brighton and London, who had discreetly accompanied the Prince through his various adventures. (Noun 1018, All Data without Repetition Dataset, Appendix A)
- b. He accused Gen Aoun of using as "sandbags" the thousands of Christians who have gathered outside the palace to deter an attack. (Noun 1372, All Data without Repetition Dataset, Appendix A)
- c. ...his first command with that rank required him to reorganise the army of Weimar which had been smashed at the battle of Ettlingen. (Noun 872, All Data without Repetition Dataset, Appendix A)

Moreover, despite their count status, these Activity nominals remain contractable, a typical feature of mass nominals. A portion of an *adventure*, *attack* or *battle* is still a valid instance of that type. They are “qualitatively the same throughout” (Langacker 2008a, p.140). This homogeneity is contradictory to the majority of count nominals that cannot be contracted. To reiterate an example stated earlier in Section 5.2.1, *the boot of a car* is not the *car* itself, only an element of the *car*'s composition. In the verbal domain, homogeneity implies unboundedness (Huyghe 2011, p.113). As shown from the examples above, this correlation is not always the case in the nominal domain. Huyghe (2011, p.113), through his work into the expression of French deverbal count Activity nominals, claims that within the nominal domain, “homogeneity and boundedness are not contradictory”. Haas and Huyghe (2010), and scholars such as Barque et al. (2009), point towards the use of a two part distinction when classifying Activity nominals: the distinction between ‘habits’ and ‘occurrences’.

Habit Activities supposedly express ‘routine Activities’ that display an extended event structure (Barque et al. 2009). As shown by the example of *printing* in example (41) below,

habit Activities differ from occurrence Activities, such as *adventure*, *attack*, and *battle*, as habit Activities express ‘pure’ Activity nominals. They are homogenous, lasting and unbounded events (Heyd and Knittel 2009; Haas and Huyghe 2010). Their mass status is directly correspondent with the notion of atelicity in the verb domain (Haas and Huyghe 2010). Occurrence Activities, contrastingly, express “individualised instances” (Barque et al. 2009, p.170), indicated by their count status. For nominals such as *adventure*, *attack*, and *battle*, each will certainly have a final bound (an *attack* may not last forever). However, the existence of this final endpoint is not imposed by the essence of the event (Moltman 2019, p.356). As Haas and Huyghe (2010, p.111) elegantly state, “the final terminal is not fixed from the inside”. The event is homogenous and may stop at some point in time, but this time is not specified without additional contextual information.

- (41) ...the most valuable being in 1875 for the making of copper rollers for calico printing.  
(Noun 872, All Data without Repetition Dataset, Appendix A)

Barque et al. (2009), Haas and Huyghe (2010) and Huyghe (2011) appear to suggest that occurrence Activities are only expressed by events which express actions alone, such as *a walk*. Nevertheless, from the examples given in this section, such as *adventure*, *attack*, and *battle*, occurrence Activities are not just limited to actions, but also include events which are not action-related. It is also interesting that the distinction of habit vs occurrence in Activity nominals is founded on the semantic behaviour of French deverbal nominalisations. In this study, we see that this distinction is also applicable to English nominals, and that nominals such as *adventure* (which is a French borrowing) can also express occurrence Activities, despite not deriving from a verb. Furthermore, while Haas and Huyghe (2010) and Huyghe (2011) each discuss the overgeneralisation of the correspondence between homogeneity and boundedness in relation to Activities within the nominal domain, they do not mention the presence of this overgeneralisation in relation to other types of nominals which do not express a temporal endpoint – namely States. Moving forward now, I will show how certain States also provide evidence of the overgeneralised correspondence between boundedness and homogeneity in the nominal domain.

The presence of countable State nominals within the data represents an intriguing finding as, like Activities outlined above, the notion of count status is typically assumed to negate certain foundational principles of State meaning, such as unboundedness, homogeneity and atelicity

(Lieber 2004, p.136). Count states are shown in examples (42a-c). From these examples, we see again that, although count status does impose the effect of making the States discrete (i.e., bounded), in that one *hope*, *obsession* and *sympathy* may be different to another *hope*, *obsession* and *sympathy*, these States are still homogenous. Reflecting the ‘occurrence Activities’ seen above, the count States are contractable and are qualitatively the same throughout. A portion of an individual *hope* is still adequately representative of a *hope*. Furthermore, these count States do not express a temporal endpoint, but instead express the capacity to continuously hold parallel to one another, without an instance of change. An individual may have one *obsession*, and then gain another in the same time-period.

- (42) a. ...this can begin to do justice to the life and work of Paul Nizan, inextricably enmeshed as it is in the illusions, deceits, hopes, aspirations, successes and failures of its time. (Noun 4844, All Data without Repetition Dataset, Appendix A)
- b. ...his vainglory, his hectoring, his insensitivity, his military obsessions, have lost him all good will, at home and abroad. (Noun 3340, All Data without Repetition Dataset, Appendix A)
- c. ...in which the radical sympathies of a prosperous nonconformist community confronted the conservatism of wealthy merchants and professional men. (Noun 2395, All Data without Repetition Dataset, Appendix A)

All of the States shown across the examples in (42) above represent ‘Psych Nouns’, expressing psychological States (Fábregas and Marín 2012). Grimm (2014, p.197) suggests that these psychological States express the potential to be counted and that their count status is dependent upon whether the State expresses an “experiencer-state” or a “stimulus”, which is a view also largely echoed in Barque et al. (2011). Experiencer-states represent the emotional state experienced by a participant, such as *anger*. Stimuli, on the other hand, represent the elements which invoke the stative emotion (Grimm 2014, p.197). In such instances, Grimm (2014, p.197) claims that stimuli can potentially relate to objects or other temporal events under the guise of a State situation type. In the examples in (42), each of these States refer to stimuli. Taking example (42b) in particular, I could propose that one military *obsession* might relate to the *use of weapons*, while another might relate to *flying an aircraft*. Both events can be understood as stimuluses for the State of *obsession*. Nevertheless, the State situation type guise

that these objects and events take remain distinctively atelic, as each of the States intuitively expresses the potential to exist in unison of one another indefinitely.

The lack of temporal endpoints and homogeneity expressed by these count States reinforces the finding that the relationship between the concepts of boundedness and homogeneity operates differently within the nominal domain when compared to the verb domain (Huyghe 2011). While these concepts can certainly be applied in both verb and nominal domains, their relationship with one another is largely influenced by the lexical category of the lexical item they are applied to. When expressed in the nominal domain, the concepts are subject to the semantic constraints of nominality. The presence of count States builds on our understanding of the relationship between the concepts of boundedness and homogeneity in the nominal domain, as it shows that homogenous non-dynamic situations, e.g., States, in addition to homogenous dynamic situations, e.g., Activities, can also afford bounded interpretations.

In the next section, we will continue our investigation into the nature and degree of nominality, and how event, state and object meaning come to be expressed in nominal form, by exploring the relationship between word formation type and situation type expression by nominals. More specifically, I will investigate the influence of the following word formation types ‘MDA’, ‘MDV’ and ‘TC’.

### **5.3. Word Formation Type: The Influence of MDA; MDV and TC**

In this section, I will examine in greater detail the relationship between word formation type and situation type expression. As stated at the end of the previous paragraph, this next section will consider the situation type distributions expressed by three word formation types: MDA, MDV and TC. The section will begin by focusing on MDA and its apparent orientation towards State situation type through the derivational morphemes *-ness* and *-ity*. The section on MDA will then also briefly account for the capacity of deadjectival nominals to express events, despite being derived from an adjective base. After covering the semantic behaviour of MDA, the section will then centre on the semantic behaviour of MDV. In doing so, I will consider the presence of *-er* nominals in the data as a core reason for the greater amount of Object expression by MDV when compared to MDA and TC. Lastly, the section will focus on the semantic behaviour of TCs, and how TCs were able to express the greatest semantic variation of any word formation type in the data.

To discuss the individual differences in situation type expression by the three word formation types noted above, I will now present Figures 5-6 and 5-7. Figure 5-6 displays information on the proportion of situation types expressed by the different word formation types. Figure 5-7 (a figure that has recurrently been presented throughout this chapter) displays a decision tree structure of information on the simultaneous influence of word formation type, count/mass status, abstract/concrete status, and genre on the situation type expression of nominals. Both these figures are presented below. I will now begin the current section by directing our attention to the influence of MDA on situation type expression.

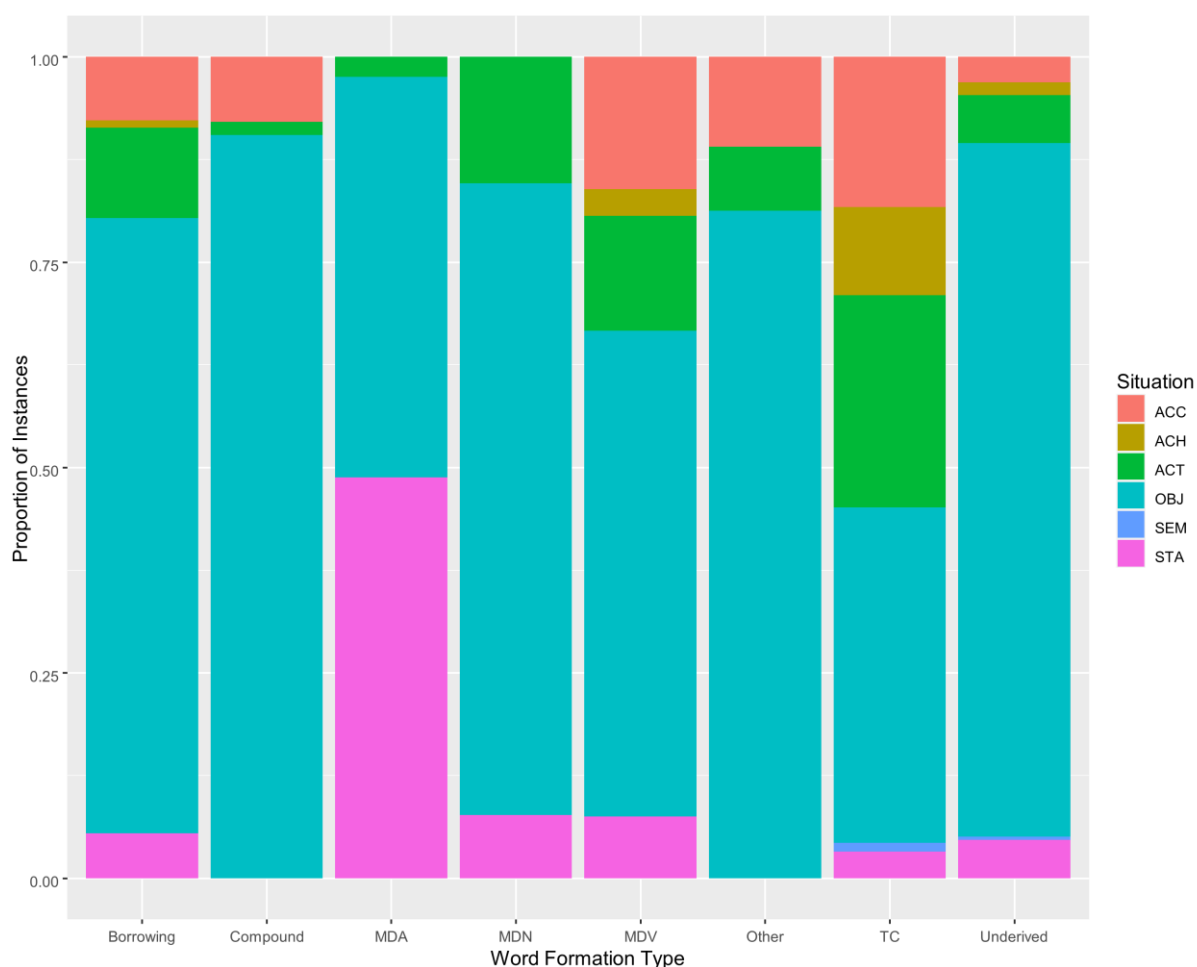


Figure 5-6. Barplot of situation type proportion expressed by word formation type.

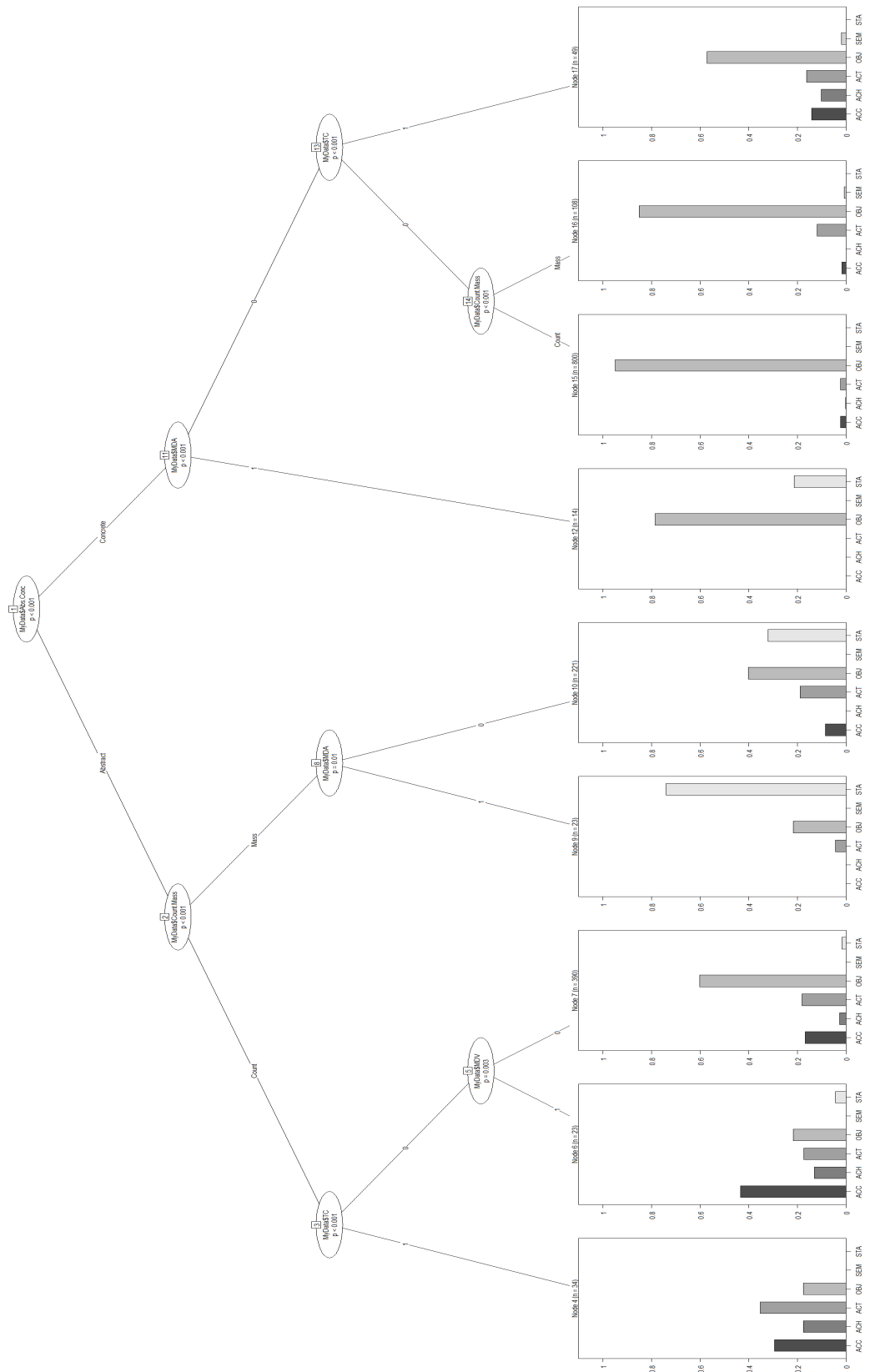


Figure 5-7. Decision tree visualising situation type ~ word formation type + count/mass status + abstract/concrete status + genre.



### 5.3.1. The Influence of MDA on Situation type Expression

From the Fisher-Exact test analyses carried out in Section 4.1.1, it is evident that MDAs express the most significantly different situation type distribution. As can be seen in both Figures 5-6 and 5-7, a large contributor to this significantly different distribution was the notably high number of States that were expressed by MDAs (48.8%). This orientation to State expression, while interesting, is not entirely surprising, given that MDAs are derived from adjectives, and adjectives primarily function to express either states or qualities. From the data, there were two derivational morphemes that were the driving force behind this high percentage of State expression: *-ness* and *-ity*. I will now examine the semantic influence of these morphemes in the following paragraphs.

#### 5.3.1.1. The Semantic Behaviour of *-ness* and *-ity* Suffixation in MDAs

The *-ness* suffix typically occurs in the deadjectival nominalisation process (Reichl 1982, p.158). Out of all the MDAs in the data, 29.2% involved *-ness*. In deadjectival nominalisation, the *-ness* suffix generally functions to describe the abstract state or property of the base adjective (Carstairs-McCarthy 2002, p.50). This is apparent in the data, as nominals formed with this suffix accounted for 45% of all States expressed by MDAs. Examples of these nominals are shown in (43a-b) below. In both examples, we see similar semantic behaviour influenced by the *-ness* suffix, as *happiness* expresses the abstract State of ‘being happy’, while *savouriness* expresses the abstract quality of ‘being savoury’.

- (43) a. It will be a great happiness to meet George MacDonald in eternity... (Noun 4683, All Data without Repetition Dataset, Appendix A)
- b. ...Of Earth teems with this joyous response; it fills one's nostrils with its fragrance; it delights one's palate with its savouriness. (Noun 650, All Data without Repetition Dataset, Appendix A)

In addition to the *-ness* suffix, the *-ity* suffix was also prevalent in the MDA word formation category; present in 19.5% of all MDAs. Like the *-ness* suffix noted above, the *-ity* suffix was highly associated with the expression of States, as nominals formed with the suffix accounted for 35% of all States expressed by MDAs. According to Bauer et al. (2013, p.257), the

deadjectival suffix *-ity* expresses a highly similar semantic influence to the *-ness* deadjectival suffix. Bauer et al. claim that, like the *-ness* suffix, nominals with the deadjectival suffix *-ity* can describe the abstract state or property of the base adjective. For instance, in examples (44a-b) below, the nominals *sensitivity* and *vulnerability* express the state of being *sensitive* and *vulnerable*, respectively. Accordingly, it is apparent that the high proportion of States expressed by MDAs is largely the result of the prevalence of the *-ness* and *-ity* derivational morphemes in the data and their semantic capacity to create nominals that describe the states and qualities that were originally expressed in their base adjectives. Following on, I will now explore how certain MDAs are able to express event readings, despite being derived from adjectives.

(44) a. The Discarded Image is a book which was written by a man with an unusual sensitivity to the differences between past and present. (Noun 1714, All Data without Repetition Dataset, Appendix A)

b. ...but perhaps what was needed was a slightly sharper, more bristly Dora with a hidden vulnerability, which came more naturally when I gave her a North London accent... (Noun 606, All Data without Repetition Dataset, Appendix A)

### 5.3.1.2. MDAs and the Expression of Events

MDAs, by their description, are nominals that are morphologically derived from adjectives. It was therefore somewhat expected that their semantic output would principally orient to the expression of States (due to the influence of their adjectival base, as shown in examples above such as *happiness*) and Objects, such as the nominals *pleasantry* and *technician*, shown in examples (45a-b) below. This expectation was largely reflected in the data. As shown in Section 4.1.1, MDAs expressed both States and Objects at an equally high rate of 48.8%, leaving 2.4% of MDAs left. This 2.4% was the result of one Activity MDA in the data, *business*, whose presence I will now discuss. The Activity MDA *business*, shown in example (46) below, is an especially interesting case, as it semantically behaves rather like a deverbal event nominal, producing a somewhat eventive structure, despite being derived from an adjective.

(45) a. As a parting pleasantry, I reported this to Mrs Brown, our Cardington Road billetor. (Noun 1274, All Data without Repetition Dataset, Appendix A)

b. RAF technician preparing for Red Arrow display by Stephen Ward Nigel Graham, an RAF technician, preparing a Red Arrow aerobatic display aircraft for a flight at RAF Scampton in Lincolnshire yesterday... (Noun 1959, All Data without Repetition Dataset, Appendix A)

- (46) You can't teach it to eat from a certain bowl in a certain place, or to do its business in a litter box. The whole idea is to get the bird to act naturally on your fist. (Noun 4026, All Data without Repetition Dataset, Appendix A)

It is important to point out that, in this example, the lexical item *business* operates as a euphemism that provides a less abrasive, indirect way of discussing a topic, i.e., the event of excretion, typically considered as 'taboo' (Allan and Burridge 1991). By exhibiting this semantically flexible behaviour, the nominal *business*, in this instance, can be defined as what Arche and Marín (2014, p.23) call an "occurrential deadjectival nominalisation". These nominals function in a semantically different way to more 'typical' deadjectival nominals that denote either State situation type or object meaning, as they can syntactically function with complements of event lexical items (Arche et al. 2021, p.23). For instance, *business*, in the sense shown above, can be 'carried out/done', 'be observed over a period of time', and can 'take place'. But why is this? According to Arche et al. (2021), this semantic behaviour is a property that is largely exclusive to particular deadjectival nominals that have been nominalised from an evaluative adjective base.

Evaluative adjectives largely function to ascribe characteristic attributes to the mind, character or behaviour of an individual (Arche and Stowell 2019). Certain evaluative adjectives express the ability to predicate sentient individuals and event subjects (Stowell 1991:106). For example, we can say *John was busy* and we can also say that *the conference that lasted two days was busy*. Following on from work concerning the semantics of evaluative adjectives by Stowell (1991), Arche (2006) and Fábregas et al. (2013), Arche et al. (2021) claim that, as certain evaluative adjectives can predicate both a sentient individual and an event, they include a covert process description within their semantic structure. Roy (2010) proposes that deadjectival nominalisations express the capacity to contain a predicative phrase. Arche et al. (2021, p.33) subsequently suggest that eventive deadjectival nominals are distinguished as they are able to embed predicative phrases whose subject is a dynamic event expressing "no time or

world parameters”. As a result of embedding this predicative phrase, the eventive deadjectival nominalisations are said to inherit the event properties of the predicative phrase they embed, which allows for the expression of occurrential events by deadjectival nominals (Arche et al. 2021, p.33).

*Business* can express an event or a series of events in which the participant(s) are understood to be *busy*. This semantic behaviour is an interesting, as while other event deadjectival nominalisations, such as those listed in Arche et al. (2021, p.10) (e.g., *cruelty* in *the cruelty lasted for months*), appear to provide an ‘evaluative stance’ (Hunston and Sinclair 1999, pp.91–101) on the events taking place, the deadjectival nominalisation *business* does not. In the specific instance in example (46), the actor (i.e., the bird) is *busy* when it undertakes the act of excretion. The act of excretion, on the other hand, cannot be described as *busy*. By employing this deadjectival nominalisation, the writer focuses on the *busy* nature of the bird during the act of excretion, allowing the writer to still express the temporal semantics associated with the act, while not expressing the taboo directly.

Therefore, contrary to the popular association between nominality and atemporal semantic structure, we can see here that the nature of nominality actively supports the expression of event semantics by nominals derived from evaluative adjectives. The deadjectival nominalisation process is evidently equipped with the quality to prize out and express the covert process descriptions in the semantic structures of certain adjectival bases when these deadjectival nominalisations are used in the appropriate linguistic contexts which support eventive readings. This discussion reinforces the suggestion from Huyghe et al. (2017) that lexical aspect does not share associations with lexical class for its expression. It is instead more dependent upon the expression of an eventuality, regardless of the lexical class that it is expressed by.

### **5.3.2. The Influence of MDV on Situation type Expression**

The Fisher-Exact test conducted in Section 4.1.1 and the CRT analysis conducted in Section 4.2 showed that MDV nominals expressed the third most distinguished situation type distribution of the word formation types. The situation type distribution of MDV nominals is available in Figures 5-6 and 5-7. Looking at the results of the CRT analysis specifically (see Figure 5-7), it is notable that while MDV word formation type was significantly influential on

the situation type distributions of abstract count nominals, it was not significantly influential in the situation type distribution of their concrete counterparts. In the following paragraphs I will examine the situation type distributions of abstract and concrete count MDVs, before then focusing specifically on the reason why concrete count MDVs did not register a significantly different situation type distribution.

The comparison of abstract count MDVs (shown in terminal node 10, Figure 5-7) with concrete count MDVs presents an interesting case of nominal semantics as, while abstract count MDVs present a rather diverse situation type distribution (note a greater orientation to Accomplishments), concrete count MDVs express a very strong orientation to Objects. Abstract mass nominals expressed Accomplishments at a frequency of 43.5%, Achievements 13%, Activities 17.4%, Objects 21.7%, and States 4.3%. In this situation type distribution, the expression of both Accomplishments and Achievements is rather high when compared to their relative proportions in the other eight terminal nodes in Figure 5-7. To give some examples, in (47a) below, the abstract count MDV *removal* expresses an Accomplishment as the *removal* of *a relatively dominant or low-ranking female* requires duration, but also expresses a temporal endpoint, i.e., when the entity has been *removed*. In (47b), on the other hand, *discovery* expresses an Achievement, as *discovering* an entity is a punctual event that ends as soon as the entity has been *discovered*.

- (47) a. ...because no consistent increase in aggression followed the removal of either a relatively dominant or low-ranking female. (Noun 2579, All Data without Repetition Dataset, Appendix A)
- b. ...but he had a sense of humour (“I am the Wallerian degeneration” in reference to his father’s eponymous discovery) and he retained an unworldly delight in science throughout his life. (Noun 2668, All Data without Repetition Dataset, Appendix A)

Concrete count nominals, in contrast, did not express such a diversity of situation type expression. In fact, concrete count MDVs displayed a very strong orientation towards Objects, at a rate of 91.8%. This exceedingly high proportion of Objects presents a unique case, as the bulk of this percentage (71.1%) consisted of nominals that expressed *-er* suffixation, while a further 4.4% included nominals that expressed its variant suffix *-or*. In the following paragraphs I will therefore focus on the dominance of *-er* suffixation in concrete count MDV

nominals, and also the more nuanced modal and aspectual semantics these nominals were able to express.

### 5.3.2.1. The Dominance of *-er* Suffixation in Concrete Count MDVs

There is certainly no shortage of literature on the behaviour of deverbal *-er* nominals, both from a descriptive standpoint, e.g., Bauer et al. (2013), Andreou (2021), and from a theoretical standpoint (Levin and Rappaport Hovav 1988; Rappaport Hovav and Levin 1992; Heyvaert 2003; 2010; Lieber 2004; 2016; Alexiadou and Schäfer 2010; Cohen 2016; Lieber and Andreou 2018 and many others). Due to the scope of this thesis, I will not try to provide an explanation of exactly how deverbal *-er* nominals are able to express the semantic meanings they do. I will instead only examine what semantic nuances the deverbal *-er* nominals in the concrete count MDV category expressed in this study. However, it is intriguing that the findings from this study do correspond with the proposals from Lieber and Andreou (2018) that deverbal *-er* nominals can suggest both modal meanings and instantiations of aspectual meaning. I will now cover this capacity of deverbal *-er* nominals in the following paragraphs.

The *-er* suffix is regarded as a productive bound derivational morpheme for forming nominals from verbal bases (Booij 2005, p.18). Most typically, it is associated with creating agents and instruments (Bauer et al. 2013, p.220), hence the association with object meaning. For instance, in example (48a), a *hunter* is a sentient individual that *hunts* another entity. But what is interesting about these types of nominals is that, although they express Objects, they also express a form of temporal meaning that is not effectively accounted for using the diagnostic syntactic tests (DSTs) in this study: namely modal readings and grammatical aspect readings, such as ‘bounded’, ‘unbounded’ and ‘habitual’ (Lieber and Andreou 2018).

- (48) a. ...they sailed away bearing his new widow and her son, aged two, who grew up to become a big game hunter in central Africa. (Noun 766, All Data without Repetition Dataset, Appendix A)

Cohen (2016) claims that the semantics of deverbal *-er* nominals predominantly express properties that are intrinsic to the subject. Following this line of thought, she argues that deverbal *-er* nominals operate in a highly similar fashion to that of dynamic modals, i.e., subject oriented modals that typically express abilities and dispositions, e.g., ‘can’ (Cohen 2016, p.96). Cohen suggests that if we take the core semantics of the deverbal *-er* suffix as a dynamic modal,

then the semantics expressed by the suffix will provide an account as to why *-er* nominals tend to express the external argument of their base verb (Lieber and Andreou 2018, p.191). For instance, example (49) below displays what could be described as a dynamic modal *-er* nominal, as a *water maker* is an object which expresses properties required to engage in *making water*. There is no instantiation of *water making* implied in this dynamic modal *-er* nominal however, as a *water maker* can adequately be a *water maker* without ever having *made* water. Cohen's (2016) theoretical interpretation of deverbal *-er* nominals certainly sheds light upon the semantics of certain deverbal *-er* nominals. However, as we shall see in examples (50a-c) below, it fails to cover the full variety of semantic readings that can also be expressed by these nominals (Lieber and Andreou 2018; Andreou 2021). In the current study, only five of the deverbal *-er/-or* nominals expressed a dynamic modal reading. In line with the work of Lieber and Andreou (2018) the majority of deverbal *-er* nominals in this study suggested the instantiation of grammatical aspect readings instead. Either, 'bounded', 'unbounded' or 'habitual'.

- (49) ...Reckmann furling system, generator, 140 litre per hour water maker, air compressor for driving bottles... (Noun 4180, All Data without Repetition Dataset, Appendix A)

Lieber and Andreou (2018) claim that deverbal *-er* nominals can express both modal and aspectual readings. Looking at grammatically aspectual readings specifically, they argue that the participants expressed by deverbal *-er* nominals suggest their involvement in either 'bounded', 'unbounded' or 'habitual' instantiations. Each of these types of deverbal *-er* nominals were present in the data. To illustrate this point, the examples in (50a-c) below suggest participant involvement in bounded (50a), unbounded (50b) and habitual (50c) instantiations. Starting with (50a), the deverbal *-er* nominal *writers* relates to two *writers* who previously wrote two shows 'The Boy Friend' and 'Salad Days'. The *writing* events suggested by the deverbal *-er* nominal represent bounded events, as once the shows are *written*, the events are terminated. Alternatively, in (50b) we witness the participant's involvement in an unbounded instantiation of *admiration* through the deverbal *-er* nominal *admirer*. In this instance, the status of an *admirer* holds as the participant's *admiring* is continuous and has not ended. Lastly, we have the expression of habituality (shown in (50c)), which was by far the most common of the three aspectual meanings expressed by deverbal *-er* nominals in the data. Here, the deverbal *-er* nominal *teachers* suggests the participants' involvement in recurrent events of *teaching*. Each one of these recurrent events may be bounded, but together they form

a continuous series of *teaching* events that form a ‘habit’ for the participants (Lieber and Andreou 2018, p.198).

- (50) a. ...bravely supposedly using the names, and the names only, of Julian Slade and Sandy Wilson (writers of *The Boy Friend* and *Salad Days*)... (Noun 36, All Data without Repetition Dataset, Appendix A)
- b. Hartley Coleridge was one of his devoted friends and admirers and said “of all landscape painters he was the most literal, the most absolute copyist of the objects on his retina. (Noun 2360, All Data without Repetition Dataset, Appendix A)
- c. ...George Lyttelton, whose published letters to Rupert Hart-Davis have given pleasure to many, were inspired teachers. (Noun 464, All Data without Repetition Dataset, Appendix A)

Accordingly, as stated earlier in this section, the findings from this study support the proposals laid out by Lieber and Andreou (2018), as the deverbal *-er* nominals exhibited modal meaning, and instantiations of grammatical aspect readings (with habitual being the most expressed). The productive nature of the *-er* suffix for MDVs in this study resulted in a situation type distribution for concrete count MDVs that was not significantly different, due to the tendency of deverbal nominals with *-er* suffixation to express Objects. However, as seen in the paragraphs above, these Object nominals present a distinctive case of nominal semantics, as while they express Objects, they are simultaneously able to suggest either modal meanings or the instantiation of grammatically aspectual meanings linked to their base verb. It is here that I will leave the semantic behaviour of MDVs and move on to cover the final word formation type in this chapter: TC. The section will first discuss situation type distribution expressed by TCs, and how TCs were able to express the greatest semantic variation out of the eight word formation types in the data.

### **5.3.3. The Semantic Variation of TC Nominals**

Transcategorisations which produce nouns are interesting, as it is suggested by Davies (2004, p.85) that transcategorisation away from nouns (e.g., from noun to verb) is a more productive word formation process. Despite this lower productivity, Halliday (1966) has suggested that



English has become nominally skewed, whereby meaning is expressed more dominantly through nominals than verbs. This claim is supported by the recent work of Fontaine (2017a), whose work on grammatical boundaries identified a preference for nominal expression over other lexical classes. Early research by Biese (1941) into the nature of nominal transcategorisation claimed that nominal TCs orient towards expressing definite beginnings and endings, and, as a result, tend to take count status. However, in opposition to Biese's work, Plag (1999, p.324) argues that transcategorisation presents the most semantically versatile nominalisation process. According to Valera (2015, p.327), TCs can express an array of different semantic meanings, as they are flexible and adapt themselves to the context they are used in. Moreover, Andreou and Lieber (2020) show that TCs are versatile lexical items that are susceptible to event, state and object expression, and receptive to both count and mass status.

This semantic versatility is evident in Figure 5-6, as TCs displayed a diverse situation type distribution of 18.3% Accomplishments, 10.8% Achievements, 25.8% Activities, 40.9% Objects, 1.1% Semelfactives and 3.2% States. Perhaps this versatility is somewhat unsurprising, given that TC encompasses all lexical class shifts towards noun, e.g., verb to noun, adjective to noun, etc.). Nevertheless, it is still interesting as, out of all the word formation types analysed within the Fisher-Exact tests (displayed in Section 4.1.1), TC recorded the second most significantly different situation type distribution. Furthermore, in the CRT analysis (see Section 4.2 and Figure 5-7), TC was involved in two significant splits in the decision tree structure, which were due to interactions between (1) abstract and count status with TC and (2) concrete status with TC. In the following paragraphs I will now explore the significantly different situation type distributions expressed by TCs in the CRT analysis, with a principle focus on the capacity of TCs to express a high variation of semantic readings.

### **5.3.3.1. The Situation Type Expression of Abstract Count TCs**

The situation type distribution expressed by abstract count TCs consisted of 29.4% Accomplishments, 17.6% Achievements, 35.3% Activities and 17.6% Objects. Intriguingly, this interaction did not express any States, although, as we saw in Section 5.2.2, this outcome is most likely the result of the typical association between States and mass status. Focusing now on the four situation types expressed by abstract count TCs, examples of each of these situation types are available in examples (51a-d) below. In (51a) we see the Achievement TC *split*. The nominal *split*, in this instance, shares and expresses the typical temporal semantics

of its verbal counterpart. This correlation between the semantics of the nominal and verbal form of *split* suggests that the lexical aspect of the base verb does hold some influence on its transcategorised nominal form (Andreou and Lieber 2020, p.356). A *split* is generally a punctual action which, due to its instantaneous nature, expresses a temporal endpoint. In (20a) *split* is used metaphorically to express that the UK indie band ‘The Smiths’ (whose members include Johnny Marr and Morrissey), would no longer create and play music together. Although, this event had been ‘brewing’ for an amount of time, the TC nominal *split* expresses only the culmination of this event.

Moving onto (51b), we have the Activity TC *endeavours*. Once again, this TC nominal does share a relationship with the typical situation type of its verbal counterpart, as to *endeavour* relates to exerting oneself and using effort (OED 2020d). In its nominal form, these durative and atelic properties are similarly expressed. Interestingly, the TC nominal *endeavours* relates to our discussion on count Activities in Section 5.2.3, as the parameters of durative and atelic properties operate differently for count Activity nominals. These *endeavours* represent occurrential atelic events that are also homogenous but bounded, i.e., discrete: a relationship unique to the nominal domain (Huyghe 2011, p.113).

In (51c) we have the Accomplishment *trip*, which represents an interesting case, as the nominal expresses an Accomplishment, despite its verbal counterpart rarely being used to express a sense of ‘travel’ (Hayriyan 2020, p.3). The nominal form expresses a different semantic meaning to its verbal counterpart, despite remaining morphologically the same. Such an instance points towards the supposed semantic flexibility of TCs (Valera 2015) and their capacity to be influenced by their surrounding linguistic context. In this instance, the expression of Accomplishment situation type is not only the product of the semantics of *trip*, but also of the subsequent complement in the clause *to New Zealand*. In this instance, *trip* expresses an Accomplishment, as it represents a durative event that will reach its final endpoint when the individual reaches New Zealand, or perhaps when the individual returns from New Zealand. Lastly, we have the Object TC *result* in (51d). Like *trip*, *result* also expresses a different lexical aspect to its verbal counterpart. While the verbal form of *result* generally expresses the meaning “to arise as a consequence, effect, or outcome of some action, process, or design” (OED 2020j), its nominal form instead focuses more upon the product of these arising process (Martsa 2013, p.183), hence its atemporal semantic structure. *Result*, in this case, is an entity brought into existence through the subsequent process of *resulting* (Bauer et al. 2013, p.212).

- (51) a. Now all this hinted, I'm sure you'll agree, that the split had been brewing, indeed planned for some time. Both Marr and Morrissey were instantly involved... (Noun 1919, All Data without Repetition Dataset, Appendix A)
- b. A lifetime's endeavours brought rich rewards (a personal estate of −£700,000 at his death) and gained him the admiration of Samuel Smiles... (Noun 2834, All Data without Repetition Dataset, Appendix A)
- c. ...but later it was decided that I could combine business and health by making a trip to New Zealand... (Noun 3565, All Data without Repetition Dataset, Appendix A)
- d. ...and we shall be under no liability for any extra costs incurred by such a person as a result of our doing so. (Noun 1139 All Data without Repetition Dataset, Appendix A)

### 5.3.3.2. The Situation Type Expression of Concrete TCs

Unlike abstract count TCs presented in the previous section, concrete TCs displayed a greater orientation to object meaning, expressing Accomplishments at a frequency of 14.3%, Achievements 10.2%, Activities 16.3%, Objects 57.1%, Semelfactives 0.2% and States 0%. This greater occurrence of Objects follows a wider pattern shown by concrete nominals that has been discussed in Section 5.1. Nonetheless, it is still interesting how these TC nominals express atemporal semantic structure - especially those that are transcategorised from verbs and express no overt derivational morphology.

Taking a closer look at these concrete TCs that express object meaning, two different types of semantic behaviour were largely at play. The first (which was the most prevalent) is displayed by the nominal *crowds*, located in example (52a). Nominals such as this express an entity that has been created as a result of the process of its verbal counterpart (Bauer et al. 2013, p.210). A *crowd* forms through the process of entities *crowding*. Alternatively, in example (52b), we see a different type of semantic behaviour that creates instruments. The concrete TC nominal *covers* relates to the action of an instrument that is generally used to perform that very action (Martsa 2013, p.183). For instance, a *cover* will typically be used to *cover* other entities.

(52) a. I wandered among turbanned headstones in the Muslim cemetery, mingled with the crowds on the Galata Bridge... (Noun 470, All Data without Repetition Dataset, Appendix A)

b. These slim volumes," Robert Hewison has observed, "with pastel covers and fragile wartime paper, became emblems of survival which no Ministry of Information publication or film... (Noun 4732, All Data without Repetition Dataset, Appendix A)

As noted in the above paragraphs, concrete TCs also expressed events. These events largely expressed physical actions, but ones which varied in relation to their duration and telicity. For instance, in (53a) below, we see the Accomplishment *walk*. The concrete TC *walk* displays a typical semantic behaviour for event TCs, as it expresses an action for an event involving that action (Martsa 2013, p.183). Intriguingly, the nominal *walk* would typically be considered an Activity, due to its durative and atelic nature (Bloch-Trojnar 2013, p.183). However, we can see from example (53a) that *walk* has been subject to aspectual coercion (see Section 2.4.1), as the complement *from London to Marlborough* provides the event with a temporal endpoint and coerces the Activity event of walking into an Accomplishment, i.e., the walk will end in Marlborough. Cases of aspectual coercion provide empirical evidence to the claims by Plag (1999), Valera (2015) and Andreou and Lieber (2020), that TC nominals are receptive to a high amount of semantic versatility.

In addition to the concrete Accomplishment TC *walk*, concrete TCs also expressed Achievements and Activities, examples of which are shown in (53b-c) below. Looking at the Achievement TC first, *crash* presents a dynamic, punctual, and telic event, as *crashes*, especially if physical, tend to be instantaneous. The concrete TC *marches*, on the other hand, does not present an instantaneous event, as *marches* generally represent dynamic, durative, and atelic events, i.e., Activities. For both *crash* and *marches* shown in (53b-c), the typical lexical aspect of their verbal counterparts is expressed in their nominal form. This semantic behaviour suggests that the lexical aspect of verbal counterparts does hold an amount of influence over the semantics of transcategorised nominals (Andreou and Lieber 2020, p.356). However, as shown in this chapter, this influence from the base verb does appear to vary depending on the nominal in question and its surrounding linguistic context.

- (53) a. ...Edward celebrated his release from St. Paul's with a three-day walk from London to Marlborough... (Noun 1801, All Data without Repetition Dataset, Appendix A)
- b. Representative Robert Tallon of South Carolina and 10 other congressmen have requested an investigation into the crash. (Noun 1943, All Data without Repetition Dataset, Appendix A)
- c. Patrick, the lifelong bachelor whose main interests are astronomy, cricket and military marches, next to Zsa Zsa Gabor! (Noun 4042, All Data without Repetition Dataset, Appendix A)

#### **5.4. Summary: Events, States, Objects and Evaluating Nominality**

Of all the predictor variables present in the study, abstract/concrete status held the greatest influence over the semantics expressed by nominals, a finding I have attributed to the enhanced capacity of concrete nominals to express object meaning. In Section 5.1.1, it was suggested that this enhanced capacity was associated with the temporal restriction conceptually imposed on certain entities through their concrete status and their primary instantiation within the domain of space. In addition to this finding, abstract nominals expressed a far greater tendency to express temporal meaning. A stronger orientation to temporal meaning was somewhat expected from abstract nominals, given that certain abstract nominals are supposedly able to preserve elements of their “relational and often processual character” during the reification process (Radden and Dirven 2007, p.79).

Abstract nominals expressed a high degree of State situation type, when compared with concrete nominals. The interaction between abstract status and mass status resulted in the strongest association with a given situation type (i.e., States) in study one. As a result of this variable interaction, count/mass status held a greater influence over the situation type expression of abstract nominals than concrete nominals. Moreover, word formation type was also influential in the expression of abstract and concrete nominals. For instance, it was identified that abstract mass nominals derived from adjectives (MDAs) were highly likely to express State situation type. This behaviour was largely the result of derivational morphemes such as *-ness* and *-ity*, which functioned to reify the abstract state or property of the nominal's base adjective.

Accordingly, the finding that abstract/concrete status holds the greatest influence over nominal semantics provides empirical support to the somewhat intuitive claim that concrete status is a hallmark of prototypical nominality (Hopper and Thompson 1984a; Langacker 1991). However, while abstract/concrete status was found to hold the greatest influence over the semantics of nominals, this study has shown that further distinct distributions in nominal semantic behaviour are observable when we consider additional variables and the order in which they interact with one another. Such a finding builds upon current continuum-based conceptions of nominal category structure by foregrounding the need to consider how interactions between lexicogrammatical, etymological, genre and various other linguistic variables can influence the semantic output of nominals.

Count/mass status was also explored in relation to its supposed association with the notion of telicity (Mourelatos 1978; Brinton 1995; Bauer et al. 2013). For Accomplishments, Achievements and States, the assumption that telic situations would correspond with count status and atelic situations would correspond with mass status generally held. This behaviour was not reflected by Activities however, which did not express a notable orientation to either count or mass status. By further investigating the semantic behaviour of Activities, a fundamental difference was identified in relation to how the concepts of bounding, homogeneity and telicity function within the verb domain and the nominal domain. Within the nominal domain, the notion of boundedness did not presuppose either the homogeneity or telicity of situations. This behaviour was clear in nominals such as *battle* (see Section 5.2.3) which are bounded but also express homogeneity and atelicity. Unlike the verb domain, bounded situations do not always describe a culmination when realised through nominals, nor do they always express heterogeneity (Huyghe 2011, p.113). As well as Activities, a lack of correspondence between the notions of bounding, homogeneity and telicity was also expressed by certain State situation types, such as *obsessions* (see Section 5.2.3). This finding is especially interesting with regards to the category of nominality, as it shows that while concepts such as bounding, homogeneity and telicity are transferable from the verb domain to the nominal domain, these concepts are subject to different semantic constraints presented by the nature of nominality. The difference in these semantic constraints across lexical categories supports the idea that although both nouns and verbs can express similar semantic behaviour, the fundamental notions of noun and verb should still be treated as distinct lexical categories.

After assessing the relationship between bounding, homogeneity and telicity within the nominal domain, I then moved onto the relationship between word formation type and situation type expression. From the post-hoc Fisher-Exact tests and the CRT analysis, it was clear that three word formation types held the largest influence over situation type expression: MDA; MDV; and TC. MDA provided the most distinct distribution, expressing mostly States. As mentioned earlier in this section, this high amount of State expression was attributed to the semantic capacity of derivational morphemes such as *-ness* and *-ity*, that reified the abstract state or property of the nominal's base adjective. Like MDAs, specific derivational morphemes were also influential in the semantics of MDV nominals. For MDVs, no derivational morpheme was more influential than the *-er* suffix. The *-er* suffix was pervasive amongst MDVs and expressed a strong orientation towards Objects, making the overall situation type distribution expressed by MDVs less distinct. However, while these deverbal *-er* nominals largely expressed Objects, many still suggested either modal or grammatical aspect readings. Deverbal *-er* nominals that suggested modal meaning often expressed the capacity of an Object to carry out an event (e.g., *water maker*), while deverbal *-er* nominals that suggested aspectual meaning typically expressed agent Objects that had either been involved in bounded events, unbounded events, or habitual events (e.g., *teacher*).

Lastly, exploring the relationship between word formation type and situation type proved insightful as certain nominal instances displayed the ability to express event meaning, despite not deriving from a verb. A particularly interesting case of this behaviour was shown by the MDA nominal *business*, derived from the adjective *busy*, which was used metaphorically to express an action that can 'be done'. I argued that, while adjectives such as *busy* do not express events, they do hold a 'covert event structure'. When these adjectives are nominalised, they are then able to enter appropriate linguistic contexts which support event readings. Accordingly, what event nominal instances such as *business* show is that nominality is not only supportive of event meaning but can be actively instrumental in the expression of temporal semantics for certain lexical items. Through nominalising lexical items such as *busy*, it allows their nominalised forms to enter syntactic contexts that facilitate the expression of temporal semantics that either cannot be construed or are not typically construed in their original lexical class.

This part in the thesis signals the end of study one. Study one has allowed us to take a broad look at the nature of nominality and observe how various variables interact with one another in

the expression of event, state and object meaning. The next chapter will mark the start of study two. While in study one I examined the inner workings of nominality, study two will take a more micro focus through examining how temporal meaning is expressed by four underived event nominals (UENs), and the extent to which the semantic behaviour of these nominals is supported by surrounding syntactic context. The next chapter outlines the methodology employed in this second study, along with the results and a short discussion on the influence definiteness and process type (defined and explained in Chapter 6) have on the semantic behaviour of four UENs.



## Chapter 6: Underived Nominals: Methodology and Results

This chapter marks the start of the second study of this thesis: *The Influence of Definiteness and Process Type on Temporal Semantics: The Case of Game; Fire; Stream and Drought*. Throughout the last three chapters, two thesis sub-aims (noted in at the beginning of Chapter 5) were addressed. Over the next two chapters, a further sub-aim (shown in bold font below) will be additionally considered, which relates to the exploration of potential relationships between the syntactic and semantic properties of underived event nominals (UENs). To examine this additional sub-aim, a narrower approach and, consequently, a second study was required to investigate how surrounding syntactic context may impact the temporal semantic behaviour of four underived event nominals (UENs). Specifically, the second study will provide greater information in relation to the extent to which definiteness and process type (see Section 6.2 for full definition) hold influence over the temporal semantics expressed across four UENs. The three sub-aims of this thesis are presented below, and, in the following paragraph, I will describe the structure of the chapter.

Sub-aims of thesis:

1. To determine how state and event meaning come to be expressed in nominal forms.
- 2. To examine the potential relationships between the syntactic and semantic properties of underived event nominals (UENs).**
3. To determine how the nature and degree of nominality can be evaluated in English.

In Section 6.1, I will detail the corpus and samples used for the second study, along with a description of the four selected UENs. I will then outline three distinct annotation procedures used to analyse the nominals in Section 6.2, including the annotation of definiteness, process type and lexical aspect. Section 6.3 will then describe the statistical analyses implemented in the current study and present the key methodological factors that motivated this decision. Moreover, Section 6.3 will present the hypotheses that were formulated for the second study. In Section 6.4, I will report and visualise the results of the statistical analyses, focusing on each of the four selected UENs in turn. The presentation of these results for each UEN will be accompanied with brief discussion on the extent to which definiteness and process type hold influence over the expression of lexical aspect.

## 6.1. Corpus and Sample

The corpus chosen for the second study is the Timestamped JSI Web Corpus 2014-2021 English (TJWCE). The TJWCE is part of a family of corpora created from IJS newfeed by Jozef Stefan Institute, Slovenia (Trampus et al. 2004). The corpus is extensive, amassing a total of more than 60 billion lexical items from the time period of 2014-2021. The corpus is a clean, real-time aggregated stream of news articles from RSS-enabled websites across the world. Moreover, the corpus is automatically tagged for part-of-speech (POS) information and timestamps have been used to augment the corpus with diachronic annotation (SketchEngine 2017).

The TJWCE was chosen because of several key factors. The first is that, as discussed above, the corpus presents a readily usable, clean corpus that is available through the corpus manager and text analysis software SketchEngine (Kilgarriff et al. 2004). The corpus is also continuously updated with new entries, providing a stream of highly current language use available to access. The large size of the corpus is also advantageous, as it provides a reliable foundation of data from which to explore the relationship between definiteness, process type and the distribution of lexical aspect expressed across specifically selected UENs (the four selected UENs are outlined in the paragraphs below). Additionally, the data within the corpus is taken largely from news articles from RSS enabled websites across the world. The decision to use data from these news articles was a motivated decision based upon the findings from the first study (see Section 4.1.4) that the genre of Broadsheet Newspaper Home and Foreign News Writing exhibited the largest amount of temporal semantics. I anticipated that this association between news article writing and temporal semantics might result in a higher expression of temporal semantics across the four nominals.

Thirty-six UENs were identified in the first study of the thesis (see Chapters 3-5). These nominals were ordered in relation to the number of lemmatised instances present within the TJWCE. From this order, two higher frequency nominals and two lower frequency nominals were selected at random for analysis. The four UENs chosen were: (1) *game*; (2) *fire*; (3) *stream*; and (4) *drought*. A table outlining these nominals and their respective frequencies within the TJWCE is available below in Table 6-1. UENs present an interesting phenomenon within the English language; they can express prototypical verbal semantics whilst also

expressing prototypical nominal form. As I have discussed at many points throughout Chapter 2, the expression of temporal semantics by morphologically underived nominals presents sufficient challenges to prototype models of lexical class that promote continua-based classifications, as their prototypical nominal form somewhat clashes with their less typical semantic behaviour, which leaves them without an explicit ‘home’ in the continuum of nominality. Accordingly, UENs occupy a unique space in our conception of nominality that requires deeper exploration.

*Table 6-1. Frequency of occurrence for four underived nominals across the Timestamped JSI Web Corpus 2014-2021 English.*

Nominal	Number of Corpus Instances
Game	64,497,533
Fire	14,534,363
Stream	2,446,813
Drought	1,103,144

Like the British National Corpus that was used earlier in the thesis, the TJWCE is lemmatised and tagged automatically for part-of-speech (POS) information. Using POS tagging, millions of nominal instances were identified for the four UENs. A random sample for each nominal was then generated and downloaded using the respective functions included in SketchEngine. A large number of nominal instances was downloaded to collect a sample size of 500 lexical nominal instances for each UEN, after reductions were made (see Section 6.1.1). The decision to analyse 500 instances was formulated on the basis of Hanks’ (2004, p.91) claim that sample sizes between 200-1000 concordance lines will provide a representative sample of a lexical item’s use.

### **6.1.1. Reductions across the Four Samples**

After the nominal instances for each sample were collected, the current study followed Balvet et al. (2011, p.1), by only considering lexical nominal instances that functioned as the syntactic head of the nominal group, as otherwise the nominal (e.g., nominal instances functioning as classifiers or possessives) could not be accurately annotated through the use of diagnostic syntactic tests (DSTs). Moreover, fixed idiomatic expressions such as the expression shown in (54) below were discounted from each sample, as the nominal instances in such expressions do

not allow for modification of the nominal group due to their fixedness (Moon 1997, p.44). As a result, they could not be accurately used with the DSTs. In addition, POS errors were also removed from each sample.

(54) Jack didn't want to stoke the fire anymore.

Once the nominals were collected and the inapplicable instances had been removed, 500 nominal instances for each UEN remained, equating to an overall sample size of 2000. These nominal instances were analysed for their definiteness, the process type of the clause they functioned in and their expression of lexical aspect. In the following section, I will describe the annotations of these variables in detail and provide reason as to why definiteness and process type were chosen for exploration. Afterwards, we will quickly recap the concept of lexical aspect and the diagnostic syntactic tests (DSTs) used to discern the different situation types expressed across the selected UENs.

## **6.2. The Annotation of the Four Underived Nominals**

As mentioned above, the nominal instances were annotated in relation to three variables: (1) definiteness; (2) process type; and (3) lexical aspect. In the following sections, I will provide background to each of these variables, whilst also specifying their relevance to the study.

### **6.2.1. Definiteness Annotation**

The first annotation conducted on the nominal instances was for their status as definite or indefinite. The notion of definiteness relates to the representation of an entity by a speaker as either identifiable or unidentifiable to a recipient (Schiffrin 2006, p.70). Definite entities relate to those that are recoverable within the discourse and/or the extra-linguistic context (Abbott 2010). In the words of Prince (1992, p.303), these entities are largely “discourse-old” entities that are familiar to the hearer. Indefinite entities, on the other hand, are not recoverable from the discourse or extra-linguistic context and are consequently not recognizable to the hearer (Jones 2018, p.1).

In the current study, nominals in isolated clauses were analysed for definiteness based on the lexico-grammatical features of the nominal group. For instance, the presence of certain determiners was taken as indicative of definite and indefinite status. While nominals preceded

by the determiner *the* typically presupposed an amount of knowledge on the side of the hearer, nominals preceded by the determiner *a* were largely analysed as indefinite, as they generally functioned to signal the introduction of new information (Schiffrin 2006, p.70; c.f. Jones 2018). In addition to the two determiners mentioned above, the presence of pronouns, proper nouns and possessives were generally taken as indicators of definiteness when they preceded a nominal. The presence of quantifiers, numerals and the lack of a prior linguistic unit (zero-article), alternatively, were understood as indicative of indefinite status when preceding a nominal. Figure 6-1 below provides an example of the definiteness annotation carried out in the *stream* dataset (Appendix C). The first nominal is annotated as definite because of the preceding *the* determiner, which presents the *stream* as recoverable to the hearer. The second instance is annotated as indefinite as there is no preceding determiner displayed. Because of the lack of the determiner, the *streams* are not specified and remain indefinite.

Number	Context	Definiteness
135	I loved fishing but the <u>stream</u> at the bottom of our garden was always black...	Definite
182	<u>Streams</u> of water blasted from firehoses into the smoke...	Indefinite

Figure 6-1. Definite and definite analysis, taken from *stream* dataset (*Stream Dataset, Appendix C*).

Definiteness provides a particularly interesting area for investigation with regards to its relationship with the lexical aspect expression of nominals, given it remains a largely unexplored research area within the field of nominal lexical aspect. Moreover, definiteness shares an evident association with count/mass status – a variable shown in Section 5.2 to influence lexical aspect expression of nominals. As discussed in the above paragraphs, the same preceding elements that influence the count/mass status of a nominal instance, e.g., definite/indefinite articles, quantifying and numeric determiners, etc., also simultaneously impact the definiteness of the nominal expression. Furthermore, the presence of definite status in the nominal group creates an essence of identifiability and type which may induce (temporal) boundaries in our conception of the nominal instance. For example, when we partner the definite article *the* with a mass nominal such as *water*, the definite article produces a specificity which acts to limit our conception of the *water* to a particular instance (Radden and Dirven

2007, p.61). While *water* is still construed as a mass substance, we can discern the *water* in question from other types of water because of the definite article. Given the typical association between boundaries and telicity (see Section 5.2), it is therefore interesting to investigate whether the definiteness of a nominal instance does hold any influence over its expression of temporal semantics.

Now that I have defined definiteness and provided motivation for its analysis, we will progress to the notion of process type. The next section will focus on process types and their categorisation, before then providing justification for its annotation.

### **6.2.2. SFL Process Type Annotation**

The process type of the clause that each nominal instance functioned in was analysed to further explore how the syntactic and semantic properties of four UENs interact with one another. The use of SFL is again relevant in the second study, as (like mentioned in Section 3.2) the principle aim of SFL is to acquire an understanding of how language works (Bloor and Bloor 2018, p.151). To achieve this fundamental aim, the framework prioritises the exploration of language within its social environment (Hasan 2009, p.37).

In SFL, the experiential metafunction focuses on the use of language to talk about one's experience of the world, including the worlds within one's mind, to describe the events, states and entities that are present within them (Thompson 2014a, p.24). The experiential metafunction is represented by the system of transitivity, which captures how experiences are encoded at the level of the clause. In the clause, Halliday and Matthiessen (2014, p.212) claim that change in the flow of events is construed through the configuration of a process, the participants involved in that process and any circumstantial elements surrounding the process (for a description of terms such as 'circumstance', see Section 2.3.1). Experience, therefore, largely centres around the process, and the type of process that is realised is highly dependent upon the main verb in the clause (Fontaine 2013, p.73). Within SFL, six different types of process, i.e., process types, are recognised: (1) Material; (2) Mental; (3) Relational; (4) Verbal; (5) Behavioural; and (6) Existential. Below, I will now describe these six different types of process and provide constructed examples to illustrate the differences between them.

Material processes are processes which relate to an “action in the material world in which someone or something does something” (Bartlett 2014, p.48). Four different participant roles are associated with the Material process. First is the Actor, i.e., the participant who conducts the process. Second is the Goal, which relates to the thing impacted by the process. Both participant roles are shown in example (55a) below, as the process of kicking *the ball* (the Goal) is a material action which is ‘done’ by an Actor, in this case *Jack*. In addition to these two participant roles, the participant roles of Beneficiary and Scope can also feature in Material processes. The participant role of Beneficiary is like Goal in that it is also impacted by the process. However, Beneficiaries are fundamentally participants that ‘benefit’ from the process or function as a recipient involved in the process (Fontaine 2013, p.74). A Material process which involves a Beneficiary is available in (55b) below. Scope, on the other hand, relates to a participant role that is not impacted by the process but construes “the domain over which the process takes place” (Halliday and Matthiessen 2014, p.239). An example of the Scope is shown in (55c).

Unlike Material processes, Mental processes do not involve material action, and instead relate to processes which transpire in one’s own consciousness (Halliday and Matthiessen 2014, p.245). Mental processes include two participant roles: the Senser and the Phenomenon. The Senser relates to the participant doing the sensing, while the Phenomenon represents the thing being sensed. For instance, in example (55d), the process of knowing an answer resides entirely within the head of a Senser, i.e., *Jack*, and therefore does not relate to an action of any kind. Additionally, *the answer* in this example represents the Phenomenon, as it is the thing that is known (sensed) by *Jack*.

Relational processes, like Mental processes, also do not relate to actions. They are instead processes in which two participants are related with respect to one another (Fontaine 2013, p.76). Relational processes can be either Attributive or Identifying. Attributive Relational processes ascribe an attribute to an entity (Bloor and Bloor 2013, p.123). They involve the participant roles of Carrier, i.e., the participant the attribute belongs to, and Attribute, i.e., the attribute itself. An example of an Attributive Relational process is shown in example (55e), where the quality (Attribute) of *thirstiness* is ascribed to the Carrier *Jack*. Identifying Relational processes, alternatively, function to identify participants (Halliday and Matthiessen 2014, p.276). These processes involve the participant roles of Identifier (the one doing the identifying) and Identified (the participant that is identified). In example (55f), the Identifier

participant *the teacher* does not attribute a quality to the Identified participant *Jack*, but instead assigns an identity to him (Fontaine 2013, p.124). These two types of Relational processes can usually be distinguished by interchanging the participants in the clause. While participants in Identifying Relational process can be seamlessly interchanged with one another (as both participants are used to identify the same entity), e.g., *the teacher is Jack*, participants in Attributive Relational clauses cannot, e.g., *\*thirsty is Jack*.

According to Halliday (1994), the three process types discussed above represent the most frequently used ways in which speakers express experience. However, like we have continuously acknowledged throughout this thesis, the process of categorisation is rarely a black and white matter. When analysing processes for process type, it is expected that instances will arise which will not slot perfectly into one process type category. Nevertheless, the accuracy of this classification can be enhanced through the addition of three further process types: Verbal; Behavioural; and Existential. Verbal processes function between the borders of Mental and Relational processes and relate to processes of saying (Bloor and Bloor 2013, p.125). As a result, these processes can involve up to three participant roles: the Sayer; the Receiver; and Verbiage. Expectedly, the Sayer represents the participant who expresses the utterance, while the Receiver relates to the participant who receives the utterance. The Verbiage participant role concerns the utterance itself. Example (55g) provides an instance of a Verbal process, as the Sayer *Jack* utters the Verbiage *he'd be there in five minutes* to the Receiver *his boss*.

In addition to Verbal Processes, we also have Behavioural processes. Halliday (1994, p.107) defines these processes as “the acting out of processes of consciousness and psychological states”. These processes only require one participant: a Behavior. Behavioural processes are conceptualised to operate between the borders of Material and Mental processes (Thompson 2014a, p.109). For instance, while in example (55h) *Jack* displays the physical action of *breathing*, this action is the product of a neurological and cognitive experience (Fontaine 2013, p.78). *Jack* cannot truly be said to be ‘acting’ in this process, nor can he be said to be completely sensing, due to the physical manifestation of this neurological and cognitive process. Accordingly, in the process of *breathing*, the process is most appropriately categorised as Behavioural, with *Jack* understood as a Behavior. Lastly, there is the Existential Process. Existential processes operate between Material and Relational processes, and function to state the existence of an entity (Halliday and Matthiessen 2014, p.215). These processes only involve



one participant, labelled the Existent. The Existent relates to the thing within the clause that exists. Clauses which feature an Existential process are rather distinct from clauses which involve other process types, as they generally include the lexical item *there* in the subject role of the clause (Halliday and Matthiessen 2014, p.220). This distinction is shown in (55i) below, where *a tree* functions as the Existent.

- (55) a. Jack kicked the ball.  
b. Jack gave the ball to Mary.  
c. Jack is playing football.  
d. Jack knows the answer.  
e. Jack is thirsty.  
f. Jack is the teacher.  
g. Jack said he'd be there in five minutes to his boss.  
h. Jack is breathing.  
i. There is a tree.

As mentioned above, the process type of the clause in which the nominal instance functions was analysed to examine the potential interactions between the syntactic and semantic properties of four UENs. The analysis of process type marks an intriguing avenue of research, as we know that the construal of lexical items is largely impacted by their surrounding syntactic context. Halliday and Martin (1993) have previously suggested an association between nominalised technical terms and Relational processes, due to the capacity of Relational processes to assist in the classification of these nominals and the establishment of taxonomies. However, despite this suggestion, the extent to which the type of experience encoded within the clause can influence the temporal semantic behaviour of particular nominals has yet to be explored. Furthermore, the fact that the nominals in this study are morphologically underived adds another interesting dimension to this research, as there is little known about the exact factors which facilitate the expression of temporal semantics by underived nominals. Accordingly, through the analysis of process type and the definiteness of the nominal instances in question, it will shine light upon the extent to which process type and definiteness hold influence over the expression of temporal semantics across the four UENs in question.

For the process type analysis, a comprehensive set of syntactic probes, taken from Fontaine (2013, pp.85–91), was employed on the main verb of the clause that each nominal instance

functioned in. Typically, the process type was discernible from the lexicogrammar. Therefore, Fontaine’s syntactic probes were only used for main verbs which did not initially express a clear orientation towards one process type. A selection of syntactic probes is shown below in Table 6-2. It is worth noting that the selection of probes in Table 6-2 is not extensive, and that the identification of process type typically relies on a combination of probes, not just one probe alone. Nevertheless, the scope of this chapter cannot cover all these probes and their combinations, so I will only present a sample of probes that were drawn upon frequently in the annotation. For a full explanation and description of syntactic probes used to help categorise process types, see Fontaine (2013, pp.85–91).

*Table 6-2. A selection of syntactic probes from Fontaine (2013, pp.85–91) for the analysis of process type.*

Number	Syntactic Probe	Process Type Probe
1	Probe: What did [participant] do? Response: What [participant] did was to...	Material
2	Probe: [participant]<perceives/knows/believes/likes/etc.> something.	Mental
3	Probe: The thing about [Carrier] is that it is [participant].	Relational Attribute
4	[participant1] is [participant2] vs [participant2] is [participant1]?	Relational Identifying
5	Can/does the clause include a Receiver as a participant (even if it does not include one in this particular case)?	Verbal
6	Is [participant] involved in an activity? (if ‘yes’, [participant] may be Behaver).	Behavioural
7	Is the Subject of the clause ‘there’?	Existential

### **6.2.3. Lexical Aspect Annotation**

Lexical aspect relates to the ‘inherent’ temporal semantic meaning expressed by individual lexical items (Smith 1997). The foundation of these temporal structures is principally based upon the lexical aspect features of dynamism, duration, and telicity (for a full overview, see Section 2.4). Through the interaction of these lexical aspect features, different temporal structures, i.e., situation types, are expressed. As you will be familiar with from reading Chapters 3-5, the analysis of lexical aspect generally concerns the use of diagnostic syntactic tests (DSTs) (Kenny 1963; Vendler 1967; Dowty 1979; Smith 1991; Rothstein 2004; Declerck 2006; Balvet et al. 2011). DSTs are an effective resource when analysing lexical aspect, as they specify whether a given lexical item can sufficiently operate within a syntactic construction indicative of a particular lexical aspect feature. If the lexical item functions satisfactorily within

the syntactic construction, it is deemed to express the lexical aspect feature. If the lexical item is unable to function adequately within the syntactic construction, it will be classified as not expressing the lexical aspect feature. When multiple DSTs that test for separate lexical aspect features are used together, they provide an empirical method that allows analysts to discern the situation types expressed by different lexical items. The five situation types analysed for in this thesis are Accomplishment; Achievement; Activity; Semelfactive; and State. Indirectly, the DSTs also analyse for atemporal structures, as an inability to efficiently operate in all the DSTs in question will signal a lack of lexical aspect, and a resultant expression of object meaning. However, in a similar vein to the previous study conducted in this thesis (see Chapters 3-5), the statistical analysis (described in Section 6.3) will only report on the situation types that occurred within the lexical aspect annotation of the nominal instances.

Eight DSTs were used for the annotation of lexical aspect. When placed together, they provide a typology of nominals, distinguishing event, state, and object nominals from one another. The eight DSTs represent the same DSTs employed in the first study of the thesis. Consequently, I will only display the eight DSTs in this section. If the reader wishes to review the specific functionality of each DST and the justification for their selection, this information is detailed in Section 3.2.2. The eight DSTs are available in Table 6-3 below.

*Table 6-3. DSTs for lexical aspect annotation and their literature source.*

<b>Syntactic test</b>	<b>Literature source</b>
Test1: X took place:	Balvet et al. (2011)
Test2: Felt X / experience X:	Balvet et al. (2011)
Test3: X lasted time:	Balvet et al. (2011)
Test4: Perform / Carry out X:	Balvet et al. (2011)
Test5: In a state of X:	Balvet et al. (2011)
Test6: X unfolded:	Balvet et al. (2011)
Test7: X is v-ing entails X has v-ed + was interrupted':	Dowty (1979); Barque et al. (2009)
Test8: X is v-ing entails X has v-ed + 'was interrupted':	Dowty (1979); Barque et al. (2009)

In the next section of this chapter, I will now detail the statistical analyses that were undertaken using the data from each of the three annotations discussed above.

## 6.3. Statistical Analysis

### 6.3.1. Multinomial Logistic Regression

For the statistical analysis of the four datasets, multinomial logistic regression analyses were carried out. Regression analysis, as shown in the first study of this thesis, concerns modelling a relationship between a dependent variable and one or more predictor variables (Levshina 2015, p.120). At its core, regression analysis is used to identify whether a variable can predict the behaviour of another. Logistic regression models differ in relation to standard linear regression models, as they cater for categorical dependent variables (Hosmer et al. 2013, p.1). Multinomial logistic regression is a particular type of logistic regression that is applicable when categorical dependent variables express more than two categories (Menard 2010, p.2). For instance, the variable *animal* expresses many different categories, i.e., *cat*, *dog*, *fish*, etc. These categories are unique and express no sense of inherent order.

A multinomial logistic regression analysis was chosen as, unlike linear regression, logistic regression does not assume a balanced data distribution (Bresnan et al. 2007, p.13). It instead provides a robust test for modelling probability based on unbalanced data. The non-parametric nature of logistic regression was beneficial for the current study, as each of the four datasets expressed an unbalanced distribution. In addition to this capacity, logistic regression does not assume a linear relationship between data variables (Bayaga 2010, p.291). As discussed above, this quality was favourable for the current study as each dataset was comprised only of categorical variables which are unable to share a linear relationship. Furthermore, multinomial logistic regression facilitates more efficient use of the data, as it allows for the simultaneous estimation of each predictor variable's influence on the dependent variable. This quality, in turn, leads to stronger statistical tests for hypotheses and more accurate estimations of parameters (Anderson and Rutkowski 2008, p.391). Lastly, multinomial logistic regression offers a more efficient alternative to other statistical measures such as the Chi-Squared test, as multinomial logistic regression informs the analyst of a significant difference in the data and further shows where that difference lies, as it provides confidence intervals that can be compared across groups.

Multinomial logistic regression is a popular statistical method in the field of Linguistics (Winter 2020, p.198). Within the current study, multinomial logistic regression was used to

identify the extent to which definiteness and process type influence lexical aspect expression across the four UENs. The analysis was again undertaken in the quantitative software environment 'R' (R Core Team 2018), this time, using the *multinom* function located in the *nnet* package (Venables and Ripley 2002). For the multinomial logistic regression analyses, an intercept category was composed. 'Baseline' models require the selection of a intercept category from the dependent variable, e.g., *Accomplishment* from situation type, and also intercept category from the predictor variables, e.g., *indefinite* from definiteness and *Material* from process type (Anderson and Rutkowski 2008, p.390). Together, these variables form what is known as the 'intercept condition'. To attain the most reliable results, the intercept condition for each UEN was composed of the most frequent category for situation type, definiteness and process type.

The effect of the intercept condition is then contrasted with the effect of every other predictor variable on the dependent variable. This effect is expressed in coefficients, in the form of 'log-odds'. Log-odds are logarithmically transformed odds that range from -infinity to infinity (Levshina 2015, p.261). They provide data on the odds of an event occurring when compared to another. While negative log-odds signal that an event is less probable, positive log-odds indicate that an event is more probable (Levshina 2015, p.261). Log-odds of zero, on the other hand, indicate no change in the probability of one event occurring than another. For instance, if our baseline category is indefinite; Material; and Accomplishment; the intercept condition may show a decrease of -1.2 when we change Accomplishment for Activity. This decrease in log-odds signals that the nominal in question is less likely to express an Activity than an Accomplishment when the nominal is indefinite and involved in a Material process. Moreover, if we look at the effect of different conditions on Activity expression, we might find that Mental process type expresses a log-odds of 0.8. These log-odds suggest that while the involvement of the nominal in a Mental process does increase probability of an Activity being expressed when compared to Material processes, the nominal is still less likely to express an Activity over an Accomplishment ( $-1.2 + 0.8 = -0.4$ ).

To assess the overall quality of the models, Nagelkerke's Pseudo  $R^2$  was computed. Nagelkerke's Pseudo  $R^2$  provides information on the amount of variation in the dependent variable that is predictable from the overall model (Bayaga 2010, p.292). Nagelkerke's Pseudo  $R^2$  measure is a frequently used measure to assess the quality of logistic regression models. It should be pointed out however, that scholars such as Hosmer and Lemeshow (2000) and

Osborne (2015, p.51) question the validity of  $R^2$  results for logistic regression models, as the measurement can be volatile and lacks clarity when compared to its use in linear regression. To compensate for the potential of this volatility, I will also report additional information on the prediction accuracy of the models.

The statistical significance of each the two predictor variables (definiteness and process type) was then identified using a Likelihood-Ratio test. The Likelihood-Ratio test acts to compare the predictive power of a null model with a more complex model that includes the influence of our predictor variables (Winter 2020, p.260). If the difference in predictive power is significantly different to that of the null model, a relationship between the predictor variable and the dependent variable is evident. For this evaluation, the level of significance was set at  $p=0.05$  (the same significance level set in the first study of the thesis).

This study will address the following research question (RQ) and two sets of hypotheses across the four UENs:

**RQ:** To what extent do definiteness and process type hold influence over the temporal semantics expressed across four UENs.

**H0:** Definiteness does not share a dependent relationship with the semantic behaviour expressed by the UEN.

**H1:** Definiteness does share a dependent relationship with the semantic behaviour expressed by the UEN.

**H0:** Process type does not share a dependent relationship with the semantic behaviour expressed by the UEN.

**H1:** Process type does share a dependent relationship with the semantic behaviour expressed by the UEN.

## 6.4. Results

In this next section, I will present the results from the multinomial logistic regression analyses. The section will split into four sub-sections that will each focus on one of the UENs. For the multinomial logistic regression analyses, the most frequent definiteness status, process type and situation type for each UEN were selected as the base category. This decision was made so that the intercept condition of each model had a sufficient number of observations, making its estimation and other estimations more reliable. It should be noted here that, for the nominal *drought*, no inferential statistics are reported, as there was not enough differentiation of situation type expression by the nominal (500 = State) to run a meaningful statistical analysis. As a result, Section 6.4.4 will only report descriptive statistics of the distribution of *drought* across definiteness and process type. Also, different axis percentages were used in different visualisations in order to best visualise the statistical results. We will first start the results by looking at the UEN *game*.

### 6.4.1. Game

In the *game* dataset, definite status, Material process type and Accomplishment situation type recorded the most observances for their respective variables. The intercept condition was therefore set as definite; Material; Accomplishment. A table presenting the output from the multinomial logistic regression model *situation type ~ definiteness + process type* is presented below, involving the raw coefficients (log-odds) and standard errors. To help the reader interpret this table, I will use the next paragraph to go through some of the log-odds recorded and what they suggest.

Table 6-4. Coefficients and std. errors produced for game by situation type ~ definiteness + process type.

Situation Type	Condition	Coefficients	Std. Errors
Activity	Intercept	-1.98	0.26
	Indefinite	-0.75	0.45
	Existential	-10.19	299.13
	Mental	-0.50	1.06
	Relational_A	-0.41	0.52
	Relational_I	-0.55	0.77
	Verbal	-7.66	139.27
Object	Intercept	-0.82	0.15
	Indefinite	0.05	0.20
	Existential	0.77	0.73
	Mental	-0.19	0.50
	Relational_A	-0.07	0.26
	Relational_I	0.52	0.30
	Verbal	2.18	1.13

In the above table, definite, Material and Accomplishment do not feature because, as noted above, these variables are acting as the base category from which we can draw predictive comparison. A comparison of probability of occurrence with the baseline categories is available in Table 6-4 in the ‘intercept’ condition of each situation type row. The remaining coefficients describe the difference in probability with the influence of other conditions (i.e., indefinite, Existential, Mental, etc.). For instance, the intercept condition shown in the Activity row indicates that items that are definite and Material will more likely produce an Accomplishment than an Activity (log-odds = -1.98). Going down the condition column, if we replace Material process type with Mental process type, the probability of Activity expression is further reduced for the nominal *game*, as the log-odds are lower than the intercept condition by -0.50. By comparing each of the log-odds from the predictor variable conditions with the intercept condition, we can identify the conditions under which *game* will most likely express a certain situation type.

Taking these coefficients, the multinomial logistic regression model generates predicted probabilities expressed in the form of percentages relating to the influence of each condition on the situation type expression of *game*. These probabilities are also accompanied by 95%



confidence intervals, which provide information on the reliability of the predicted percentage. The 95% confidence intervals tell us that if we carried out the estimation process repeatedly on different samples from the population, there would be a 95% probability that the stated confidence range would contain the true parameter value (Levshina 2015, p.98). Both the prediction rate percentage and the confidence 95% interval range are available below in Table 6-5. These results were collected using the *get\_model\_data()* function in the *sjPlot* R package (Lüdecke 2021). The results are also visualised in Figures 6-2 and 6-3, which were produced using the *plot\_model* function, similarly found in the *sjPlot* R package. It should be noted here that for each UEN, we will see a number of extreme prediction rates and large confidence intervals (e.g., Activity instances in Existential clause types, shown in Table 6-5 below). These prediction rates and large confidence intervals were largely the result of a lack of recorded observations for certain categories, not because of an increase in variation. To ensure the reliability of discussion throughout Section 6.4, I will not be focusing on these results, as their interpretation is problematic.

Table 6-5. Prediction rate and confidence intervals produced for game by situation type ~ definiteness + process type.

Situation Type	Condition	Predicted	95% Confidence Interval
Accomplishment	Definite	64%	26% - 90%
	Indefinite	65%	45% - 81%
	Existential	51%	20% - 81%
	Material	64%	59% - 70%
	Mental	70%	49% - 85%
	Relational_A	67%	58% - 76%
	Relational_I	55%	42% - 67%
	Verbal	20%	3% - 69%
Activity	Definite	6%	0% - 100%
	Indefinite	3%	0% - 100%
	Existential	0%	0% - 100%
	Material	7%	4% - 10%
	Mental	4%	1% - 25%
	Relational_A	5%	2% - 11%
	Relational_I	3%	1% - 12%
	Verbal	0%	0% - 100%
Object	Definite	30%	16% - 51%
	Indefinite	32%	23% - 44%
	Existential	49%	19% - 80%
	Material	29%	24% - 34%
	Mental	26%	12% - 47%
	Relational_A	28%	20% - 38%
	Relational_I	42%	30% - 54%
	Verbal	80%	31% - 97%

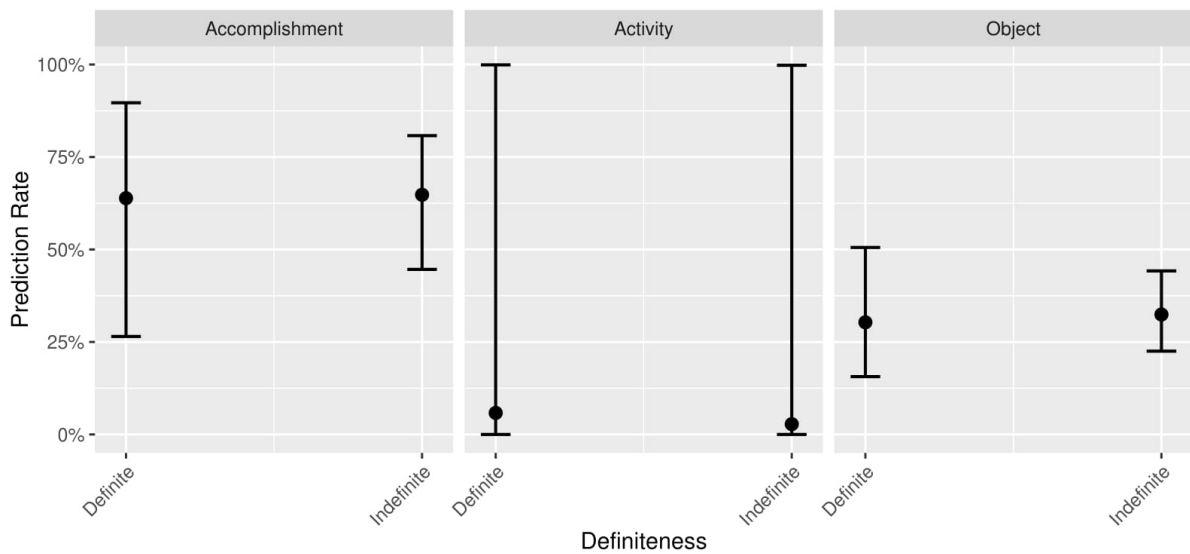


Figure 6-2. Prediction rate and confidence intervals of definiteness across situation type for game.

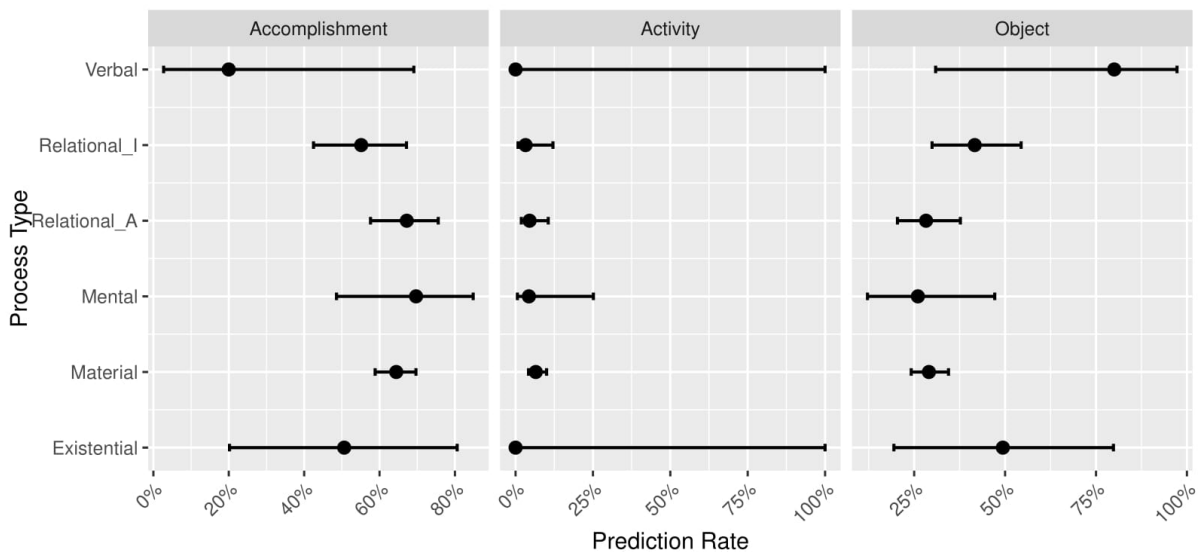


Figure 6-3. Prediction rate and confidence intervals of process type across situation type for game.

Table 6-5 and Figures 6-2 and 6-3 show that Accomplishment was the most highly predicted situation type by both definite and indefinite status and every process type, apart from Verbal at 20%. Looking specifically at Figure 6-2, there is very little differentiation in definiteness across situation type. For instance, definite instances (64%) and indefinite instances (65%)

predict Accomplishment at very similar rates. This similarity is also displayed for Objects (30% vs 32%) and Activities (6% vs 3%).

Across the process type categories, Accomplishment was generally the highest predicted situation type. This was followed by Object and then Activity. Unlike definiteness, there does appear to be slightly more differentiation in the prediction rates of situation type across the process type categories. The condition of Material process type generates a prediction rate of 64% for Accomplishment. This is then increased to 70% if the process type is changed to Mental. We also see a large decrease in prediction rate if the condition is changed to Verbal, with only a 20% prediction rate for Accomplishment. However, only five verbal instances were recorded in the *game* dataset, so the validity of this prediction rate remains questionable.

Moreover, the condition of Relational\_I interestingly signals a greater variation of prediction rate across situation type than Relational\_A. When *game* was expressed as either an Accomplishment or an Activity, the nominal was more likely to be described by an adjective in a Relational clause than it was to be equated with another nominal group. In Relational\_A clauses that expressed *game* as an event, writers more typically presented their opinions about *games*, with *game* functioning in the Carrier participant role (as shown in example (56a)). In addition, the expression of possession was also prevalent in Relational\_A clauses. In clauses that indicated possession, *game* was identified in both the Carrier and Attribute participant roles. This distinction is shown below, where *game* functions in the Carrier participant role in (56b) and the Attribute participant role in (56c). On the other hand, Relational clauses that functioned to identify participants, i.e., Relational\_I clauses, also showed the capacity to include the event *game* as a participant, shown in (56d). However, this type of process showed a stronger patterning than Relational\_A towards the Object expression of *game*. A notable number of Relational\_I clauses concerned types of video *games* (66.6%), where the Relational\_I clause functioned to state the foundation and/or aspects of the *game*, as shown in example (56e).

(56) a. ...this game isn't quite over. (Noun 304, *Game* Dataset, Appendix C)

b. The games had playoff atmospheres... (Noun 408, *Game* Dataset, Appendix C)

- c. Both teams have games with county rivals scheduled for next Friday... (Noun 79, *Game Dataset*, Appendix C).
- d. Still, that game was the only one that City dropped points at home in during their title-winning season. (Noun 208, *Game Dataset*, Appendix C)
- e. ...PEPI BATH is a role-play game where children learn about hygiene in a fun way. (Noun 347, *Game Dataset*, Appendix C)

The multinomial logistic regression model for *game* returned a low Nagelkerke Pseudo  $R^2$  (0.03969419). While this calculation does aim to report the amount of variance in the dependent variable that is explained by the model, I will remind the reader that, as stated in Section 6.3.1, the reporting of  $R^2$  for logistic regression models should be taken cautiously, as the measurement is prone to volatility (Osborne 2015, p.51). To augment this statistic, I will report further information on the model by noting its prediction accuracy. For the prediction accuracy, an intermediate prediction accuracy (63.6%) was identified, meaning that the model accurately predicted the categories of 63.6% of observations. To assess the statistical significance of each of the two predictor variables (definiteness and process type), the Likelihood-Ratio test was performed, which compares two models with and without each variable. A non-significant relationship was found for both definiteness:  $X^2(2) = 3.3638$ ,  $p = 0.1860$  and process type:  $X^2(10) = 12.2907$ ,  $p = 0.2661$ . Therefore, despite some slight alterations in semantic behaviour by the nominal *game*, the results suggest that definiteness and process type do not hold a significant influence over the semantic behaviour of *game*. As a result, we cannot discard the two null hypotheses.

#### 6.4.2. Fire

The next UEN that was analysed is *fire*. In the *fire* dataset, indefinite status, Material process type and Activity situation type recorded the highest number of observations. Consequently, for the multinomial logistic regression model, the intercept condition was set at indefinite; Material; Activity. The output of the model *situation type ~ definiteness + process type* is presented below, including the raw coefficients and standard errors.

Table 6-6. Coefficients and std. errors produced for fire by situation type ~ definiteness + process type.

Situation Type	Condition	Coefficients	Std. Errors
Accomplishment	Intercept	-18.82	470.59
	Definite	-15.62	<0.01
	Existential	-0.05	9.13
	Mental	0.19	14.05
	Relational-A	15.14	470.59
	Relational-I	2.78	519.62
	Verbal	-0.30	12.51
Object	Intercept	-0.92	0.14
	Definite	-1.42	0.32
	Existential	0.83	0.70
	Mental	-44.70	NaN* <sup>6</sup>
	Relational-A	-0.94	0.42
	Relational-I	0.08	0.59
	Verbal	-0.52	0.79
State	Intercept	-3.64	0.43
	Definite	0.43	0.55
	Existential	-29.49	<0.01
	Mental	-39.02	<0.01
	Relational-A	0.74	0.62
	Relational-I	0.52	1.09
	Verbal	-37.08	<0.01

In similar vein to the previous nominal *game*, the coefficients above were used to generate predictive percentages for each condition on the situation type expression of the nominal *fire*. These percentages are presented in Table 6-7 and Figures 6-4 and 6-5 below.

<sup>6</sup> Stands for 'not a number'. It could not be converted, as its value is either too large or too small.

Table 6-7. Prediction rate and confidence intervals produced for fire by situation type ~ definiteness + process type.

Situation Type	Condition	Predicted	95% Confidence Interval
Accomplishment	Definite	0%	0% - 100%
	Indefinite	0%	0% - 100%
	Existential	0%	0% - 100%
	Material	0%	0% - 100%
	Mental	0%	0% - 100%
	Relational-A	0%	0% - 0%
	Relational-I	0%	0% - 100%
	Verbal	0%	0% - 100%
Activity	Definite	97%	95% - 98%
	Indefinite	89%	86% - 91%
	Existential	64%	32% - 87%
	Material	79%	74% - 83%
	Mental	100%	100% - 100%
	Relational-A	86%	77% - 92%
	Relational-I	76%	53% - 90%
	Verbal	87%	60% - 97%
Object	Definite	3%	2% - 5%
	Indefinite	11%	8% - 13%
	Existential	36%	13% - 68%
	Material	19%	15% - 24%
	Mental	0%	0% - 0%
	Relational-A	8%	4% - 17%
	Relational-I	20%	8% - 44%
	Verbal	13%	3% - 40%
State	Definite	0%	0% - 1%
	Indefinite	0%	0% - 0%
	Existential	0%	0% - 0%
	Material	2%	1% - 5%
	Mental	0%	0% - 0%
	Relational-A	6%	2% - 14%
	Relational-I	4%	1% - 24%
	Verbal	0%	0% - 0%

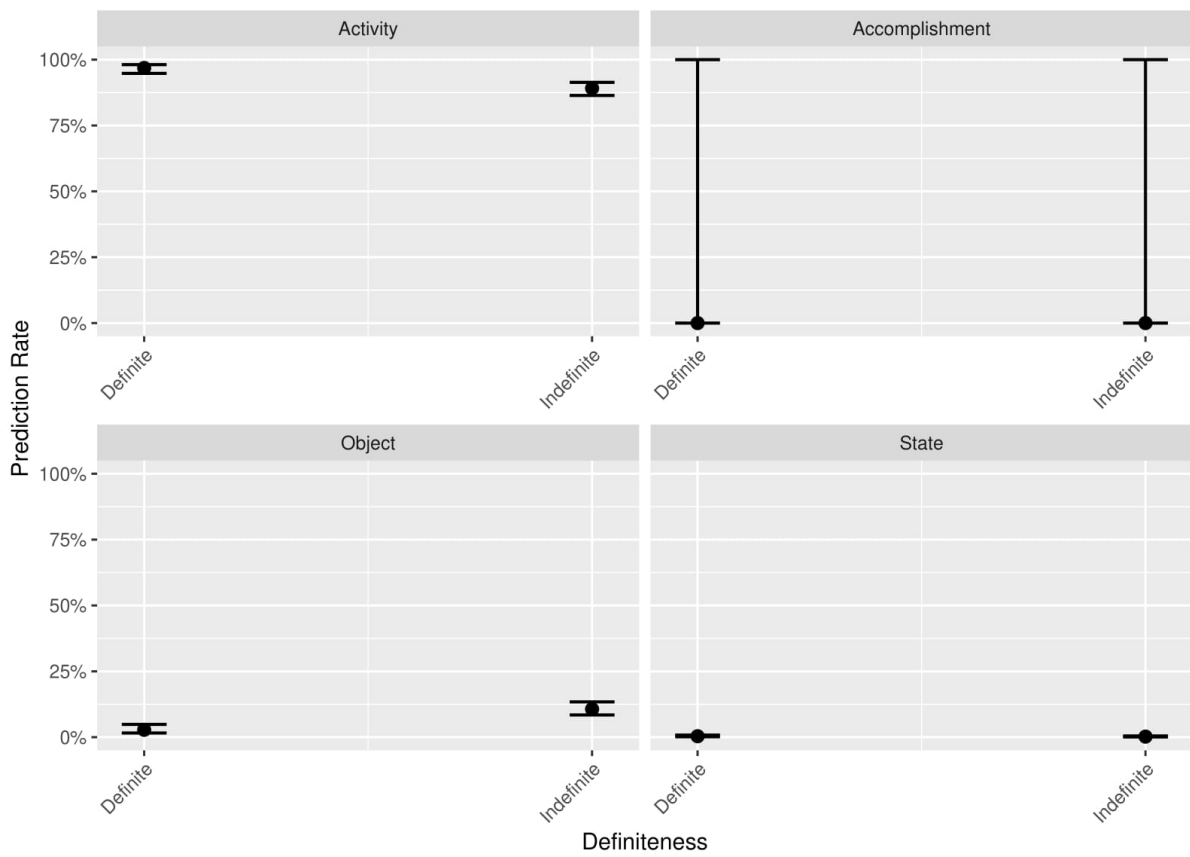


Figure 6-4. Prediction rate and confidence intervals of definiteness across situation type for fire.



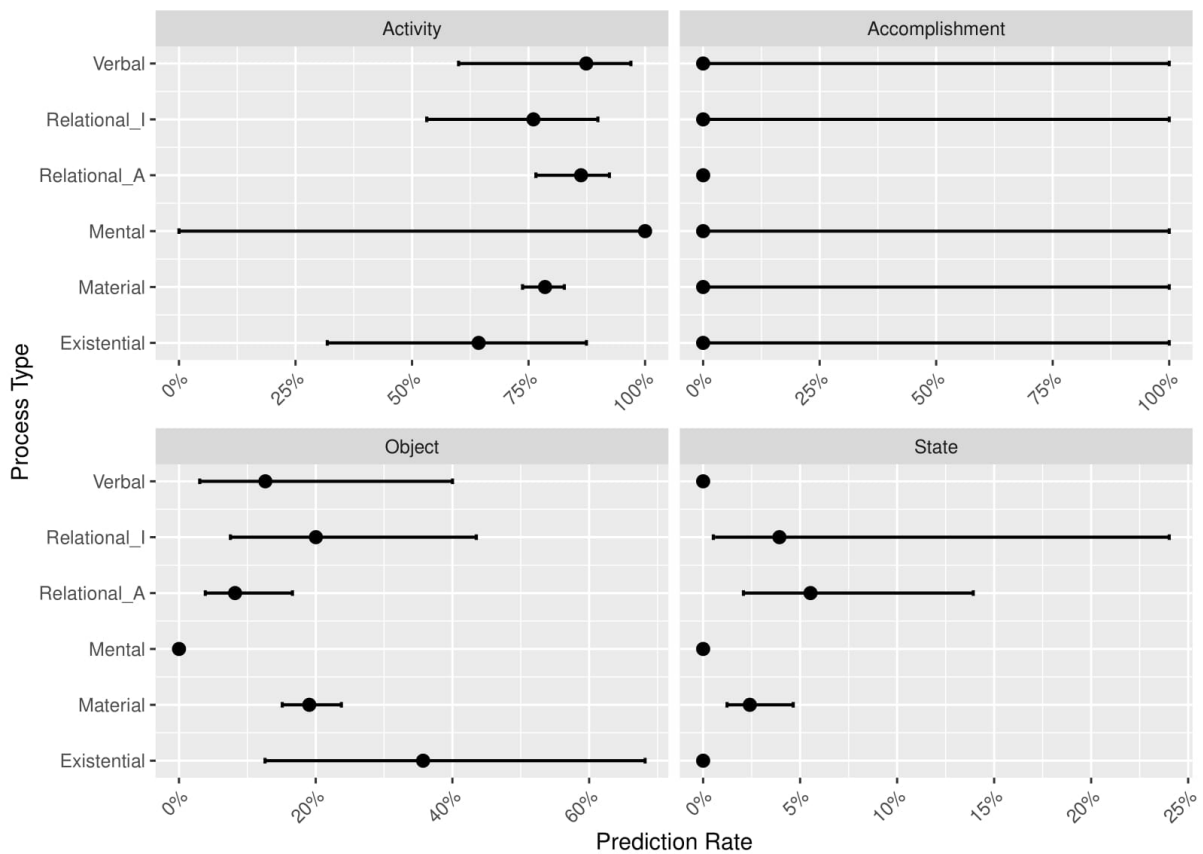


Figure 6-5. Prediction rate and confidence intervals of process type across situation type for fire.

In Table 6-7 and Figures 6-4 and 6-5, Activity is the highest predicted situation type for *fire*. Interestingly, despite expressing four different situation types (contrasted with the three expressed by *game*), *fire* displays less variation than *game* in its situation type expression, with a high concentration of all conditions expressing Activity. In addition, there is an apparent difference in prediction rate between definite (97%) and indefinite (89%) status, as the confidence intervals of definite and indefinite Activity instances do not overlap. This result suggests that definite status is more influential in the expression of Activity by *fire* than indefinite status. On the other hand, the prediction rate expressed by indefinite status for Object (11%) is notably higher than that expressed by definite status (3%). This greater expression of Object by indefinite instances was largely the result of zero-articles within the nominal group, which signal indefiniteness. A definite Activity instance (57a), an indefinite Activity instance (57b) and an indefinite Object instance (57c) are presented below.

- (57) a. The small fire broke out at the Oxford University research lab... (Noun 32, *Fire Dataset*, Appendix C)
- b. ...a fire or flood can destroy a lifetime of memories overnight. (Noun 246, *Fire Dataset*, Appendix C)
- c. ...many of the rooms in the hotel had been substantially destroyed by fire... (Noun 420, *Fire Dataset*, Appendix C)

Contrasting the examples (57a-c) above, our semantic interpretation of (57c) is distinctly different to (57a-b). Examples (57a-b) both display Activity events, as these instances of fire express dynamicity, duration, and do not express an inherent temporal endpoint. Example (57c), on the other hand, represents *fire* as an atemporal physical substance, i.e., an Object, that caused damage, as opposed to a dynamic event in which the damage was caused. This distinction is particularly intriguing, as the absence of an article in the nominal group (i.e., zero-article) in (57c) appears to be influential in our interpretation of *fire*. For instance, if we take the first DST used in the lexical aspect annotation (‘X took place’) and apply it to each example in (57) above, we see that while the instances of *the fire* and *a fire* function satisfactorily, the instance of *fire* alone, i.e., *fire took place*, does not function to the same degree of satisfaction. Here, the use of the zero-article is key, as it assists in the construal of *fire* as a mass substance, and not an individuated instance of fire.

Radden and Dirven (2007, p.73) suggest that when an entity is construed as a substance, it can become restricted to a certain domain, e.g., *squid* restricted to the domain of *food*. They continue to state that removing conceptual boundaries of an entity can direct attention away from the intricacies in the entity’s “whole and internal composition” (Radden and Dirven 2007, p.73). For instance, if our conceptualisation of *fire* is only limited to its status as a type of physical substance, its temporal qualities become inaccessible. In addition, when partnered with mass nominals, zero-articles help establish an indeterminateness between the start and endpoints of the entity in question (Maekelberghe and Heyvaert 2016, pp.320–321). Because of this indeterminateness, we are also unable to discern the potential event boundaries of the nominal, and, as a result, it is conceptualised as a static substance rather than a dynamic event.

When considering the influence of process type on situation type expression, there was little variation in prediction rates across the different process type categories. Specifically, each process type displayed a high prediction rate for Activity. One of these high prediction rates that deserves a little more attention is Mental process type, which expressed a 100% prediction rate for Activity. As shown in example (58) below, *fire* was only construed as an event when used within a Mental process type clause. Construals such as these were largely evident in clauses that involved *seeing* the *fire* or *thinking* about the *fire* to some degree. However, we should consider that the confidence interval for Activities expressed in Mental process type clauses is very large, so more data is certainly required to assess the validity of this result.

- (58) ...an off-duty volunteer firefighter from Lincoln County was driving by and saw the fire (Noun 1071, *Fire* Dataset, Appendix C)

The MLR *fire* model returned a low Nagelkerke Pseudo  $R^2$  (0.1302877). However, this Pseudo  $R^2$  result is higher than what the Nagelkerke Pseudo  $R^2$  reported for the MLR *game* model. The MLR *fire* model also reported a fairly high prediction accuracy (77.8%), once again higher than the MLR *game* model. From the Likelihood-Ratio test, it was identified that definiteness shared a significant relationship with situation type expression:  $X^2(3) = 27.093$ ,  $p = <0.001$ , but a non-significant relationship was found for process type:  $X^2(15) = 19.985$ ,  $p = 0.1725$ . These results suggest that, while definiteness does hold an influence over the semantic behaviour of the nominal *fire*, process type does not. Accordingly, we can disregard the null hypothesis for definiteness, but we cannot disregard the null hypothesis for process type. If we compare these results with the previous results of the MLR *game* model, there is an apparent difference in the influence definiteness has over semantic behaviour between the nominals *game* and *fire*. Such a difference suggests that variables such as definiteness can express different degrees of influence on semantic behaviour, dependent on the lexical item in question.

### 6.4.3. Stream

The third UEN explored in the study was *stream*. In the *stream* dataset, indefinite status, Material process type and Object meaning recorded the highest number of observations. Therefore, the intercept condition was set at indefinite; Material; Object. The output of the MLR model *situation type ~ definiteness + process type* is presented below, including the raw coefficients and standard errors.

Table 6-8. Coefficients and std. errors produced for stream by situation type ~ definiteness + process type.

Situation Type	Condition	Coefficients	Std. Errors
Accomplishment	Intercept	-1.17	0.17
	Definite	0.32	0.24
	Behavioural	16.75	<0.01
	Existential	-0.67	1.09
	Mental	-0.08	0.83
	Relational-A	-0.09	0.30
	Relational-I	-0.30	0.48
	Verbal	0.26	1.24
Activity	Intercept	-0.89	0.16
	Definite	-0.19	0.24
	Behavioural	-3.04	<0.01
	Existential	1.07	0.58
	Mental	0.28	0.72
	Relational-A	-0.53	0.34
	Relational-I	0.12	0.40
	Verbal	-12.06	480.80

Once again, the coefficients above were used to generate percentages related to the influence of each condition on the situation type expression of the nominal *stream*. These percentages are available in Table 6-9 and Figures 6-6 and 6-7 below.

Table 6-9. Prediction rate and confidence intervals produced for stream by situation type ~ definiteness + process type.

Situation Type	Condition	Predicted	95% Confidence Interval
Accomplishment	Definite	24%	8% - 55%
	Indefinite	18%	5% - 51%
	Behavioural	100%	100% - 100%
	Existential	8%	1% - 39%
	Material	20%	16% - 25%
	Mental	18%	4% - 50%
	Relational-A	21%	13% - 30%
	Relational-I	15%	7% - 30%
	Verbal	31%	4% - 83%
Activity	Definite	17%	0% - 98%
	Indefinite	22%	0% - 99%
	Behavioural	0%	0% - 0%
	Existential	49%	25% - 73%
	Material	22%	18% - 27%
	Mental	28%	9% - 59%
	Relational-A	15%	9% - 24%
	Relational-I	26%	14% - 41%
	Verbal	0%	0% - 100%
Object	Definite	58%	11% - 94%
	Indefinite	60%	6% - 97%
	Behavioural	0%	0% - 0%
	Existential	44%	21% - 69%
	Material	58%	52% - 63%
	Mental	55%	27% - 80%
	Relational-A	65%	54% - 74%
	Relational-I	59%	43 - 73%
	Verbal	69%	17% - 96%

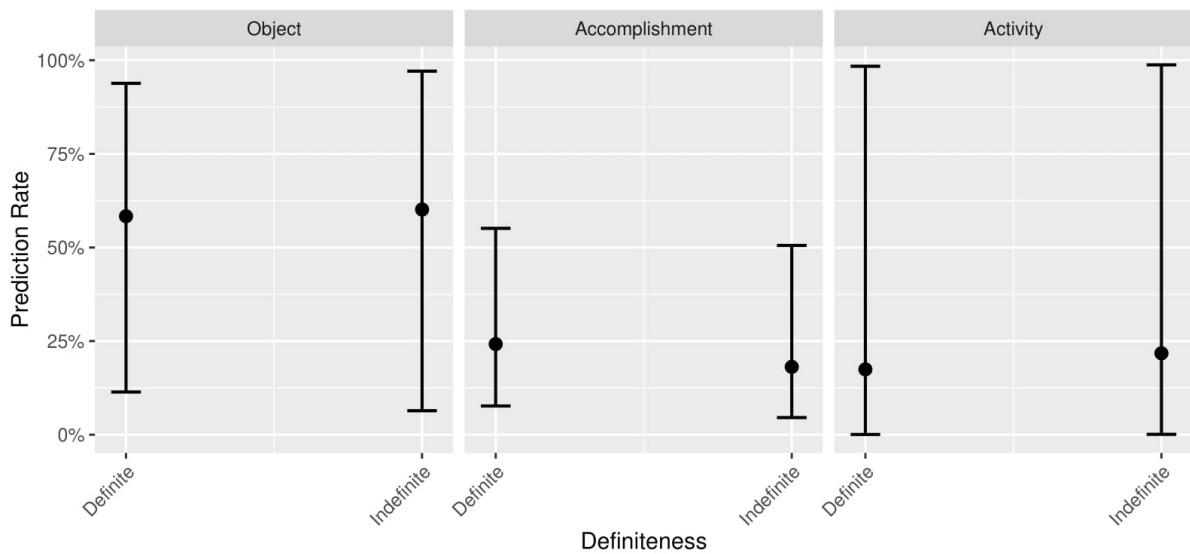


Figure 6-6. Prediction rate and confidence intervals of definiteness across situation type for stream.

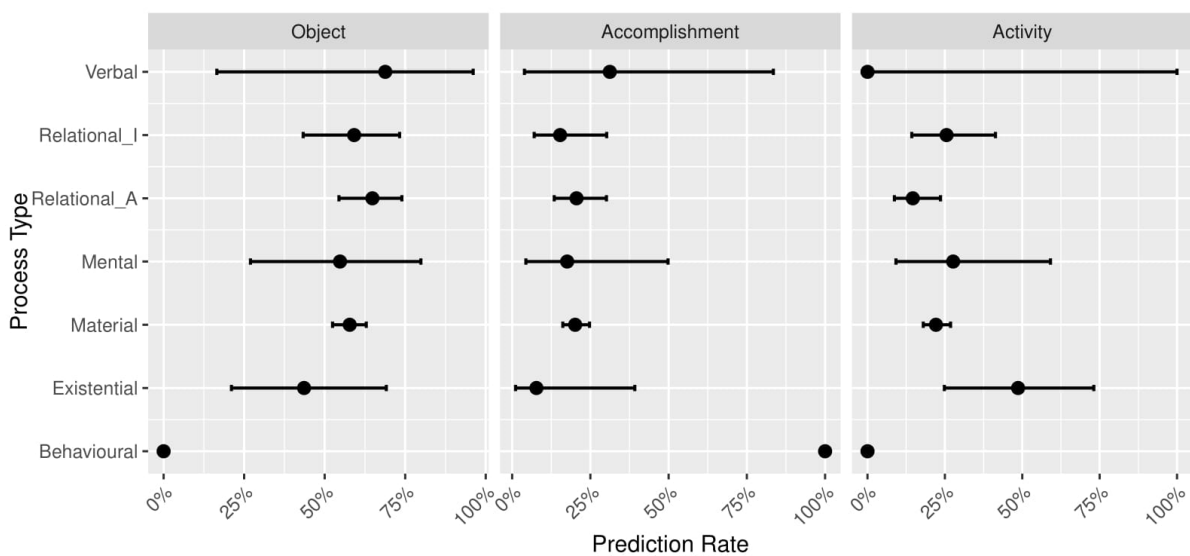


Figure 6-7. Prediction rate and confidence intervals of process type across situation type for stream.

In Table 6-9 and Figures 6-6 and 6-7, Object is generally the highest predicted situation type by the conditions, except for Behavioural (0%) and Existential (44%) process type. For definiteness, little difference was shown between the behaviour of definite and indefinite instances in relation to situation type expression. For process type, a lack of variation was also largely shown in prediction rates. However, we can see there is some difference in the rates to

which process types predict Accomplishment and Activity, with certain process types predicting higher percentage rates for Activity, and others for Accomplishment. For instance, while Existential (49%), Material (22%), Mental (28%) and Relational\_I (26%) all produce higher prediction rates for Activity, Behavioural (100%), Relational\_A (21%) and Verbal (31%) all generated higher prediction rates for Accomplishment. Example instances of the Object (59a), Accomplishment (59b) and Activity (59c) meanings expressed by *stream* are shown in the examples below. The semantic expression of *stream* will be discussed further in Section 7.4.

- (59) a. Officials supporting the mine purchase have said it will preserve hundreds of good jobs and a revenue stream for the tribe (Noun 2, *Stream* Dataset, Appendix C)
- b. The YouTube stream where you can watch the debate is embedded below. (Noun 20, *Stream* Dataset, Appendix C)
- c. Click on bald eagle live stream. (Noun 150, *Stream* Dataset, Appendix C)

The MLR *stream* model returned a low Nagelkerke Pseudo  $R^2$  (0.04121647). This Pseudo  $R^2$  result is slightly higher than what the calculation produced for the MLR *game* model, but lower than what was reported for the MLR *fire* model. The MLR *stream* model also reported an intermediate prediction accuracy (58.8%), although this prediction accuracy is the lowest of the three models considered in this chapter. After conducting the Likelihood-Ratio test, it was identified that both definiteness:  $X^2(2) = 3.1664$ ,  $p = 0.2053$  and process type:  $X^2(12) = 13.4054$ ,  $p = 0.3403$  shared a non-significant relationship with the semantic behaviour of *stream*. Therefore, we cannot disregard the null hypothesis for either definiteness or process type. Such a result is again interesting as it further cements the suggestion presented in the previous section that the influence of variables like definiteness on semantic behaviour can alter depending on the nominal in question. While definiteness was shown to significantly influence the semantic behaviour of the nominal *fire*, this effect holds less influence over the semantic behaviour of *game* and *stream*.

#### 6.4.4. Drought

The last UEN that we will focus on in this chapter is *drought*. As mentioned at the beginning of Section 6.4, *drought* presents an interesting case for the exploration of temporal semantics, as its semantic structure, from the lexical aspect annotation, appears exceptionally rigid. So much so that every instance of *drought* expressed State. Because of this semantic rigidity, running inferential statistics on the dataset would most probably not assist us anymore than descriptive statistics in answering our research question for this study. Accordingly, below I have presented two tables. Table 6-10 concerns the relationship between definiteness and situation type for the nominal *drought*, and Table 6-11 relates to the relationship between process type and situation type for *drought*.

Table 6-10. Distribution of situation type across definiteness for *drought*.

Situation Type	Definite	Indefinite
State	188	312

Table 6-10 shows that, although *drought* expressed no variation in situation type, the nominal was expressed more as an indefinite expression than a definite expression. This finding is interesting as while more indefinite State instances were expressed by *drought*, there were still a considerable number of definite instances. States, as covered extensively in Section 5.2.2, are largely abstract entities that tend to be homogenous and unbounded. In Section 5.2.2, we found that count/mass status was highly influential in the expression of States due to the homogenous and unbounded qualities typically associated with mass status. Interestingly, despite the potential association between definiteness and (temporal) boundedness as noted above in Section 6.2.1, it appears that for the nominal *drought*, definiteness holds very little influence over its semantic behaviour.

Table 6-11. Distribution of situation type across process type for *drought*.

Situation Type	Existential	Material	Mental	Relational-A	Relational-I	Verbal
State	11	345	16	80	35	13



In addition to definiteness, we can see from Table 6-11 that process type holds very little influence over the semantic behaviour of *drought*. As is evident from the table above, most instances occurred within Material processes. However, as we know, in every process type, State was the only situation type expressed. Accordingly, unlike the previous nominals *game*; *fire*; and *stream*, the semantic structure of *drought* appears incredibly fixed, with no variation in semantic behaviour shown across process type. In Section 7.3, I will come back to this semantic rigidity by revisiting the nominals *fire* and *drought* to consider how understanding differences in the semantic flexibility of nominals can assist us in further understanding the nature and degree of nominality.

## 6.5. Towards an Exploration into Semantic Flexibility

In this chapter, we have seen that, interestingly, process type did not share a dependent relationship with the semantic expression of any of the four UENs. Moreover, we saw that the influence of definiteness on semantic expression of the UENs varied, as while definiteness did not share a dependent relationship with the semantic expression of *game*; *stream*; and *drought*, it did share a significant relationship with the semantic expression of *fire*. As discussed in Section 6.4.2, this influence was largely the result of the absence of an article in the nominal group (zero-article) partnered with the mass status of the nominal. When zero-articles and mass status co-occurred, this led to an indeterminateness between the start and endpoints of the entity in question. Due to this indeterminateness, the potential event boundaries of the nominal were difficult to identify and, consequently, resulted in the conceptualisation of fire as an atemporal substance, rather than a dynamic event.

In the following chapter, I will drop our focus on the influence of definiteness and process type on the semantic behaviour of the four UENs and turn to the lexical aspect distributions expressed by each UEN. In particular, I will assess how the nominals *game* and *stream* expressed a certain amount of semantic flexibility, while the nominals *fire* and *drought* generally did not.

## Chapter 7: Semantic Flexibility, Stability, and the Boundaries of Nominality

In the previous chapter, I explored the extent to which definiteness and process type held influence over the temporal semantic behaviour of four selected nominals: *game*; *fire*; *stream*; and *drought*. Overall, it was shown that definiteness expressed varying degrees of influence over the temporal semantic behaviour of the four nominals. Specifically, it was identified that while definiteness did not hold a significant relationship with the semantic behaviour of *game*; *stream*; and *drought*, it was influential in the semantic behaviour of *fire*. On the other hand, we also discovered that process type held little influence over the semantic behaviour of all the nominals concerned. Interestingly, from the lexical aspect annotation, it was apparent that *game* and *stream* expressed a large semantic range. This semantic flexibility contrasted with the temporal semantic behaviour of nominals such as *fire* and *drought*, which generally presented themselves as semantically stable.

In this chapter, I will explore in detail this notion of semantic flexibility and investigate why certain nominals appear to allow for a greater variation in semantic expression than others. I will first revisit the notion of meaning, and how certain lexical items are able to express different temporal semantics. After doing so, I will focus on the nominal *game*, and the relationship the lexical item shares with its various ‘microsenses’ (a term introduced by Cruse (2001), discussed in Section 7.2.1), as a result of its hypernymic structure. Moreover, I will direct our attention to how the notion of ‘rules’ can affect the semantic structure of the nominal *game* and how our conceptions of ‘lexical facets’ (described in Section 7.2.2) can also impact this structure. When discussing rules in this chapter, I will use the term ‘Human Imposed Parameters’ (HIPs), to emphasise the influence human agency seemingly holds over the expression of temporal semantics by the four nominals. The chapter then will turn to nominals that express semantically stable temporal structures (such as *fire* and *drought*) and cover how a lack of association with human agency influences their meaning potential. In addition to this discussion on the relationship between semantic flexibility and human agency, I will focus on the nominal *stream* and how semantic extensions driven by human agency have increased its semantic flexibility. The final section of this chapter will provide a summary of our discussion and lead us into the concluding chapter.

## 7.1. Meaning Potential and Semantic Structure

As has been discussed many times throughout this thesis, meaning presents a particularly complicated area for linguists. There is an uncertainty over the extent to which the meanings expressed by lexical items are influenced by the syntactic context in which they are used and the extent to which these meanings are influenced by more internal properties assigned to the lexical items themselves. While it would certainly be inaccurate to claim that meaning is completely governed by the internal structure of the lexical item, it would also perhaps be over-committal to suggest that all meaning is entirely the result of the syntactic context a lexical item enters. Evans (2010, p.21) suggests that the meaning of a lexical item cannot be discerned without both the utterance in which it is used and the subsequent encyclopaedic knowledge that the lexical item is naturally associated with. He argues that lexical items present lexical concepts (which include our encyclopaedic knowledge) and that this meaning is then narrowed when the lexical item is placed in a syntactic context. This perspective correlates roughly with what Hanks (2013) calls ‘meaning potential’ (covered in Section 3.2.2), which we will now revisit in the next paragraph.

For Hanks’ (2013) ‘meaning potential’ relates to idea that lexical items only express a meaning ‘potential’ when in isolation, and that this potential is the product of a number of components generated from our previous experience of the lexical item’s linguistic behaviour and, further, our ontological knowledge of what the lexical item can potentially represent. When a lexical item is used in context, certain components of this meaning potential are activated via contextual cues (Allwood 2003, p.29). For instance, using Hanks’ example of *bank*, if the lexical item *bank* is used when discussing the topic of money, like in the sentence *he put the money in the bank*, we can discern that the *bank* expressed here relates to a financial institution, and not a piece of land that slopes down towards a river or lake. According to Hanks (2000, p.210), meaning is a dynamic event (the coming together of a number of specific meaning components) and not a static entity. During this event, certain meaning components will be more strongly activated in a given use of a lexical item than in others.

The syntactic context, according to Hanks, restricts the meaning potential of the lexical item to a specific instance of meaning, which does seem intuitive. By detailing how various semantic components can combine, we can economically account for a diverse range in which lexical items are used (Hanks 2013, p.72). However, while Hanks does appear to consider the

influence of ontological knowledge on the meaning potential of lexical items in his discussions on implicature (2007), semantic types (2012), and exploitation (2013), this acknowledgement is not necessarily explicit. While syntactic context evidently influences the meaning expression of a lexical item, it can be argued that the semantic nature of the lexical item itself is also partly influential in the constructions it can successfully enter.

When thinking of the meanings of lexical items separate from their use in syntax, we do not just think of the lexical item's grammatical form, but our experience of what that lexical item represents and the various contexts it can potentially be used in. Taylor (2003, p.87) states that our interpretation of linguistic expressions is dependent on the activation of our world knowledge. For instance, to use Taylor's (2003, p.86) example, the concept of *dog* does not exist independently of our knowledge of dogs. It is instead highly intertwined with our understanding of dogs as a species in the animal kingdom, the different breeds of dog that exist, our knowledge of their behaviour and characteristics, along with their relationship with humans. This information helps construct our mental representation of *dog*, which in turn can affect the way the concept is expressed through language.

Considering both the syntactic behaviour of the lexical item and relevant world knowledge, each lexical item shares an association with a "network of paradigmatic and syntagmatic connections that condition a specific sense" (Polguère 2015, p.73). Moreover, because of the limitations placed on a lexical item's expression of meaning by our understanding of world knowledge, certain regions of a lexical item's semantic structure display a greater 'density' than others, and, as a result, express more stability in the face of contextual variation (Cruse 2000, p.30). This notion of stability will be revisited throughout this chapter, and will be particularly relevant in Section 7.3, where we will explore the semantic stability of the nominals *fire* and *drought*. In the next section however, I will first investigate the semantic behaviour of the nominal *game*, paying particular attention to the nominal's hypernymic structure, its association with inherent Human Imposed Parameters (HIPs) and our conception of certain 'lexical facets' as large contributors to the nominal's apparent semantic flexibility.

## **7.2. The Semantic Flexibility of Game**

In the second study of this thesis, *game* largely expressed Accomplishment situation type. However, importantly, *game* also expressed a notable percentage of Object and Activity

instances. For our current discussion on the temporal flexibility of *game*, it is the distinction between Accomplishment and Activity instances which perhaps proves the most insightful.

### **7.2.1. Hyponymy and Rules: The Difference of Accomplishment and Activity**

Shown by Wittgenstein (1953), and also discussed in Section 2.1.2, the question of what constitutes a game is notoriously difficult to answer. Aarts et al. (2004, p.18) describe the meanings expressed by lexical items like *game* as inherently fuzzy. We encounter complications when attempting to define *game*, as there are countless activities that could be understood to constitute games, despite displaying rather different characteristics (Aarts and Haegeman 2006, p.123).

One reason for this difficulty relates to the hypernymic structure of the lexical item. Hyponymy concerns the hierarchical relation between a more general lexical item, e.g., *game*, and a more specific instance, i.e., a type, of that lexical item (Jackson and Zé Amvela 2007, p.117), e.g., *football*. When expressed in the form of *game*, these specific instances act as ‘microsenses’, which are the discrete sense units of a lexical item that differ from one another, but function at the same hierarchical level (Croft and Cruse 2004, pp.126–127) e.g., contrast the microsenses of *football* and *basketball*. These discrete sense units are subsumed under the hypernym *game* and each microsense is associated with a far greater set of contextual relations than its superordinate term (Michaelis 2006, p.900). Across the *game* dataset, a large variation of microsenses expressed Accomplishment situation type. Of these microsenses, 99.2% of Accomplishment readings were related to sports related games.

For sports games that are interpretable as events, namely sports like football and basketball etc., we generally associate some sort of temporal endpoint with their semantic structure. But why is this? Games like these have HIPs, which involve, in a professional context, a time limit. Football, for instance, is played professionally for 90 minutes (at least when in a league competition), and after the addition of added time, the game is ended. The end of this game is not arbitrary, but inherent to the game itself, as it is encoded within its HIPs. Before the game is played, players, spectators and the like will all know that the game will be restricted to that amount of time. As we understand that HIPs are embedded within the very nature of professionally played sports games, a temporal endpoint is also inherent within the structure of

the nominal when used to talk about these games. Examples of Accomplishment readings for *game* are shown in (60a-b). Example (60a) relates to the game of ice hockey, while example (60b) relates to the game of American football. In both examples, the games mentioned are construed as having already reached their internal temporal endpoint.

- (60) a. The six-foot-three, 209-pound left-shot defenceman was traded from the Oilers to the Flames and has played 24 games with his new club. (Noun 16, *Game Dataset*, Appendix C)
- b. The Seahawks absolutely dominated that game, outgaining the Saints 429 yards to 188... (Noun 34, *Game Dataset*, Appendix C)

Well-defined HIPs have previously been thought of as essential for the concept of *game* (Aarts et al. 2004, p.468). However, from the data analysed in this study, well-defined HIPs are not always required for the conceptualisation of something as a game. We find this specifically in example (61) below. In what we can presume to be a slightly uncharacteristic use of the lexical item *game*, the expression of *mind games* in example (61) relates to the process of ‘dating’ and how the brain can trick one into thinking they share more commonalities with another person than they actually do. Here, these games lack structure and, as they exist in the mind, can be invoked at any moment. We do not have winners and losers in such games. One may only speak, instead, of ‘overcoming’ the *games*, as shown below. Any ending of these games is therefore arbitrary and not encoded within their semantic structure.

- (61) One of the hallmarks of emotional intelligence, which itself is key to overcoming these mind games. (Noun 52; *Game Dataset*: Appendix C).

Interestingly, the referential scope of *game* also allows for *mind games* to be played between multiple people. When multiple participants are conceptualised, *mind games* typically relate to certain strategies individuals can employ to hold interpersonal influence over another (Kellermann and Cole 1994). Kellerman and Cole (1994) identify 64 distinct strategies of how this influence can be exerted. While I do not have space to cover all these strategies here, we can see that such strategies include ‘challenging others to comply with oneself’; ‘criticising others so that they comply’; and also ‘threatening others into compliance’ (Kellermann and Cole 1994, pp.7–12).

The example of *mind games* is particularly interesting, as the playing of these games involves no HIPs and a notable lack of structure. This lack of structure results in the conception of these games as continuous, as they express no internal endpoint. The desired goal of such games resides entirely within the heads of the participants. It can be suggested that our conception of mind games as a type of game stems from their somewhat competitive nature and the fact that they are ‘played’ by (or at least involve) human(s). When multiple participants are considered, winning in these types of mind games essentially relates to whether one participant can successfully manipulate the behaviour of the other. If the other does alter their behaviour in accordance with what the one who initiated the mind games desires, then the participant who initiated the mind games has technically won. However, as stated above, these games may be invoked and continually re-invoked, and thus, never truly express a natural temporal endpoint.

Evidently then, there is a natural lack of telicity that can be identified in the meaning potential of *game*. Whether this atelicity is expressed or not appears dependent upon the type of game being played, and whether that game inherently enforces a specific set of HIPs. Accordingly, I suggest that it is the presence/non presence of HIPs in the semantic expression of *game* that is largely responsible for the semantic flexibility of the nominal between Accomplishment and Activity readings. Interestingly, we can only really distinguish different types of games by observing the distributional context of the noun. These distinctions between Accomplishment and Activity readings therefore point to a major development in the field of lexical aspect for nominals, as despite the framework claiming to relate to the inherent temporal structures of the lexical item (Vendler 1967; Dowty 1979; Smith 1991; Rothstein 2004), the situation type Accomplishment may be typically more phrasal than lexical for certain nominals.

The conceptualisation of games having the capacity to be invoked and continually re-invoked is displayed by many other senses of the nominal as well. Examples (62a-b) below provide intriguing examples of *game* construed almost as a property of an individual/team that is reflective of something they continually do. For instance, in (62a), where the context relates to an ice hockey team, *game* is construed somewhat like an aspect of the team’s overall playing style, i.e., the team’s defensive game. But this aspect is based off the foundation of previous defensive displays by the team. This defensive game has a continuity to it, as it can be invoked and carried out at any time by the team in future matches. If the team were to change defensive strategy, then the team’s future defensive game would also alter.

In another example in (62b), we are told to ‘spice up’ our *burger game*. In the similar vein to example (62a), our burger game is construed as a continuous property, relating to our ability to make burgers. This *game* is established on all our previous instantiations of burger making, and how successful they were. Example (62b) is construed as an Activity as it suggests that our burger game is something continuous, that almost remains dormant until drawn upon by an agent, i.e., when they make a burger. Example (62b) also suggests that future instantiations of our burger game can be improved through the addition of specific condiments and sauces. Because of this supposed need for improvement, we also see an element of competition attached to our burger game, even if that competition is only with ourselves to improve burger style and strategy etc.

- (62) a. If we can find some more offensive while keeping our defensive game the same, we’ll be in good shape. (Noun 351, *Game Dataset*, Appendix C)
- b. Spice up your burger game with these tasty condiments and sauces. (Noun 102, *Game Dataset*, Appendix C)

Such a temporal structure is largely relatable to the notion of habitual aspect, which was discussed in relation to verbs in Section 2.4.1 and deverbal *-er* nominals in Section 5.3.2.1. To remind the reader, habitual aspect relates to the construal of an entity as continuously instantiated over a given time frame, as opposed to focusing on one specific instantiation (Binnick 2006, p.248). From these examples, it is apparent that the expression of habituality by nominals is not always dependent on the presence of certain morphology (e.g., *-er*) and can be expressed by nominals which are regarded as morphologically simple.

### **7.2.2. Video Games and the Conceptualisation of Lexical Facets**

In my final discussion on *game*, I will now consider how the nominal can express object meaning, and how this relates to the concept of ‘lexical facets’. From the data, *game* expressed Object 32.4% of the time, making Object the second most expressed temporal structure in the dataset. The expression of object meaning was largely due to the frequent discussion of *video games*, a use that displayed some interesting distinctions in meaning that I will now detail below.



One main reason for these interesting meanings comes in the form of lexical facets. Lexical facets, a term proposed by Cruse (2000), relate to how certain lexical items can invoke different perspectives of the same conceptual meaning. A frequent example used to explain lexical facets is the noun *book*. While an utterance noting the weight of a book relates primarily to the physicality of the book, a different utterance might describe a book as ‘exciting’, which relates to the contents of the book. Here we have two facets that are not semantically derived from one another, but together behave as a single gestalt (Abrard and Stosic 2021, p.12). Facets are essentially the different readings we can identify within a given sense, that relate to different aspects of the same concept (Paradis 2004, p.252). Facets operate differently to microsenses (which we saw earlier in this chapter), as facets represent different ontological types so cannot be classified under one hypernym (Croft and Cruse 2004, p.130). Consequently, facets are not typically differentiated by their domain of occurrence, while microsenses are (Cruse 2001, p.49).

When looking at instances of *game* related to video games in the data, the expression of two lexical facets was apparent: the expression of the physical nature of the game (this lexical facet could also be expressed as a non-physical downloadable file), and the content of the game. Alternatively, it was also identified that these facets sometimes combined to express a “global reading” of the concept (Cruse 2000, p.30). This distinction in lexical facets is shown in examples (63a-c) below. In example (63a), the facet of *game* as a downloadable file/application is expressed. The expression of this facet is clear, as the game cannot be deleted without the removal of the file/application from one’s mobile device. Example (63b), alternatively, focuses on the content facet of the game, as the writer describes the content of games included in the Grand Theft Auto series as *missions surrounded by optional objectives in a large environment*. It would not make sense to describe either a physical object or file/application in this way. Lastly, example (63c) displays a global reading of the *game* concept, as when stating that there have been nine Mario Kart games, this utterance can simultaneously relate to nine different physical/downloadable releases of the game and subsequently nine different variations of game content.

- (63) a. Once you tire of Flappy Bird you may want to delete the game... (Noun 174, *Game Dataset*, Appendix C)

b. Even open-world games such as the Grand Theft Auto series are simply story missions surrounded by optional objectives in a large environment... (Noun 7, *Game Dataset*, Appendix C)

c. But there have been, in fact, nine Mario Kart games, come along, count with us -- and take a look at some of the commercials made for them too! (Noun 439, *Game Dataset*, Appendix C).

The expression of object meaning by video game related games, especially those that express the facet of game content, is a particularly interesting area, as it points towards a difference in linguistic behaviour when compared to the other types of games that we have already visited in this section, e.g., sports games. Given that the content of a video game can be played over a period of time by an agent, we might expect the expression of temporal structure. However, this was rarely the case. Most typically, when video games were expressed, they did not relate to a specific instantiation of a given video game being played. This behaviour can be compared with expression of sports games which, alternatively, nearly always drew upon a specific instantiation. But what reason is there for this difference in semantic behaviour?

I suggest that this behavioural difference stems from the fact that the content of video games exists independently of being played, while specific sports games, such as football matches, are not conceived like this. An individual can talk about the content of a video game, without ever playing it. As a result, and as shown in example (64a) below, a considerable amount of the dataset which concerned video games centred around detailing the nature and features of video games, with *game* expressed as a non-dynamic entity. On the other hand, we appear to less frequently talk about specific sports games, such as individual football matches, without an instantiation of the game, whether that be a past, future, imaginary or even implied instantiation. Individual sports games, such as these football matches just mentioned, are largely brought into being by their players participating in the game, and, as a consequence, these games are conceptualised as events. This is not to say that sports games are always construed as events. As shown in example (64b), sports games can also be construed as Objects when they are construed as an entire concept without temporality.

(64) a. Mobile games are great in that there is very little forgiveness. (Noun 346, *Game Dataset*, Appendix C)

b. FIFA boss Sepp Blatter has vowed to make soccer Sri Lanka's most popular game by overtaking cricket... (Noun 975, *Game* Dataset, Appendix C)

From the examples we have seen throughout Section 7.2, it is evident that *game* displays a high degree of semantic flexibility. This semantic flexibility was fundamentally the result of three distinct features of *game*: (1) the hypernymic structure of the nominal and its consequent association with multiple microsenses; (2) the establishment and variation of time-related HIPs and whether they are inherent to the structure of the particular type of game; and (3) the construal of different lexical facets in the expression of video games and how our conception of these facets (game physicality/downloadability and game content) largely exist independent of instantiation. In the next section, I will now explore how nominals such as *fire* and *drought* express semantic stability in the face of contextual change.

### **7.3. Fires, Droughts, and Semantic Stability**

One of the most striking findings from the second study of this thesis was how nominals such as *fire* and *drought* expressed a high degree of semantic stability in their temporal expression. *Fire* expressed Activity at a frequency of 77.8%, Object 19.2%, State 2.8%, and Accomplishment 0.2%. Alternatively, *drought* displayed an even greater semantic rigidity, expressing States 100%. Below, I will investigate how nominals such as *fire* and *drought* are able to express this stability, while nominals like *game* (discussed in the previous section) do not display this capacity. I will first explore how these nominals are generally conceptualised as unpredictable, natural events (I should note here that not all fires occur naturally, and human initiated fires will be discussed in Section 7.3.2). Taking this information, I will then focus on how the largely unpredictable and natural nature of these events/states reflects an absence of human agency encoded within the semantic structure of the nominals, which, in turn, can result in a lack of temporal boundaries and greater semantic stability.

#### **7.3.1. Natural Events and Human Agency**

Unlike many of the event nominals we have so far come across in this thesis, *fire* and *drought* are of particular interest as they can constitute (drought especially) a type of event understood as ‘natural’, that it is induced by meteorological conditions. Examples of these natural events are shown below in (65a-b). Events related to weather have piqued the curiosity of linguists in

the past, as the participants encoded by these events are notoriously difficult to identify (Dong et al. 2020, p.2). Firstly, while our conception of meteorological events might involve participant-like entities, these entities are generally not conceived as referentially independent from the event itself (Kittilä et al. 2010, p.568). Take example (65a) for instance. While it may be the individual flames of a forest fire that are causing the damage, these flames are part of, and therefore, generally indistinguishable from the forest fire event in its entirety. Moreover, and perhaps most importantly to our current discussion, Kittilä et al. (2010, p.570) suggest that the semantic encoding of participants within both meteorological events and events that are the product of meteorological activity is hard to specify, as the cause of such events is beyond human control. These events are instead typically construed as either just transpiring or ‘happening’ to humans (Bromhead 2006, p.123). As a result of this lack of control and agency, Gross and Kiefer (1995, p.58) claim that events which construe natural disasters tend to co-occur with specific verbs, such as *devastated*, *destroyed*, etc., which frame humans as powerless witnesses to the event. An instance of a natural event occurring with verbs such as these is shown specifically in example (65b), where a drought *ravages* a village. In this instance, the drought and its impact are construed as out of the control of the village’s inhabitants.

- (65) a. For instance, there were 36 incidences of forest fires in 2010 alone, which burned more than 9,162.81 acres of forests. (Noun 220, *Fire Dataset*, Appendix C).
- b. As is the case every few years, drought ravages a village. (Noun 152, *Drought Dataset*, Appendix C).

We touched on the concept of human agency above, and how it clearly plays a role in how we conceptualise events. Olsen (1994, p.375) describes human agency as the capacity of humans to effect change, related to the notion of control. Agency itself is inextricably linked with the notion of time, as we, as humans, do not have the capacity to cause change just like that (Ferrero 2022, p.329). On the contrary, one of the central characteristics of human agency is that humans can carry out activities for long durations, perhaps even for a lifetime (Ferrero 2022, p.329). Up until this point, agency has typically been explored in the verb domain. Formanowicz et al. (2021) suggest that agency is associated with prototypical verbs, as prototypical verbs generally express actions (Vigliocco et al. 2011) and actions typically require agency for their instantiation.

Relating to lexical aspect, Kearns (2011, p.168) claims that agency and lexical aspect do not appear to be entirely independent of each other. As discussed in Section 2.4.2, the expression of agency is typically understood to mark a dividing line between events and states, as Accomplishments, Achievements and Activities are understood to require an input of energy to occur (Declerck 2006, p.51). However, within the nominal domain, especially shown through specific examples of *fire*, this requirement of agency is not always apparent. Forest fires caused by meteorological activity do not express any human agency within their semantic structure, yet they are conceptualised as dynamic unfolding events that occur over time. The inherent temporal structure of fires such as these is continuous, as fires can theoretically last as long as supporting conditions (such as fuel and oxygen) allow. As a result, there is no exact preconditioned ending to fires, and any successful human attempts to extinguish these fires result in an arbitrary ending, as it is not inherently encoded within the event's temporal structure.

### 7.3.2. Accidents, Arson, and Atelicity

Focusing specifically on the nominal *fire*, fires can also be directly caused by humans. A human may light a fire, whether that be on purpose or by accident. I note this contrast in examples (66a-b). While in (66a), the fire event is construed fortuitously as the accidental result of a misplaced cigarette, the fire events in (66b) are construed as the deliberate product of arson. In instances such as these, human agency is involved in the root cause of the event. However, despite this notion of human agency, there is still a notable lack of control attributed to humans after the event has been initiated. To pick up on a comment from Bromhead (2006, p.123) mentioned earlier in this section, events like large fires are typically construed as 'happening' to humans, which suggests a lack of human control. A fire may be started by a human, but then become too large and dangerous for the instigator to control. In such events, destruction is typically caused, and professional fire brigades are required to extinguish the fire. However, as discussed above, the endpoint of a fire remains largely arbitrary, and tends not to be encoded in the nominal's temporal structure. At the point that a fire is started, we are generally unable to say definitively the exact time that the fire will end.

- (66) a. Authorities said the men lived in the home without electricity, so officials believe the fire may have started because of a cigarette. (Noun 408, *Fire* Dataset, Appendix C).

- b. The now 18-year-old accused pleaded guilty to one count of arson endangering life, admitting he set two separate fires at the hospital... (Noun 326, *Fire Dataset*, Appendix C)

This semantic behaviour brings us back to the semantic stability of nominals such as *fire* and *drought*. I propose that the lack of human control associated with our conception of these events/states affects the influence surrounding syntax can have on the temporal structure expressed by each nominal, as we are aware of that no agents are present to control the event/state and induce change. To best explain this point, I will compare the semantic behaviour of *fire/drought* with the nominal *game* (discussed in Section 7.2) and explore how the semantic expressions of each nominal behave with the addition of temporal pre-modifiers in the nominal group.

First, I will examine the semantic behaviour of *fire/drought* when used in the construction *the one-month fire/drought*. When used in this construction, we understand that the fire/drought in question lasted one month. However, the semantic structure expressed by *fire/drought* is still construed as atelic, despite the influence of the temporal pre-modifier *one-month*, which tries to syntactically coerce an inherent endpoint. Our temporal understanding of *fire/drought* in this instance is not altered by the presence of the pre-modifier, as we associate a lack of human control with the continuity of these events/states. The continuity of these events/states is established instead by natural forces, which leads to an atelic reading, where any ending is considered arbitrary.

We can then contrast this semantic behaviour with the nominal *game* in the syntactic construction *the two-hour game*. In this instance, the temporal pre-modifier can coerce an inherent temporal endpoint in the construal of the nominal, as it is not unreasonable to expect a game to inherently last two hours. This coercion in temporal semantics largely points towards the association of games with inherent HIPs, which was discussed in Section 7.2.1. But what inherent HIPs also point to is a conception of human agency and control over certain games and the human capacity to enforce these HIPs, e.g., an agreed time limit. As an apparent result of our conceptions of human agency in relation to these nominals, the (a)telicity expressed by *fire* and *drought* tends to be far more rigid in its semantic expression when compared to nominals such as *game*, where (a)telicity signals a greater susceptibility to semantic coercion

when used in different syntactic contexts. This behaviour is incredibly important in our discussion of the nature and degree of nominality because human agency is a lexical variable and not syntactic. The influence of human agency over the semantic expression of nominals shows that, contrary to the view of lexical contextualism (that lexical items do not possess any meaning of their own and are governed entirely by their context of use (Nowak 2021, p.217)) our real-world knowledge of the phenomena nominals represent also affects nominal semantic behaviour and receptibility to semantic coercion.

In the next section, I will continue with our discussion of semantic flexibility, but will explore how our conceptions of human agency with certain nominals can lead to semantic extensions in meaning. In doing so, I will visit the last of the four nominals in the second study: *stream*.

#### **7.4. Stream: Data Transfer and Contextual Dependency**

Within the dataset, *stream* expressed Object at a frequency of 58.4%, Activity 21.8%, and Accomplishment 19.8%. As we can see from this lexical aspect distribution, Object was notably the most frequent temporal semantic behaviour expressed. Diving deeper into the dataset, it is apparent that this tendency to express object meaning was driven largely by two different types of stream: (1) water-based streams, relating to a continuous flow of liquid; and (2) monetary-based streams, a metaphoric extension of water-based streams, associated with sources of finance. This distinction in meaning is displayed below in examples (67a-b). However, whilst Object was the main temporal meaning expressed, it is the considerable amount of Accomplishment and Activity expression by *stream* that is most interesting to this thesis, which will now be discussed in the following paragraphs.

- (67) a. I loved fishing but the stream at the bottom of our garden was always black apart from the last week in July... (Noun 135, *Stream* Dataset, Appendix C).
- b. He has multiple income streams; his music, partnership with Effen Vodka, SMS headphones and stake in Frigo underwear all help him rake in money. (Noun 186, *Stream* Dataset, Appendix C).

Activity and Accomplishment readings were largely due to semantic extensions of water-based streams. In a similar vein to monetary-based streams noted above, i.e., a metaphoric extension

related to the flow of money, a considerable number of Accomplishment and Activity readings of *stream* expressed the transfer of data, in the form of live stream video and audio content. A distinction in Accomplishment and Activity readings is available below in examples (68a-c). Examples (68a-b) express Accomplishment readings, while example (68c) expresses an Activity reading.

- (68) a. When is the 2017 CONCACAF Gold Cup and how can you watch the live stream?  
(Noun 699, *Stream* Dataset, Appendix C)
- b. The EP generated more than 2.5 million streams in the first month... (Noun 592, *Stream* Dataset, Appendix C).
- c. Click on Bald Eagle Live Stream. (Noun 150, *Stream* Dataset, Appendix C).

The distinction between Accomplishment and Activity shown in examples (68a-c) is intriguing, as *stream* appears to behave almost like a ‘shell-noun’. Shell nouns are nouns which express the capacity to be used as conceptual ‘shells’ (i.e., containers) for complex information (Schmid 2000, p.4). This capacity is typically shown in nominals such as *fact*, *problem*, *idea*, etc., where we can only understand the full meaning of each nominal after taking into account a typically accompanying proposition, e.g., *the fact that sea levels are rising*. While the nominal *stream* does not have the full capacity shown by nominals that we would typically regard as shell-nouns, it is evident that the temporal structure of the nominal is highly dependent on certain surrounding syntactic information, i.e., the thing being streamed. For instance, example (68a) is regarded as an Accomplishment, as it is a live stream of the 2017 CONCACAF Gold Cup, which is an international football competition played by North American, Central American, and Caribbean nations. The competition has a natural endpoint, as once a team has won the competition and the final game has been played, the competition will end, until it returns two years later. As the competition expresses a natural endpoint, so does the stream of the event; there is no content left for the stream to present once the competition has finished. Consequently, the stream of the event is inextricably tied to the event being streamed. This semantic behaviour is also shown in example (68b), which instead relates to audio content. A stream of an E.P. can only last the length of the E.P. Once the E.P. finishes, the event has ended. As a result, (68b) is also analysed as an Accomplishment, as the event is dynamic and durative, but is restricted temporally by the entity being streamed.



However, we can see in example (68c) that the entity being streamed does not always impose temporal restrictions on the semantic structure expressed by stream. Certain streams, such as *the bold eagle live stream* in (68c), are continuous, and have no established endpoint. The purpose of the stream is to continuously monitor bold eagle activity and enable internet users the opportunity to watch the stream whenever they please. The stream may end after a certain time frame, e.g., one month, but this ending is not encoded inherently in the stream event. The endings of such events are therefore interpreted as arbitrary. Accordingly, streams such as these were analysed as Activities. What we can glean from this behaviour is that within the nominal domain, there may be almost a continuum in relation to meaning generation when expressing events, with certain nominals more contingent on surrounding syntax for their meaning, and others less so. On one side of the continuum, we find nominals such as *stream*, which are vastly influenced by the syntactic constructions they enter, while on the other side we find nominals such as *fire* and *drought* which express an element of resistance to the influence of syntax on semantic expression.

## 7.5. Chapter Summary

Over this chapter, I have investigated the semantic behaviour of the four nominals chosen for the second study of this thesis. This section will now briefly return to the main points of the chapter and will discuss them in relation to our exploration of nominality and its evaluation.

The second study of this thesis marked a necessary shift in focus, from considering properties such as abstract/concrete status, count/mass status and word formation type to focusing more on the role of syntax in semantic expression and how the syntactic behaviour of nominals is established. From the results of the multinominal logistic regression analyses (presented in chapter 6), it was evident that syntactic variables such as definiteness (aside from its influence on the semantic behaviour of *fire*) and process type generally held little influence over the semantic behaviour of the four nominals under investigation. However, what was found in this study was that certain nominals expressed a greater amount of semantic flexibility than others, as evidenced by their expression of lexical aspect.

As discussed in Section 7.2 above, *game* posed a very interesting case in our understanding of nominality, as the nominal was able to construe a range of diverse meanings. This diversity in

semantic expression was largely due to the hypernymic structure of the nominal and its association with various microsenses (e.g., football; basketball etc.) that were established largely on the conception of different HIPs. It was uncovered that without rules, the nominal *game* appears to have a natural relationship with Activity expression (for instance, *mind games*; *burger game*), and it is the enactment of these HIPs that is the primary driver in the distinction between the expression of *game* as an Accomplishment or as an Activity. As this distinction between different types of games is principally reliant on the distributional syntactic information surrounding the nominal, this finding interestingly points towards a large development in the field of lexical aspect which is that, despite the framework claiming to relate to the inherent temporal structures of the lexical item (Vendler 1967; Dowty 1979; Smith 1991; Rothstein 2004), the expression of Accomplishment may be more phrasal for certain nominals than lexical.

After focusing on *game*, the chapter then moved to discussing the notion of semantic stability in the nominal domain and its apparent association with human agency. In the second study of the thesis, the nominals *fire* and *drought* expressed a large amount of Activity and State situation type respectively. *Fire* and *drought* were of particular interest, as they largely expressed natural events which implied either no human agency in their instantiation, or, at the very least, a lack of human control in their continuation. Because of this lack of control, both *fire* and *drought* frequently construed a lack of telicity, with endings resulting from human intervention (e.g., extinguishing a fire) largely considered arbitrary and not encoded within their inherent temporal structure. Intriguingly, associations of human agency were a primary driver in the ability of syntactic constructions to coerce certain semantic readings from these nominals.

This finding was evident from the difference in influence constructions involving temporal pre-modifiers (contrast *the one-month fire/drought* with *the two-hour game*) had over the semantics expressed by *fire/drought* and *game*. For nominals such as *fire and drought*, temporal pre-modifiers held little influence over their semantic behaviour. The lack of human control encoded in their semantic structure construed the atelicity of these events/states as the result of natural forces which seemingly rendered any endings to fires/droughts as arbitrary. On the other hand, we saw that temporal pre-modifiers were able to coerce a telic reading from the nominal *game*. We know from our ontological knowledge that games can have HIPs, such as a time limit, and that these HIPs require a degree of human agency for their enactment. To

summarise this point, it is evident that human agency was highly important variable in relation to the extent to which the four nominals were susceptible to temporal semantic coercion from their surrounding syntax. While nominal events/states which involve human agency appeared more semantically flexible, nominal event/states that did not require human agency were seemingly more stable in their expression of temporal semantics. As discussed above in Section 7.3.2, such a finding is important in our exploration of nominality, as it opposes the view of lexical contextualism and suggests an interplay between our real-world knowledge of phenomena and surrounding syntactic context in the semantic expression of nominals.

The final nominal that was discussed in this chapter was *stream*. Event readings of *stream* were certainly interesting, as they commonly related to the transfer of data, which meant that the temporal structure of these streams was entirely dependent on the type of data being ‘streamed’. In these cases of data transfer, *stream* did not express a temporal semantic behaviour of its own but adopted the temporal semantic behaviour of the event being streamed. The distinction in Accomplishment and Activity was generated on whether the stream was of an event with a natural endpoint, such as an international football competition, or an event that was continuous, such as the constant filming of bird activity. These Accomplishments and Activities mark an interesting contrast in the nominal domain with previously discussed nominals like *fire* and *drought*. While the meanings of *stream* are highly reliant on the syntactic constructions they enter when expressed as an event, nominals like *fire* and *drought*, as discussed above, are not as reliant.

What this distinction in semantic behaviour again suggests is that nominals that express human agency in some capacity are more dependent on surrounding syntax for their expression of meaning than nominals that do not express human agency. Accordingly, from the semantic behaviour of the four selected nominals, there appears to be almost a continuum in relation to meaning generation, with certain nominals more contingent on surrounding syntax for their meaning, and others less so. Previous work (Halliday and Hasan 1976; Francis 1986; Schmid 2000; Flowerdew 2003; Jiang and Hyland 2015 and many others) has shown that certain nominals, like shell-nouns, are heavily reliant on their syntactic surroundings for their semantic expression. However, this research has not investigated more ‘general’ nouns and the influence lexical features, such as human agency, may have on their syntactic dependency for meaning expression. Future research into this area that tries to quantify this distinction in a continuum-based model would contribute greatly to our understanding of nominality, as it would provide

some clarity as to how individual nominals interact with syntax and how they are able to express temporal semantics.

It is here that we reach the end of this chapter, which also marks the end of the second study in the thesis. In contrast to the first study of this thesis (see Chapters 3-5), this second study has enabled us to take a closer look into the behaviour of four UENs and the extent the temporal semantics they express share a relationship with their surrounding syntactic context. The current chapter has built on the results of the previous chapter, with a discussion exploring the role of syntax in semantic expression and investigating how the syntactic behaviour of nominals is established.

## Chapter 8: Conclusion: Evaluating the Nature and Degree of Nominality

This thesis has set out to offer an empirical exploration of the nature and degree of nominality. In doing so, the primary aim of this thesis has been to carry out a detailed examination into nominal semantic behaviour and how different temporal meanings come to be expressed. To achieve this aim, three sub-aims (noted below) were considered in a multi-perspective investigation, incorporating two distinct studies. The first study took a wide stance on the structure of nominality, focusing on nominal category features (such as abstract/concrete status; count/mass status; word formation type; and genre) and how their influence and interactions with one another played a role in the semantic expression of nominals. The second study, alternatively, took a smaller-scale approach, focusing specifically on four UENs (*game*; *fire*; *stream*; and *drought*) and the extent to which their temporal semantic expression shared a relationship with surrounding syntactic context (definiteness; process type of the clause).

Sub-aims of thesis:

1. To determine how state and event meaning come to be expressed in nominal forms.
2. To examine the potential relationships between the syntactic and semantic properties of underived event nominals (UENs).
3. To determine how the nature and degree of nominality can be evaluated in English.

This final chapter comprises the main findings from each study and considers how they refine current understandings of nominality and nominal semantic behaviour. Section 8.1 will cover the first two sub-aims of the thesis, by focusing on event and state meaning in the nominal domain and the temporal capacities of UENs. Section 8.2 will then tackle the last sub-aim of the thesis, by outlining how our findings can assist in our evaluation of nominality. Afterwards, Section 8.3 will address limitations of the thesis and Section 8.4 will briefly outline areas of future research, with a specific focus on agency in the nominal domain.

## **8.1. Events, States and UENs**

In this section, I will consider how event and state meanings come to be expressed by nominals and, further, how UENs express temporal meaning. In doing so, Section 8.1.1 will first cover how the importance of nominal category features and their interactional capacities seemingly play a crucial role in the expression of temporal meaning.

### **8.1.1. Nominal Feature Importance and Interaction**

It was discussed in Chapter 2 that current conceptions of lexical classes such as noun and verb are “unsatisfactory”, due to their dependence on a mixture of morphological, syntactic and semantic criteria which do not necessarily correspond in particular instances (Lyons 1977, p.423). While current theories of categorisation, such as Prototype theory, have been beneficial in capturing our pre-theoretical intuitions that some members of a given class are better examples of that class than others (Geeraerts 2006, p.144), it was discussed that work was still required to address a number of issues that arise when trying to accurately categorise members of a lexical class. Three specific issues were particularly relevant to the first study of this thesis and our aim of understanding how event and state meaning is expressed in nominal forms. These issues were: (1) varying category feature importance; (2) how category features interact in the establishment of prototypicality; and (3) the influence genre (context) holds on what is considered prototypical. In this section, I will use my findings to address these issues and show how nominals express event and state meaning.

The first study of this thesis was principally conducted to investigate the extent to which abstract/concrete status, count/mass status, word formation type and genre held influence over the distribution of semantic behaviour displayed by nominals. The study also explored how the influence of these variables on semantic behaviour differed, and the extent to which variable interactions affected semantic behaviour. In doing so, Fisher-Exact tests and Cramer’s V tests were conducted to analyse the relationship each variable shared with situation type expression. Moreover, a Classification and Regression Tree (CRT) analysis was undertaken to display whether there were any significant feature interactions that were influential in situation type expression.

The outcome of this study identified several interesting findings. The first was that abstract/concrete status, count/mass status, word formation type and genre each reported

significant relationships with situation type expression. However, when these predictor variables were used in the CRT analysis, genre, unlike the three other predictor variables, did not contribute significant splits in the data. Of the predictor variables that did produce significant data splits, abstract/concrete status held the strongest influence over semantic behaviour. This influence was reflected by the somewhat intuitive pattern that events and states were more likely expressed by abstract nominals, while objects were more likely expressed by concrete nominals. Such a result is interesting and important in our journey to understanding the nature and degree of nominality, as it presents two key insights. The first is that the distinction between abstract/concrete appears to be conceptual, rather than grammatically encoded. Accordingly, this influence suggests that our ideas about the phenomena nominals represent is important in nominal semantic expression.

The second insight is that certain nominal features clearly hold a greater strength over semantic behaviour than others. When attempting to categorise nominals on the prototypicality of their semantic expression, we must consider that certain nominal category features express a stronger association with prototypical nominal semantics. Consequently, when exploring gradience, whether that is within a lexical class (subjective gradience) or between lexical classes (intersective gradience) (Aarts 2004b, p.6), it may not simply be the number of nominal features expressed that is central to a nominal's conception as prototypical (linking back to Croft and Cruse's (2004) critique of over-simplified lists in the construction of prototypicality), but, more complexly, which features are expressed. A nominal that expresses a small amount of highly influential category features tied to prototypical nominal semantic behaviour may be regarded as more prototypical than a nominal that expresses a broad range of category features that are not regarded as influential in the establishment of prototypical semantics.

Staying on the effects of abstract/concrete status, the first study of this thesis also revealed that count/mass status held a stronger influence over the semantic expression of abstract nominals than concrete nominals. This pattern was largely driven by the orientation of abstract mass nominals towards State situation type. As discussed in Section 5.2.2, the behaviour of these nominals is supportive of the suggestion that nominalised mass nominals share a strong relationship with Activity and State situation type, due to the effect of their unbounded nature on the expression of a temporal endpoint (Mourelatos 1978; Brinton 1991). Moreover, this result also shines light on the suggestion by Cruse (2011, p.64) and the findings from Medin

and Shoben (1988) that category feature interactions are prevalent in our conceptualisation and subsequent categorisation of entities.

In this research, the semantic behaviour of nominals was the product of multiple feature interactions, including abstract/concrete status, count/mass status and word formation type. For instance, the semantic distribution of abstract mass nominals was significantly different to that of abstract count nominals. Therefore, when categorising nominals on the prototypicality of their semantic expression, it is not enough to simply explore which category features they express. Instead, we must also investigate where interactions can occur between features and how these interactions are responsible for the production of distinct semantic behaviour. By taking into consideration feature interactions, we can more accurately view category features as dynamic entities which are active in the construction of meaning, as opposed to stable entities which form a ‘checklist’ of prototypicality. In the next section, I will now address the influence of word formation type on the semantic expression of nominals. I will specifically focus on nominals that are morphologically derived from adjectives and the event-creating capacity of nominals.

### **8.1.2. The Influence of Word Formation on Nominal Semantics**

In addition to the influence of abstract/concrete status and count/mass status, we saw that word formation type also held a considerable influence over semantic behaviour. This influence was predominantly the product of three specific word formation types: nominals morphologically derived from adjectives (MDA); nominals morphologically derived from verbs (MDV); and transcategorised nominals (TC). Of the three types, MDA produced the most semantically divergent distribution, differing significantly from every word formation type, even when Object instances were removed from the data. The primary reason for this significant difference was the large proportion of States expressed by MDAs, as this was not replicated by any other word formation type. This finding somewhat plays to our intuitions, given that adjectives primarily function to express states or qualities. In line with the work of Roy (2010), it was somewhat expected that this capacity would be retained to an extent during the nominalisation process. However, what was even more interesting was the capacity of MDAs to express events.



As shown in Section 5.3.1.2, the MDA nominal *business* expressed event semantics, as it represented something that could be ‘done’ and could be observed ‘over a period of time’. In line with the work of Arche et al. (2014; 2021), it was suggested that certain nominals express events when they are nominalised from an evaluative adjective base, such as *busy*. These evaluative adjective bases can predicate sentient individuals and event subjects (Stowell 1991, p.106), that include a ‘covert’ process description in their semantic structure. For instance, we can say that *John was busy*. When nominalised, nouns with an adjectival base can embed a predicative phrase (Roy 2010) whose subject is a dynamic event expressing “no time or world parameters” (Arche et al. 2021, p.33). Consequently, the deadjectival nominalisations are somewhat able to inherit the eventive properties of the predicative phrase they embed, which allows for the expression of an event (Arche et al. 2021, p.33). Therefore, *business* has the potential to express any event or a series of events in which a participant can be described as *busy*. Fundamentally then, the expression of event semantics by MDAs concerns how nominals can capture the dynamism of participants, i.e., (human) agency, during the nominalisation process, and how this agency assists in the creation of temporal meaning. We will visit this notion of agency again in greater detail in our discussion of UEN semantic behaviour (Section 8.1.4).

The capacity of MDAs to express event semantics is certainly intriguing as it shows that, contrary to the popular association of nominality with stability and atemporal semantic expression (Givón 2001, p.51; Halliday and Matthiessen 2014, p.236), the nature of nominality, given supporting circumstances, can actively play a large role in the original production of event semantics. Such a finding requires some reflection on our current understanding of nominality, as (along with UENs as discussed in Section 8.1.4) it directly disputes the suggestion that the expression of lexical aspect is a verbal trait (Vendler 1967; Dowty 1979; Smith 1991; Declerck 2006). Regardless of lexical class, if a lexical item expresses temporal meaning, it will also express lexical aspect features (Huyghe et al. 2017). Accordingly, findings such as these problematise predominant understandings of nominality, as the current framing of the lexical class tends to vastly overemphasise the time-stable nature of a selection of nominals (Hopper and Thompson 1984a; Croft 1991; Langacker 1991; Givón 2001; Halliday and Matthiessen 2014) and provides very little to no attention to the capacity of the lexical class to act as an event-creating resource. In future, greater recognition should be given to how nominals are able to capture notions of (human) agency in their meaning during the nominalisation process, and how this assists in the expression of event semantics. In the next

section, I will now explore the parameters around which lexical aspect features are expressed (e.g., the notions of boundedness and homogeneity) and how the relationship between these parameters is different for nouns and verbs.

### **8.1.3. Boundedness, Homogeneity and Telicity in the Nominal Domain**

As I showed earlier in Section 8.1, the count/mass distinction is an important notion when exploring nominal semantics. In Chapter 5, it was stated that count/mass status in the nominal domain is viewed as similar to the imperfect/perfect distinction in the verb domain (Mourelatos 1978; Talmy 1978; Bach 1986; Jackendoff 1990; 1991; Brinton 1991; 1998; Kearns 2000 and many others), and consequently, also understood to share ties with the behaviour of lexical aspect (Mourelatos 1978; Brinton 1995; Bauer et al. 2013). This hypothetical relationship between count/mass status and lexical aspect is principally founded on the notions of boundedness, homogeneity, and telicity. Within the verb domain, there exists a very strong implication between the notions of boundedness and homogeneity (for more information on these terms, see Section 5.2) (Huyghe 2011, p.113). For instance, verbs which are unbounded such as *running* are homogenous, as their semantic composition is the same throughout (any instance in the running event is the same event), and atelic, as they are construed as continuous without an inherent endpoint. Accordingly, as count nominals are bounded, Bauer et al. (2013) suggest that they will naturally favour heterogeneous telic expressions. Alternatively, as mass nominals are unbounded, it is proposed that they will display a greater gravitation to homogenous atelic expressions (Mourelatos 1978; Brinton 1995). From the results of the study, this relationship generally held, as Accomplishments and Achievements were more likely to be expressed by count nominals, and States were more likely to be expressed by mass nominals. However, this correspondence was interestingly not always adhered to, particularly by Activities.

From the results of the first study, the relationship between boundedness, homogeneity and telicity operated differently in the nominal domain when compared to the verb domain. Activity nominals such as *adventure*; *attack*; and *battle* expressed homogenous atelic events, despite being bounded by their countable grammatical structure. Such a finding acts to confirm research by Barque et al. (2009), Haas and Huyghe (2010), Huyghe (2011) and Grimm (2014) that bounded readings are not contradictory to homogeneity in the nominal domain. In addition to Activities, certain nominal States also expressed the capacity to express atelicity and

homogeneity, despite being bounded by count status, such as *hopes*; *obsessions*; and *sympathies*. The presence of count States in the data was intriguing as it shows that non-dynamic situations, in addition to dynamic situations, can afford bounded interpretations in the nominal domain.

The lack of correspondence between boundedness, homogeneity and telicity expressed by certain nominals lays down an important milestone in our empirical investigation of nominality. Work, such as Croft (1991); Givón (2001); Sasse (2001); and Vinson and Vigliocco (2002) among others, has suggested that there exists a continuum between nouns and verbs in relation to their semantic behaviour. The above findings from study one complement this research by suggesting that while semantic notions such as boundedness, homogeneity and telicity can easily be transferred between verb and nominal domains (i.e., a sharing of semantic space), each lexical class presents different semantic constraints on how these concepts behave. Interestingly, it is the verbal domain that imposes stricter rules that act to align these concepts of boundedness and homogeneity when verbs are expressed. On the other hand, the nominal domain presents a more flexible alternative, where the correlation of these concepts, while somewhat adhered to, is not compulsory. The relationship between boundedness, homogeneity and telicity therefore draws a dividing line between the verb and nominal domain and constitutes a semantic boundary between the two lexical classes that can be empirically identified.

In the following section, we will now visit the semantic behaviour of underived event nouns (UENs). In doing so, I will consider how UENs can express event meaning and the extent to which conceptions of (human) agency associated with these nominals contribute to this semantic behaviour.

#### **8.1.4. The Semantic Behaviour of Underived Event Nouns**

As discussed earlier in this chapter, but also more comprehensively in sections 2.1.4; 2.4.4 and Chapters 6 and 7, UENs have so far posed a problematic area for prototype models of nominal categorisation, as they exhibit prototypical nominal forms, while expressing temporal semantics. This behaviour is particularly interesting given that nominals, especially those which are morphologically underived, share a popular association with atemporal semantic expression and notions of time-stability (Givón 2001, p.51; Halliday and Matthiessen 2014,

p.236). As a result of their ability to express temporal semantics, UENs are often ambiguously categorised somewhere in the middle of the continuum of nominality without an explicit 'home'. The middle space in prototype models has often been a source of regular critique. While it does allow us to categorise instances such as UENs as nouns, it is generally unclear what this space actually constitutes (Geeraerts 2006, p.149). In addition, it is also largely unknown how UENs can express event semantics, given their morphologically underived status. Accordingly, the second study of this thesis was dedicated to exploring the semantic behaviour of four UENs (*Game*; *Fire*; *Stream*; and *Drought*), and how these nominals were able to express temporal meaning.

The second study used multinomial logistic regression modelling to identify the extent to which definiteness and process type were adequate predictors of the situation types expressed by the four UENs. Using the likelihood-ratio test, it was identified that definiteness was largely uninfluential on the semantics expressed by *game*, *stream*, and *drought*, but did hold influence over the semantics expressed by *fire*. Moreover, process type was shown to be uninfluential in the semantic expression of each nominal. From the above results, the only significant result recorded was the influence of definiteness on *fire*. As shown in Section 6.4.2, this influence was largely due to the effect of zero-articles (the complete absence of an article preceding the head noun in the nominal group) in the construal of *fire* as an atemporal substance.

When articles were absent in the nominal group, the conceptual boundaries of the entity were given less focus and, thus, attention was directed away from the composition of the entity (Radden and Dirven 2007, p.73). Furthermore, when zero-articles are placed with mass nominals, such as *fire* in these instances, they can create an indeterminateness between the start and end points of the entity in question (Maekelberghe and Heyvaert 2016, pp.320–321). Because of this indeterminateness, potential event boundaries were difficult to discern, and, consequently, these instances of *fire* were conceptualised as a static substance rather than a dynamic event. Accordingly, from these findings, I suggest that variables such as definiteness can express different degrees of influence on semantic behaviour, dependent on the lexical item in question.

This emphasis on the semantic behaviour of individual lexical items was carried into Chapter 7, where the situation type distributions of each UEN were explored and discussed in relation to nominal semantic flexibility. Through exploring these distributions, it was apparent that

while nominals such as *game* and *stream* expressed a flexibility in their semantic expression and a potential for semantic coercion from surrounding syntax, nominals such as *fire* and *drought* were far more semantically stable and were not as receptive to semantic coercion. After a deeper examination of the nominal's semantic behaviour, I proposed that the semantic flexibility of the nominals shown and their receptiveness to semantic coercion was the result of several elements which each shared an association with the notion of human agency.

The first UEN, *game*, expressed a variation of semantic output. The most interesting semantic behaviour exhibited by *game* was the distinction in Accomplishment and Activity readings. This distinction was largely the result of whether the game in question was conceived as having human-imposed parameters (HIPs), i.e., rules, or not. Games that were structured by HIPs expressed an inherent endpoint (e.g., a game of football is typically 90 minutes), whereas games that were not structured by HIPs (such as mind games) did not. Due to this variation, I proposed that that these meanings were fundamentally dependent upon the surrounding syntactic context for their expression and not the result of the internal semantics of the lexical item, as it was only through distributional evidence that the specific type of game in question could be interpreted. A similar behaviour was recognised for the UEN *stream*, where information streams were highly dependent upon their surrounding syntactic context for their temporal semantic expression. While a stream of a football competition will naturally end when the competition ends (there will be no information left to stream), a constant live stream of bird activity has no inherent ending and therefore lacks telicity. These findings interestingly point towards a major development in the field of lexical aspect, as despite the framework claiming to relate to the inherent temporal structures of the lexical item (Vendler 1967; Dowty 1979; Smith 1991; Rothstein 2004), it is apparent from the behaviour of *game* and *stream* that the situation type Accomplishment may be more phrasal for certain nominals than lexical. This proposal was further supported by the findings on the semantic stability of *fire* and *drought*, which I will outline below.

The nominals *fire* and *drought* presented a great difference in their semantic behaviour when compared to the other two UENs, as their expression of temporal semantics was highly stable. *Fire* and *drought* were somewhat unique, as the temporal semantic behaviour they expressed generally did not require human agency for its manifestation and continuation. Previous research has suggested that the semantic encoding of participants in natural events is ambiguous, as the cause of such events is typically beyond human control (Kittilä et al. 2010,

p.570). Natural events such as fire are typically construed as either just transpiring or ‘happening’ to humans (Bromhead 2006, p.123). Now, interestingly, because of the naturality of these events and the subsequent lack of human control, any ending of a *fire* or *drought* is viewed as arbitrary and not encoded within the temporal semantics of the nominal. For instance, a fire will continue so long as supporting conditions (such as fuel and oxygen) allow, and any successful human intervention to stop the fire is not encoded within the temporal structure of *fire*. As shown in Section 7.3.2, perhaps the most efficient way to highlight differences in semantic flexibility is by looking at the semantic behaviour of different nominals when placed in a nominal group with temporal pre-modifiers (e.g., one-month; two-hour; etc.).

Looking at the nominal group *the one-month fire/drought*, we understand that the fire/drought lasted one month, however the semantic structure of the nominal is still construed without a temporal endpoint. Our temporal understanding of fire/drought is not altered by the presence of the pre-modifier, as we associate a lack of human control with the durative nature of these events/states. The absence of human control construes the lack of telicity by these events/states as the result of natural forces, which tends to render any endings to fires/droughts arbitrary. We will generally not know for definite how long a fire/drought will last before it has started. We can only accurately say *the one-month fire/drought* after the event has reached/surpassed this temporal milestone. Alternatively, in the construction *the two-hour game*, the temporal pre-modifier coerces an inherent temporal endpoint in the construal of the nominal, as it is not unreasonable for a game to inherently only last two hours. This coercion in temporal semantics largely relates to the association of games with HIPs and their active enforcement by human agents.

This distinction between semantically flexible and stable nominals leaves a lot to unpack for our exploration of the four UENs. Firstly, if it was not already obvious, certain underived nominals clearly display the capacity to express temporal semantics, once again reiterating that the framework of lexical aspect is fundamentally associated with semantics and not lexical class (Huyghe et al. 2017). Secondly, despite its lack of coverage in the nominal domain, associations with agency (especially human agency) clearly played a large role in the expression of telicity by the four UENs. This relationship between human agency and telicity is principally established on the naturality of the event/state, be that how the event/state is caused or how it is continued. Natural events/states were far more resistant to semantic coercion from surrounding syntax, as their internal temporal structure is already vastly determined by

our real-world knowledge of how these events/states transpire. Alternatively, nominals which did express some sort of association with human agency were more susceptible to semantic coercion, as we are aware of the potential for human agents to exert some control over the event/state and therefore enact change (e.g., the two-hour game).

What the relationship between human agency and lexical aspect in the nominal domain points to is that our real-world knowledge of the phenomena nominals represent can be influential in the syntactic behaviour of nominals. Contrary to the view of lexical contextualism, that lexical items do not possess any meaning of their own and are governed entirely by their context of use (Nowak 2021, p.217), I have shown in this thesis that real-world knowledge can act to assist in a lexical item's meaning potential (Hanks 2013). From the findings of study two, it is apparent that the influence of our real-world knowledge on meaning potential is different for each lexical item, as certain nominals were more receptive to semantic coercion than others. In fact, I suggest that, within the nominal domain, there may be a continuum in relation to meaning generation, ranging from nominals that are more contingent on surrounding syntax for their meaning to nominals that are less contingent (we will revisit this again in Section 8.2). Therefore, when investigating nominal semantic behaviour, there will be times where we must examine our real-world knowledge of phenomena, to greater understand interactions between syntax and semantics in the nominal domain. In doing so, we will be able to better understand the limitations of an individual lexical item's meaning potential which impact its typical syntactic use.

## **8.2. The Nature and Degree of Nominality**

As was discussed above in Section 8.1.1, the finding that nominal category features display variation in influencing semantic behaviour, and that category feature interactions affect semantic behaviour, is central to our future conception of nominality. Traditional prototype models of nominal semantics view category features as stable elements, which help constitute an entity's categorical status. However, as discussed in 2.1.4, these feature lists are typically regarded as vastly 'oversimplified' (Croft and Cruse 2004, p.87). From what we have seen in study one, these category features should not be viewed as stable, but dynamic entities that actively work together in the construction of meaning.

Perhaps an appropriate description for this behaviour within the nominal domain pertains to the concept of ‘intersectionality’ (Crenshaw 1989). Although typically used to describe how different social and political identities (e.g., race; gender; class etc.) overlap and allow for the propagation of inequality and privilege (Garcia 2016, p.102), intersectionality fundamentally relates to how categories are not distinct but interrelated, and participate in the continuous mutual construction of one another (Hill Collins and Bilge 2020, p.16). This notion of co-construction has been shown in our research into semantics in the nominal domain, as the amount of influence one category feature displays over nominal semantics is evidently dependent upon the influence of another. The most notable example of this was shown in Section 5.2.2, where count/mass status shared a stronger relationship with the situation type expression of abstract nominals than concrete nominals. Consequentially, I suggest that, due to the intersectionality of nominal category features, we are unable to accurately discern the true influence of an individual category feature without considering its behaviour in relation to other category features. It is therefore almost imperative that future category models of nominal semantics take a quantitative slant and look to apply multivariate predictive modelling of some description, to accurately capture the dynamic landscape of the nominal field.

In addition to the intersective nature of nominal category features, it was noted in Section 5.2.3 that the notions of boundedness and homogeneity appear to share a different relationship in the nominal domain, when compared to the verb domain. As discussed in Section 8.1.3, this distinction was the result of how the nominal domain does not impose a direct relationship between the concepts of boundedness and homogeneity. Such a distinction is important in our evaluation of nominality as it marks a dividing line in the semantic space between nouns and verbs. It firstly supports the work of Croft (1991); Givón (2001); Sasse (2001) and Vinson and Vigliocco (2002) among others that suggests that nouns and verbs can be placed on a continuum in relation to their semantic expression. However, it also builds on this work by showing that although both nouns and verbs can express the same semantic concepts, these concepts are subject to different rules when expressed by different lexical classes.

The relationship between boundedness, homogeneity and telicity in the nominal domain therefore also suggests that the lexical class of noun may be more lenient on the behaviour of semantic concepts. Interestingly, this lenience might almost be a design feature of the lexical class, as nominals are frequently at the forefront of meaning creation, due to their ability to name new classes of entities. As Halliday (1966/2003, pp.53–54) discusses, when entities come



into existence, language must make provisions to express them, and many of these provisions are made by the lexical class noun. If the lexical class imposed stricter rules on the semantic behaviour of its members, it might hinder the natural capacity of noun to make these provisions. Accordingly, from this research, the lexical class of noun, at least semantically, appears to be a naturally more permissive lexical class than verb, which supports nominality in the expression of new meaning. This section will now lastly discuss the notion of (human) agency and how it may be an important variable in relation to our understanding of nominality.

The semantic expression of (human) agency is certainly interesting in our evaluation of nominality, for two core reasons. The first is that while research has explored how human agency is manifested through the use of verb constructions (Kenny 1963; Mourelatos 1978; Duranti 2004; Ahearn 2010; Haase 2022), very little research has devoted time to investigating human agency in the nominal domain and the influence it may hold over semantic expression. The second is that human agency represents a lexical variable. We know whether a nominal expresses (human) agency through our real-world knowledge of the phenomenon expressed by the nominal, rather than by the nominal's grammatical behaviour. Interestingly, and like discussed in Section 8.1.4, study two showed that human agency was highly influential on the extent to which four nominals were receptive to semantic coercion from their surrounding syntax. While the temporal semantics expressed by nominals such as *game* could be modified through temporal pre-modifiers in the nominal group, the temporal semantics expressed by nominals such as *fire* and *drought* could not. What this difference in semantic behaviour points to is that nominals that express (human) agency in some capacity may be more dependent on surrounding syntax for their expression of meaning than others. While previous work (Halliday and Hasan 1976; Francis 1986; Schmid 2000; Flowerdew 2003; Jiang and Hyland 2015 and many others) has shown that certain nominals, like shell-nouns, are highly dependent on their syntactic surroundings, this research has not explored more 'general' nouns and the influence lexical features may have on their syntactic dependency for meaning expression. Accordingly, future research that tries to quantify this distinction in a continuum-based model would certainly prove an insightful venture, as it would provide some clarity as to how individual nominals interact with syntax and where they draw potential event meaning from.

Moreover, (human) agency was also briefly shown in study one to be involved in the original creation of event meaning during the nominalisation of the MDA *business*. The event status of *business* was found to be highly related to notions of agency, as it expressed an event in which

something, i.e., an agent, was *busy*. It should be noted that this finding was purely founded on the semantic behaviour of one nominal. However, given the associations we have seen between temporal semantics and (human) agency throughout this thesis, (human) agency should be regarded as a serious variable of interest in the expression of temporal semantics by nominals moving forward. Throughout this thesis, we have seen that associations of agency can provide a sense of dynamism and control to a nominal's semantic expression, regardless of the nominal's derivational history (expressed by both MDA and Underived nominals). These qualities are highly important aspects in the expression of different temporal semantic structures, and, consequently, assist in the expression of temporal meaning in the nominal domain.

### 8.3. Limitations

Although this research has contributed insights into how events and states are expressed in the nominal domain, how UENs are able to express temporal semantics, and also our evaluation of nominality, the multi-perspective approach taken resulted in some restrictions for both studies. In the following paragraphs, I will detail these restrictions and the motivations for each approach.

The first study of this thesis extracted 5000 nominals from the British National Corpus (BNC), from four distinct genres. The 5000 nominals were then cut to 2700 nominal instances with repetitions, and 1662 nominal instances without repetitions, after unanalysable instances were discarded. The second study drew upon data from the Timestamped JSI Web Corpus 2014-2021 English (TJWCE) and, following Hanks' (2004, p.91) comments on adequate sample representation for lexical item behaviour, extracted enough data to attain 500 annotatable instances for the UENs *game*; *fire*; *stream*; and *drought* (n=2000). The choice to analyse this number of instances in each study was founded on finding a practical balance between having the data quantity to attain reliable statistical results whilst also providing a feasible number of instances for manual annotation. Having larger sample sizes and additional annotators would have positively influenced the reliability of this research, given that these studies tried to capture the semantic behaviour of many different types of nominals. Despite this limitation, many interesting findings were identified across the two studies that have created avenues for future research. Accordingly, it should be noted that this research does not attempt to provide a complete account for every aspect of nominal semantic behaviour. I instead encourage the

reader to view this research as an explorative base from which future research into the nature and degree of nominality can follow.

Moreover, not all semantic analyses can be entirely objective, with some instances requiring a subjective stance for their analysis (Bache 1985). In addition, the choice of research topic and statistical analyses (Braun and Clarke 2013, p.36), as well as the selection of lexical items to analyse (Baker 2006, p.12), can all be regarded as subjective practices that have influenced this research project. To circumvent some of the natural subjectivity of the semantic analysis, diagnostic syntactic tests (DSTs) were used in both studies (Dowty 1979; Barque et al. 2009; Balvet et al. 2011), where nominals would display temporal information from their capacity to function in certain syntactic constructions. These DSTs generally provided a reliable foundation on which to categorise nominals based on their situation type expression. For exceptional cases, where the nominal's functionality in DSTs was questionable, a second analyst (familiar with the field of lexical aspect) was called upon to provide a supplemental perspective on the semantic categorisation of the nominal.

Furthermore, while the use of multinomial logistic regression in the second study of this thesis provided an appropriate statistical analysis to identify the extent to which definiteness and process type influenced lexical aspect expression across the four UENs, low quantity observations (typically displayed by process types such as Behavioural, Existential and Verbal) led to an inflation in confidence intervals. As a result, the reliability of the predictions for these process types was diminished, as they displayed an increased interval range (e.g., 1%-100%). Nevertheless, as these process type categories only returned a small number of observations, they bared little influence on the central findings of the research. To pick up on an earlier limitation, future research which can make use of a larger sample size will be able to subvert this issue, as larger sample sizes would produce a greater number of observations for these process types, resulting in more accurate confidence intervals.

Lastly, the present study only explored the semantics of nominals within the English language. Therefore, the results of this study may not reflect the typical semantic behaviour of nominals within different cultures/languages, as many cultures have lexical items which boast 'culture-specific' meanings (Goddard and Wierzbicka 2014, p.2), while certain languages, such as 'Mundari' (a Munda language spoken in India), do not have dedicated lexical classes such as noun and, instead, display the semantic functions of nouns and verbs within one singular,

maximally flexible lexical class (Van Lier and Rijkhoff 2013, p.6). Accordingly, a cross-language comparison of nominal semantic behaviour between English and a separate language would prove an interesting area of future research.

## 8.4. Future research

Several paths for future research have been highlighted throughout this thesis. Most notably, these include: the use of multivariate predictive modelling in the categorisation of nominal semantics (Section 4.2); category feature importance and interaction (sections 5.1-5.3); the relationship between boundedness, homogeneity, and telicity as a semantic boundary between noun and verb (6.3); semantic flexibility expressed by UENs (Chapter 7); and the influence of (human) agency on the semantic expression of nominals (Section 5.3.1.2 and Chapter 7). In the following section, I will specifically focus the influence of (human) agency in detail.

In the discussion of the influence of MDA word formation type on nominal semantic behaviour (5.3.1.2), and more prevalently throughout Chapter 7, the notion of (human) agency appeared influential in the temporal semantics expressed by certain nominals. In study two, events that did not encode human agency in their semantic description expressed a greater resistance to semantic coercion from their surrounding syntactic context, while nominals which did encode human agency were more semantically flexible. However, this influence was not statistically examined. Furthermore, study two only considered the semantic behaviour of four UENs, therefore limiting the generalisability of these findings to nominals of differing word formation types. Accordingly, in a similar vein to the first study of this thesis, a large sample of nominals could be obtained from a corpus available in SketchEngine (Kilgarriff et al. 2004) and analysed for a series of nominal category features (e.g., abstract/concrete, count/mass etc.), with the additional analysis of (human) agency. For a more detailed insight into the influence of agency, the expression of agency could be annotated as an ordinal variable, where analysts, or perhaps even participants, would evaluate how explicit agency is expressed by each nominal on a ‘Semantic Differential Scale’ (Williams 2021, p.94). One end of the scale could signal an absence of agency, while the other end could signal a highly explicit expression of agency. By doing so, the research would be better equipped to capture nuances in agency expression (and their influence on temporal semantics) that would be largely inaccessible from a binary variable.

By applying multivariate predictive modelling in the form of an appropriate regression analysis (dependent upon which category features are chosen as predictor variables and how they are analysed), we could gain valuable statistical insights into the influence of (human) agency on nominal semantics and the behaviour of the variable in conjunction with other nominal features. As (human) agency is also expressed by verb constructions, such a study would prove an insightful endeavour in expanding our currently limited knowledge of (human) agency in the nominal domain and what temporal affordances and restrictions are provided to the notion specifically by nominality.

## References

- Aarts, B. 2004a. Conceptions of gradience in the history of linguistics. *Language Sciences* 26(4), pp. 343–389.
- Aarts, B. 2004b. Modelling linguistic gradience. *Studies in Language* 28(1), pp. 1–49.
- Aarts, B. 2006. Conceptions of categorization in the history of linguistics. *Language Sciences* 28(4), pp. 361–385.
- Aarts, B., Denison, D. and Popova, G. 2004. *Fuzzy grammar: a reader*. OUP Oxford.
- Aarts, B. and Haegeman, L. 2006. English word classes and phrases. In: McMahon, A. and Aarts, B. eds. *The handbook of English linguistics*. Malden, MA, USA: Blackwell Publishing, pp. 116–145.
- Abbott, B. 2010. *Reference*. Oxford: Oxford University Press.
- Abrard, O. and Stosic, D. 2021. *Polysémie standard et facettes de sens vues au travers du prisme de la diversité linguistique : entre la variation et la régularité*. Toulouse: Université Toulouse.
- Ahearn, L. 2010. Agency and language. In: Jaspers, J., Verschueren, J., and Östman, J.-O. eds. *Society and language use*. Amsterdam; John Benjamins Pub. Co., pp. 28–49.
- Alexiadou, A. and Grimshaw, J. 2008. Verbs, nouns, and affixation. In: Schäfer, F. ed. *SinSpeC (1): Working papers of the SFB 732*. Stuttgart: University of Stuttgart, pp. 1–16.
- Alexiadou, A., Iordăchioaia, G., Cano, M., Martin, F. and Schäfer, F. 2013. The realization of external arguments in nominalizations. *The Journal of Comparative Germanic Linguistics* 16(2–3), pp. 73–95.
- Alexiadou, A. and Schäfer, F. 2010. On the syntax of episodic vs. dispositional -er nominals. In: Alexiadou, A. and Rathert, M. eds. *Syntax of nominalizations across languages and frameworks*. Berlin: DE GRUYTER MOUTON, pp. 9–38.
- Allan, K. and Burrige, K. 1991. *Euphemism & dysphemism: language used as shield and weapon*. New York: Oxford University Press.

- Allwood, J. 2003. Meaning potentials and context: some consequences for the analysis of variation in meaning. In: Cuyckens, H., Dirven, R., and Taylor, J., R. eds. *Cognitive approaches to lexical semantics*. Berlin: De Gruyter Mouton, pp. 29–66.
- Anderson, A.J., Murphy, B. and Poesio, M. 2014. Discriminating taxonomic categories and domains in mental simulations of concepts of varying concreteness. *Journal of Cognitive Neuroscience* 26(3), pp. 658–681.
- Anderson, C. and Rutkowski, L. 2008. Multinomial logistic regression. In: Osborne, J. ed. *Best practices in quantitative methods*. Thousand Oaks, California, United States of America: SAGE Publications, pp. 390–409.
- Andreou, M. and Lieber, R. 2020. Aspectual and quantificational properties of deverbal conversion and -ing nominalizations: the power of context. *English Language & Linguistics* 24(2), pp. 333–363.
- Andreou, M.A. 2021. Personal/participant/inhabitant in morphology. In: Lieber, R. ed. *The Oxford encyclopedia of morphology*. Oxford: Oxford University Press
- Arche, M.J. 2006. *Individuals in time tense, aspect and the individual/stage distinction*. Amsterdam: John Benjamins.
- Arche, M.J., Fábregas, A. and Marín, R. 2021. On event-denoting deadjectival nominalizations. *The Linguistic Review* 38(2), pp. 191–231.
- Arche, M.J. and Marín, R. 2014. Edgy Nominalizations. *Proceedings of the 31st West Coast Conference on Formal Linguistics. Cascadilla Proceedings Project*, pp. 21–30.
- Arche, M.J. and Stowell, T. 2019. Dispositional adjectives: characterizing and episodic predication. In: *Paper presented at the workshop on the individual/stage-level contrast, discussion series on persistent issues in language analysis*. Greenwich: University of Greenwich
- Aronoff, M. and Fudeman, K.A. 2011. *What is morphology?* 2nd ed. Malden MA: Blackwell Pub.
- Bach, E. 1986. The algebra of events. *Linguistics and Philosophy* 9(1), pp. 5–16.

- Bache, C. 1985. The semantics of grammatical categories: A dialectical approach. *Journal of Linguistics* 21(1), pp. 51–77.
- Baker, M.C. 2003. *Lexical categories: verbs, nouns, and adjectives*. Cambridge; New York: Cambridge University Press.
- Baker, P. 2006. *Using corpora in discourse analysis*. London; Continuum.
- Balvet, A., Barque, L., Condette, M.-H., Haas, P., Huyghe, R., Marín, R. and Merlo, A. 2011. Nomage: an electronic lexicon of French deverbal nouns based on a semantically annotated corpus. In *WoLeR 2011 at ESSLLI, International Workshop on Lexical Resources*, pp. 8–15.
- Barque, L., Fabregas, A. and Marín, R. 2011. On the (un)countability of stative deverbal nouns. Aston University, Birmingham, pp. 1–10.
- Barque, L., Marín, R., Jugnet, A. and Huyghe, R. 2009. Two types of deverbal activity nouns in French. Pisa, Italy, pp. 169–175.
- Bartlett, T. 2014. *Analysing power in language: a practical guide*. Milton Park, Abingdon, Oxon: Routledge.
- Bateman, J., A. 2017. The place of systemic functional linguistics as a linguistic theory in the twenty-first century. In: Bartlett, T. and O’Grady, G. eds. *The Routledge handbook of Systemic Functional Linguistics*. London: Routledge, pp. 11–26.
- Bauer, L., Lieber, R. and Plag, I. 2013. *The Oxford reference guide to English morphology*. 1st ed. Oxford: Oxford University Press.
- Bayaga, A. 2010. Multinomial logistic regression: usage and application in risk analysis. *Journal of Applied Quantitative Methods* 5(2), pp. 288–297.
- Bennett, P. 2014. Langacker’s Cognitive Grammar. In: Taylor, J. R. and Littlemore, J. eds. *The Bloomsbury companion to Cognitive Linguistics*. London: Bloomsbury, pp. 29–48.
- Berlin, B. and Kay, P. 1969. *Basic color terms: their universality and evolution*. London: University of California Press.



Bielak, J. and Pawlak, M. 2013. Introduction to Cognitive Grammar. In: Bielak, J. and Pawlak, M. eds. *Applying Cognitive Grammar in the foreign language classroom: teaching English tense and aspect*. Berlin, Heidelberg: Springer, pp. 7–56.

Biese, Y.M. 1941. *Origin and development of conversion in English*. Helsinki: Annales Academiae Scientiarum Fennicae B XLV.

Binnick, R.I. 2006. Aspect and aspectuality. In: McMahon, A. and Aarts, B. eds. *The handbook of English linguistics*. Malden, MA, USA: Blackwell Publishing, pp. 244–268.

Bloch-Trojnar, M. 2013. Deverbal nominalizations in English: an LMBM approach. In: Hacken, P. ten and Thomas, C. eds. *The semantics of word formation and lexicalization*. Edinburgh: Edinburgh University Press, pp. 180–203.

Bloch-Trojnar, M. and Malicka-Kleparska, A. 2017. *Aspect and valency in nominals*. Boston: De Gruyter, Inc.

Bloor, M. and Bloor, T. 2007. *The practice of critical discourse analysis: an introduction*. London: Hodder Arnold.

Bloor, M. and Bloor, T. 2018. Systemic Functional Linguistics. In: Flowerdew, J. and Richardson, J. E. eds. *The Routledge handbook of critical discourse studies*. London: Routledge

Bloor, T. and Bloor, M. 2013. *The functional analysis of English: a Hallidayan approach*. Third Edition. London; New York: Routledge/Taylor & Francis Group.

BNC 2007. The British National Corpus, version 3 (BNC XML Edition). Available at: <http://www.natcorp.ox.ac.uk/>.

Bolognesi, M., Burgers, C. and Caselli, T. 2020. On abstraction: decoupling conceptual concreteness and categorical specificity. *Cognitive Processing* 21(3), pp. 365–381.

Booij, G. 2006. Inflection and derivation. In: Brown, K. ed. *Encyclopedia of Language & Linguistics*. Oxford: Elsevier, pp. 654–661.

Booij, G.E. 2005. *The grammar of words: an introduction to linguistic morphology*. Oxford: Oxford University Press.

- Braun, V. and Clarke, V. 2013. *Successful qualitative research: a practical guide for beginners*. Los Angeles; SAGE.
- Bresnan, J., Cueni, A., Nikitina, T. and Baayen, R.H. 2007. Predicting the Dative Alternation. In: Bouma, G., Krämer, I., and Zwarts, J. eds. *Cognitive foundations of interpretation*. Amsterdam: Royal Netherlands Academy of Science, pp. 69–94.
- Brinton, L. 1995. The aktionsart of deverbal nouns in English. In: Bertinetto, P. M. ed. *Temporal reference, aspect and actionality: semantic and syntactic perspectives*. Rosenberg & Sellier, pp. 27–42.
- Brinton, L., J. 1991. The mass/count distinction and aktionsart: the grammar of iteratives and habituais. *Belgian Journal of Linguistics* 6(1), pp. 47–69.
- Brinton, L., J. 1998. Aspectuality and countability: a cross-categorial analogy. *English Language and Linguistics* 2(1), pp. 37–63.
- Broccias, C. 2013. Cognitive Grammar. In: Hoffmann, T. and Trousdale, G. eds. *The Oxford handbook of Construction Grammar*. Oxford: Oxford University Press, pp. 191–210.
- Broccias, C. and Hollmann, W.B. 2007. Do we need summary and sequential scanning in (Cognitive) grammar? *Cognitive Linguistics* 18(4), pp. 487–522.
- Bromhead, H. 2006. The semantics of bushfire in Australian English. In: Ye, Z. and Bromhead, H. eds. *Meaning, Life and Culture: In conversation with Anna Wierzbicka*. 1st ed. Australian National University: ANU Press, pp. 115–134.
- Brown, R.W. 1957. Linguistic determinism and the part of speech. *The Journal of Abnormal and Social Psychology* 55(1), pp. 1–5.
- Brysbaert, M., Warriner, A.B. and Kuperman, V. 2014. Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods* 46(3), pp. 904–911.
- Butler, C.S. 2013. Systemic Functional Linguistics, Cognitive Linguistics and Psycholinguistics: Opportunities for dialogue. *Functions of language* 20(2), pp. 185–218.
- Carstairs-McCarthy, A. 2002. *An introduction to English morphology: words and their structure*. Edinburgh: Edinburgh University Press.

Cohen, A. 2016. A semantic explanation for the external argument generalization. *Morphology* 26(1), pp. 91–103.

Collier, E. 2021. *How is manner encoded in motion constructions? A systemic functional linguistics exploration into expressions of manner. MA Thesis.* Cardiff University: Unpublished.

Comrie, B. 1976. *Aspect: an introduction to the study of verbal aspect and related problems.* Cambridge: Cambridge University Press.

Crenshaw, K. 1989. Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine. *University of Chicago Legal Forum*, pp. 139–168.

Croft, W. 1991. *Syntactic categories and grammatical relations: the cognitive organization of information.* Chicago: University of Chicago Press.

Croft, W. 2001. *Radical construction grammar: syntactic theory in typological perspective.* Oxford: Oxford University Press.

Croft, W. 2009. Toward a social cognitive linguistics. In: Evans, V. and Pourcel, S. eds. *New directions in Cognitive Linguistics.* Amsterdam; Philadelphia: John Benjamins, pp. 395–420.

Croft, W. 2012. *Verbs: aspect and causal structure.* Oxford: Oxford University Press.

Croft, W. and Cruse, D., A, 2004. *Cognitive Linguistics.* Cambridge: Cambridge University Press.

Cruse, D., Alan 2000. Aspects of the micro-structure of word meanings. In: Ravin, Y. and Leacock, C. eds. *Polysemy: theoretical and computational approaches.* Oxford: Oxford University Press, pp. 30–51.

Cruse, D.A. 2001. Microsenses, default specificity and the semantics-pragmatics boundary. *Axiomathes* 12(1), pp. 35–54.

Cruse, D.A. 2011. *Meaning in language: an introduction to semantics and pragmatics.* 3rd ed. Oxford: Oxford University Press.

Crystal, D. 1967. English. *Lingua* 17, pp. 24–56.

Dąbrowska, E. and Divjak, D. 2015. *Handbook of Cognitive Linguistics*. Walter de Gruyter GmbH & Co KG.

Damasio, A., R. and Tranel, D. 1993. Nouns and verbs are retrieved with differently distributed neural systems. *Proceedings of the National Academy of Sciences of the United States of America* 90(11), pp. 4957–4960.

Danesi, M. 2015. Advertising Discourse. In: Tracy, K., Sandel, T., and Ilie, C. eds. *The international encyclopedia of language and social interaction*. Chichester, Malden, MA: Wiley-Blackwell, pp. 35–44.

Davies, C. 2004. *A corpus-based investigation of noun to verb conversion in English*. Liverpool: University of Liverpool.

Davies, M. 2008. The Corpus of Contemporary American English (COCA): 520 million words, 1990-present. Available at: <http://corpus.byu.edu/coca/>.

Declerck, R. 1989. Boundedness and the Structure of Situations. *Leuvense Bijdragen* 78, pp. 275–308.

Declerck, R. 2006. *The grammar of the English verb phrase*. Berlin: Mouton de Gruyter.

Declerck, R. 2007. Distinguishing between the aspectual categories ‘(a)telic’, ‘(im)perfective’ and ‘(non)bounded’. *Kansas Working Papers in Linguistics* 29, pp. 48–64.

Depraetere, I. 1995. On the necessity of distinguishing between (un)boundedness and (a)telicity. *Linguistics and Philosophy* 18(1), pp. 1–19.

van Der Auwera, J. and Gast, V. 2011. Categories and prototypes. In: *The Oxford handbook of linguistic typology*. Oxford: Oxford University Press, pp. 166–189.

Derewianka, B. 2020. Changing approaches to the conceptualization and teaching of grammar. In: Gao, X. ed. *Second handbook of English language teaching*. Cham: Springer International Publishing, pp. 821–845.

Deshors, S.C. and Götz, S. 2020. Common ground across globalized English varieties: a multivariate exploration of mental predicates in World Englishes. *Corpus Linguistics and Linguistic Theory* 16(1), pp. 1–28.

- Dickey, S. 2016. Lexical and grammatical aspect. In: Riemer, N. ed. *The Routledge handbook of semantics*. London; New York, NY: Routledge, pp. 338–353.
- Dixon, R.M.W. 1982. Where have all the adjectives gone? In: *Where have all the adjectives gone? and other essays in semantics and syntax*. Berlin, New York: DE GRUYTER MOUTON, pp. 1–62.
- Dong, S., Huang, C.-R. and Ren, H. 2020. Towards a new typology of meteorological events: A study based on synchronic and diachronic data. *Lingua* 247, pp. 1–14.
- Dowty, D.R. 1979. *Word meaning and Montague Grammar: the semantics of verbs and times in Generative semantics and in Montague's ptq*. Dordrecht, Holland; London: D. Reidel.
- Drożdż, G. 2017. *The puzzle of (un)countability in English. A study in Cognitive Grammar*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- Du Bois, J. 1980. Beyond definiteness: the trace of identity in discourse. In: Chafe, W. L. ed. *The pear stories: cognitive, cultural, and linguistic aspects of narrative production*. Advances in discourse processes. Norwood, N.J: Ablex PubCorp, pp. 203–274.
- Duranti, A. 2004. Agency in language. In: Duranti, A. ed. *A companion to linguistic anthropology*. 1st ed. Malden, Mass.; Blackwell, pp. 451–473.
- Evans, V. 2010. On the nature of lexical concepts. *Belgrade English Language and Literature Studies* 2(1), pp. 11–45.
- Evans, V. and Green, M. 2006. *Cognitive Linguistics: an introduction*. Edinburgh: Edinburgh University Press.
- Fábregas, A., Leferman, B. and Marín, R. 2013. Evaluative adjectives are Davidsonian states. *Proceedings of Sinn und Bedeutung* 17, pp. 237–253.
- Fábregas, A. and Marín, R. 2012. The role of aktionsart in deverbal nouns: state nominalizations across languages. *Journal of Linguistics* 48(1), pp. 35–70.
- Ferrero, L. 2022. Introduction to part 7. In: Ferrero, L. ed. *The Routledge handbook of philosophy of agency*. Abingdon, Oxon; Routledge, pp. 329–335.
- Filip, H. 2012. *Lexical aspect*. Oxford: Oxford University Press.

Flowerdew, J. 2003. Signalling nouns in discourse. *English for Specific Purposes* 22(4), pp. 329–346.

Fontaine, L. 2013. *Analysing English grammar: a systemic functional introduction*. Cambridge: Cambridge University Press.

Fontaine, L. 2017a. Fluid boundaries and the categorization of nominal expressions. In: Neumann, S., Wegener, R., Fest, J., Niemietz, P., Hatzen, N., and International Systemic Functional Congress eds. *Challenging boundaries in linguistics: systemic functional perspectives*. Frankfurt am Main; Peter Lang, pp. 19–38.

Fontaine, L. 2017b. The early semantics of the neologism BREXIT: a lexicogrammatical approach. *Functional Linguistics* 4(1), pp. 1–15.

Fontaine, L. 2017c. The noun, grammar and context. *Linguistics and the Human Sciences* 11(2–3), pp. 178–202.

Fontaine, L. 2019. When leaving is a thing: construing Britain’s exit from the European Union. In: Carter-Thomas, S. and Hamilton, C. E. eds. *Science, Systemic Functional Linguistics and Language Change*. Cambridge: Cambridge Scholars Publishing

Fontaine, L., Jones and Schönthal, D. 2023. *Referring in Language: an integrated approach*. Cambridge: Cambridge University Press.

Fonteyn, L. 2016. From nominal to verbal gerunds: a referential typology. *Functions of Language* 23(1), pp. 60–83.

Fonteyn, L. 2019. A corpus-based view on the (aspectual-)semantics of modern English nominalizations. *Language Sciences* 73, pp. 77–90.

Formanowicz, M., Pietraszkiewicz, A., Roessel, J., Suitner, C., Witkowska, M. and Maass, A. 2021. “Make it happen!”: Verbs as markers of agency increase message effectiveness. *Social Psychology* 52(2), pp. 75–89.

Fowler, R. 1991. *Language in the news: discourse and ideology in the British press*. London: Routledge.

Francis, G. 1986. *Anaphoric nouns*. Birmingham: English Language Research, University of Birmingham.

Garcia, L. 2016. Intersectionality. *Kalfou (Santa Barbara, Calif.)* 3(1), pp. 102–106.

Geeraerts, D. 1985. *Paradigm and paradox: explorations into a paradigmatic theory of meaning and its epistemological background*. Leuven: Leuven University Press.

Geeraerts, D. 2000. Saliency phenomena in the lexicon: a typology. In: Albertazzi, L. ed. *Meaning and cognition: a multidisciplinary approach*. Amsterdam: John Benjamins Publishing Company, pp. 79–101.

Geeraerts, D. 2006. Prototype theory: prospects and problems of prototype theory. In: Geeraerts, D. ed. *Cognitive Linguistics: basic readings*. Berlin: Walter de Gruyter, pp. 141–168.

Geeraerts, D. 2010. *Theories of lexical semantics*. Oxford: Oxford University Press.

Givón, T. 1981. On the development of the numeral ‘one’ as an indefinite marker. *Folia Linguistica Historica* 15, pp. 35–54.

Givón, T. 2001. *Syntax: an introduction*. Amsterdam; Philadelphia: John Benjamins.

Gleason, H.A. 1965. *Linguistics and English grammar*. Holt, Rinehart & Winston.

Glynn, D. 2014. Techniques and tools: corpus methods and statistics for semantics. In: Glynn, D. and Robinson, J. A. eds. *Corpus methods for semantics: quantitative studies in polysemy and synonymy*. Amsterdam: John Benjamins Publishing Company, pp. 307–342.

Goddard, C. and Wierzbicka, A. 2014. *Words and meanings: lexical semantics across domains, languages, and cultures*. Oxford: Oxford University Press.

Grimm, S. 2014. Individuating the abstract. *Proceedings of Sinn und Bedeutung* 18, pp. 182–200.

Grimm, S. and McNally, L. 2013. No ordered arguments needed for nouns. *Proceedings of the 19th Amsterdam colloquium*, pp. 123–130.

Grimshaw, J. 1990. *Argument Structure*. Cambridge, Mass: MIT Press.

- Gross, G. and Kiefer, F. 1995. La structure événementielle des substantifs. *Folia Linguistica* 29(1–2).
- Haas, P. and Huyghe, R. 2010. Les propriétés aspectuelles des noms d'activités. *Cahiers Chronos* 21, pp. 103–118.
- Haase, M. 2022. Agency, events and processes. In: Ferrero, L. ed. *The Routledge handbook of philosophy of agency*. Abingdon, Oxon; Routledge, pp. 47–58.
- Halliday, M.A.K. 1961. Categories of the theory of grammar. *WORD* 17(2), pp. 241–292.
- Halliday, M.A.K. 1966. Grammar, society, and the noun. In: Halliday, M. A. K. and Webster, J. eds. *On Language and Linguistics. Collected Works of MAK Halliday 3*. London: Continuum, pp. 50–73.
- Halliday, M.A.K. 1985. *An introduction to functional grammar*. London: Edward Arnold.
- Halliday, M.A.K. 1994. *An introduction to functional grammar*. 2nd ed. London: Edward Arnold.
- Halliday, M.A.K. 2009. *The essential Halliday*. London; New York: Continuum.
- Halliday, M.A.K. and Hasan, R. 1976. *Cohesion in English*. London: Longman.
- Halliday, M.A.K. and Martin, J.R. 1993. *Writing science: literacy and discursive power*. London: Routledge, Taylor & Francis Group.
- Halliday, M.A.K. and Matthiessen, C.M.I.M. 1999. *Construing experience through meaning: a language-based approach to cognition*. London; New York: Continuum.
- Halliday, M.A.K. and Matthiessen, C.M.I.M. 2004. *An introduction to functional grammar*. 3rd ed. London: Arnold.
- Halliday, M.A.K. and Matthiessen, C.M.I.M. 2014. *Halliday's introduction to functional grammar*. Fourth edition. London: Routledge.
- Hanks, P. 2000. Do word meanings exist? *Computers and the Humanities* 34(1/2), pp. 205–215.



- Hanks, P. 2004. Corpus pattern analysis. In: Williams, G. and Vessier, S. eds. *Proceedings of the 11th Euralex International Congress*. Lorient, France: UBS, pp. 87–98.
- Hanks, P. 2007. Preference syntagmatics. In: Khurshid, A., Brewster, C., and Stevenson, M. eds. *Words and Intelligence II: Essays in Honor of Yorick Wilks*. Dordrecht, The Netherlands: Springer, pp. 119–136. doi: 10.1007/1-4020-5833-0\_13.
- Hanks, P. 2012. How people use words to make meanings: Semantic types meet valencies. In: Boulton, A. and Thomas, J. eds. *Input, process and product: developments in teaching and language corpora*. Masarykova Univerzita: Masaryk University Press, pp. 54–69.
- Hanks, P. 2013. *Lexical analysis norms and exploitations*. Cambridge, Mass.: MIT Press.
- Hanks, P. 2020. How context determines meaning. In: Corpas Pastor, G. and Colson, J.-P. eds. *Computational phraseology*. Amsterdam; John Benjamins Publishing Company, pp. 297–310.
- Hao, J. 2018. Reconsidering “cause inside the clause” in scientific discourse – from a discourse semantic perspective in systemic functional linguistics. *Text & Talk* 38(5), pp. 525–550.
- Harnad, S. 2017. To cognize is to categorize. In: Cohen, H. and Lefebvre, C. eds. *Handbook of categorization in cognitive science*. Amsterdam: Elsevier, pp. 21–54.
- Hasan, R. 2009. Wanted: a theory for integrated sociolinguistics. In: Webster, J. ed. *The collected works of Ruqaiya Hasan, vol 2: semantic variation – meaning in society and in sociolinguistics*. London; Oakville, CT: Equinox, pp. 5–40.
- Haspelmath, M. 1996. Word-class-changing inflection and morphological theory. In: Booij, G. and van Marle, J. eds. *Yearbook of morphology 1995*. Dordrecht: Springer Netherlands, pp. 43–66.
- Hayriyan, A. 2020. Deverbal nominals possess no internal VP structure. *Lingua* 248, pp. 1–23.
- Hess, A.S. and Hess, J.R. 2017. Understanding tests of the association of categorical variables: the Pearson Chi-Square test and Fisher’s exact test. *Transfusion* 57(4), pp. 877–879.
- Heyd, S. and Knittel, M.-L. 2009. Les noms d’activité parmi les noms abstraits: propriétés aspectuelles, distributionnelles et interprétatives. *Lingvisticae investigationes* 32(1), pp. 124–148.

Heyvaert, L. 2003. *A cognitive-functional approach to nominalization in English*. Berlin: Mouton de Gruyter.

Heyvaert, L. 2010. A cognitive-functional perspective on deverbal nominalization in English. Descriptive findings and theoretical ramifications. In: Rathert, M. and Alexiadou, A. eds. *The semantics of nominalizations across languages and frameworks*. Berlin ; New York: DE GRUYTER MOUTON, pp. 51–81.

Heyvaert, L. 2015. *A cognitive-functional approach to nominalization in English*. Reprint 2015. Berlin; Boston: De Gruyter Mouton.

Heyvaert, L., Maekelberghe, C. and Buyle, A. 2019. Nominal and verbal gerunds in present-day English: aspectual features and nominal status. *Language Sciences* 73, pp. 32–49.

Hill Collins, P. and Bilge, S. 2020. *Intersectionality*. 2nd edition. Cambridge, UK; Polity Press.

Hinton, P.R. 1995. *Statistics explained: a guide for social science students*. London; New York: Routledge.

Hopper, P.J. 1991. On some principles of grammaticalization. In: Traugott, E. C. and Heine, B. eds. *Approaches to grammaticalization*. Amsterdam: Benjamins, pp. 17–16.

Hopper, P.J. and Thompson, S.A. 1984a. The discourse basis for lexical categories in universal grammar. *Language* 60(4), pp. 703–752.

Hopper, P.J. and Thompson, S.A. 1984b. The discourse basis for lexical categories in universal grammar. *Language* 60(4), pp. 703–752.

Hosmer, D.W. and Lemeshow, S. 2000. *Applied logistic regression*. 2nd ed. New York: Wiley.

Hosmer, D.W., Lemeshow, S. and Sturdivant, R.X. 2013. *Applied logistic regression*. 3rd ed. Hoboken, N.J: Wiley.

Hothorn, T., Hornik, K. and Zeileis, A. 2006. Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics* 15(3), pp. 651–674.

Hudson, R.A. 2010. *An introduction to word grammar*. Cambridge; New York: Cambridge University Press.

- Hunston, S. and Sinclair, J. 1999. A local grammar of evaluation. In: Hunston, S. and Thompson, G. eds. *Evaluation in text: authorial stance and the construction of discourse*. Oxford: Oxford University Press, pp. 74–101.
- Huyghe, R. 2011. (A)telicity and the mass-count distinction: The case of French activity nominalizations. *Recherches linguistiques de Vincennes* 40, pp. 101–126.
- Huyghe, R., Barque, L., Haas, P. and Tribout, D. 2017. The semantics of underived event nouns in French. *Italian Journal of Linguistics* (1), pp. 117–142.
- Jackendoff, R. 1983. *Semantics and cognition*. Cambridge: MIT Press.
- Jackendoff, R. 1990. *Semantic structures*. Cambridge, Mass.: The MIT Press.
- Jackendoff, R. 1991. Parts and boundaries. *Cognition* 41(1–3), pp. 9–45.
- Jackson, H. and Zé Amvela, E. 2007. *Words, meaning and vocabulary: an introduction to modern English lexicology*. 2nd ed. London; Continuum.
- Jenset, G.B. 2008. Basic statistics for corpus linguistics. In: *Handout for methods seminar in English linguistics*. University of Bergen: Unpublished, pp. 1–20. Available at: <http://www.uib.no/People/gje037/statTutorialR.pdf>.
- Ježek, E. and Ramat, P. 2009. On parts-of-speech transcategorization. *Folia Linguistica* 43(2), pp. 391–416.
- Jiang, F. (Kevin) and Hyland, K. 2015. ‘The fact that’: stance nouns in disciplinary writing. *Discourse Studies* 17(5), pp. 529–550.
- Jones, K. 2014. *Towards an understanding of the use of indefinite expressions for definite reference in English discourse*. Cardiff: Cardiff University.
- Jones, K. 2018. ‘A man who revels in his own ignorance, racism and misogyny’: Identifiable referents trump indefinite grammar. *Functional Linguistics* 5(1), pp. 1–22.
- Kearns, K. 2000. *Semantics*. Houndmills, Basingstoke, Hampshire; London: Macmillan Press.
- Kearns, K. 2011. *Semantics*. 2nd ed. New York: Palgrave Macmillan.

- Kellermann, K. and Cole, T. 1994. Classifying compliance gaining messages: Taxonomic disorder and strategic confusion. *Communication Theory* 4(1), pp. 3–60.
- Kenny, A. 1963. *Action, emotion and will*. London: Routledge & Keegan Paul.
- Kilgarriff, A., Rychly, P., Smrz, P. and Tugwell, D. 2004. The Sketch Engine. *Information Technology*, pp. 105–116.
- Kiss, T., Pelletier, F.J., Husic, H. and Poppek, J. 2017. Issues of mass and count: dealing with ‘dual-life’ nouns. In: *Proceedings of the 6th Joint Conference on Lexical and Computational Semantics (\*SEM 2017)*. Vancouver, Canada: Association for Computational Linguistics, pp. 189–198.
- Kittilä, S., Eriksen, P.K. and Kolehmainen, L. 2010. The linguistics of weather: Cross-linguistic patterns of meteorological expressions. *Studies in Language* 34(3), pp. 565–601.
- Labov, W. 1973. The boundaries of words and their meanings. In: Bailey, C. J. N. and Shuy, R. W. eds. *New ways of analyzing variation in English*. Washington D.C: Georgetown University Press.
- Lakoff, G. 1970. *Irregularity in syntax*. New York: Holt, Rinehart and Winston.
- Lakoff, G. 1987. *Women, fire, and dangerous things: what categories reveal about the mind*. Chicago; London: University of Chicago Press.
- Langacker, R. 2008a. *Cognitive Grammar: a basic introduction*. New York: Oxford University Press.
- Langacker, R.W. 1987a. *Foundations of cognitive grammar. Vol.I, Theoretical prerequisites*. Stanford: Stanford University Press.
- Langacker, R.W. 1987b. Nouns and Verbs. *Language* 63(1), pp. 53–94.
- Langacker, R.W. 1990. *Concept, image, and symbol: the cognitive basis of grammar*. Berlin: Mouton de Gruyter.
- Langacker, R.W. 1991. *Foundations of cognitive grammar. Vol.II, descriptive application*. Stanford: Stanford University Press.

Langacker, R.W. 2000. *Grammar and conceptualization*. Berlin; New York: Mouton de Gruyter.

Langacker, R.W. 2004. Remarks on nominal grounding. *Functions of Language* 11(1), pp. 77–113.

Langacker, R.W. 2008b. Sequential and summary scanning: a reply. *Cognitive Linguistics* 19(4), pp. 571–584.

Langacker, R.W. 2009. *Investigations in Cognitive Grammar*. Berlin; New York: Mouton de Gruyter.

Langacker, R.W. 2013. *Essentials of Cognitive Grammar*. New York: Oxford university press.

Langacker, R.W. 2017a. Cognitive Grammar. In: Dancygier, B. ed. *The Cambridge handbook of Cognitive Linguistics*. Cambridge: Cambridge University Press, pp. 262–283.

Langacker, R.W. 2017b. Conceptualization, symbolization and grammar. *Journal of Cognitive Linguistics* 1(1), pp. 31–64.

Langacker, R.W. 2019. Construal. In: Dąbrowska, E. and Divjak, D. eds. *Cognitive Linguistics - foundations of language*. Berlin; Boston: Walter de Gruyter, pp. 140–166.

Levin, B. and Rappaport Hovav, M. 1988. Nonevent -er nominals: a probe into argument structure. *Linguistics* 26(6), pp. 1067–1084.

Levshina, N. 2015. *How to do linguistics with R: data exploration and statistical analysis*. Amsterdam, Netherlands; Philadelphia, Pennsylvania: John Benjamins Publishing Company.

Li, P. and Shirai, Y. 2000. *The acquisition of lexical and grammatical aspect*. Berlin; New York: Mouton de Gruyter.

Lieber, R. 2004. *Morphology and lexical semantics*. Cambridge: Cambridge University Press.

Lieber, R. 2010. *Introducing morphology*. Cambridge; New York: Cambridge University Press.

Lieber, R. 2016. *English nouns: the ecology of nominalization*. Cambridge: Cambridge University Press.

Lieber, R. and Andreou, M. 2018. Aspect and modality in the interpretation of deverbal -er nominals in English. *Morphology* 28(2), pp. 187–217.

Lieber, R. and Plag, I. 2022. The semantics of conversion nouns and -ing nominalizations: A quantitative and theoretical perspective. *Journal of linguistics* 58(2), pp. 307–343.

Lüdecke, D. 2021. *\_sjPlot: Data visualization for statistics in social science\_*. Available at: <https://CRAN.R-project.org/package=sjPlot>.

Lyons, J. 1977. *Semantics / Vol.2*. Cambridge: Cambridge University Press.

Maekelberghe, C. and Heyvaert, L. 2016. Indefinite nominal gerunds, or the particularization of a reified event. *English Studies* 97(3), pp. 317–340.

Malinowski, B. 1923. *The problem of meaning in primitive languages*. New York: Harcourt, Brace, and Co.

Martin, A., Haxby, J.V., Lalonde, F.M., Wiggs, C.L. and Ungerleider, L.G. 1995. Discrete cortical regions associated with knowledge of color and knowledge of action. *Science* 270(5233), pp. 102–105.

Martin, J.R. 2022. Systemic Functional Linguistics (SFL): A social theory of language. In: Ding, A. and Evans, M. eds. *Social Theory for English for Academic Purposes: Foundations and Perspectives*. London: Bloomsbury Publishing, pp. 59–84.

Martsa, S. 2013. *Conversion in English: a cognitive semantic approach*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Matthiessen, C.M.I.M., Wang, B., Ma, Y. and Mwinlaaru, I.N. 2022. *Systemic functional insights on language and linguistics*. Singapore: Springer Singapore. doi: 10.1007/978-981-16-8713-6.

Medin, D.L. and Shoben, E.J. 1988. Context and structure in conceptual combination. *Cognitive Psychology* 20(2), pp. 158–190.

Menard, S.W. 2010. *Logistic regression from introductory to advanced concepts and applications*. Thousand Oaks, Calif.: SAGE.

Michaelis, L., A. 2006. Cognitive Linguistics. *Language* 82(4), pp. 898–902.

Moltman, F. 2019. Nominals and event structure. In: Truswell, R. ed. *Oxford handbook on event structure*. Oxford: Oxford University Press, pp. 342–367.

Moon, R. 1997. Vocabulary connections: multi-word items in English. In: Schmitt, N. and McCarthy, M. eds. *Vocabulary: description, acquisition and pedagogy*. Cambridge: Cambridge University Press, pp. 40–63.

Mourelatos, A. 1978. Events, processes, and states. *Linguistics and Philosophy* 2(3), pp. 415–434.

Murphy, M.L. 2010. *Lexical meaning*. New York: Cambridge University Press.

Niemeier, S. and Reif, M. 2008. Making progress simpler? Applying cognitive grammar to tense-aspect teaching in the German EFL classroom. In: Knop, S. D. and Rycker, T. D. eds. *Cognitive Approaches to Pedagogical Grammar: A Volume in Honour of René Dirven*. Berlin; New York: Walter de Gruyter, pp. 325–356.

Nowak, E. 2021. No context, no content, no problem. *Mind & language* 36(2), pp. 189–220.

OED 2020a. advance, n. and adj. *OED Online*. Available at: <http://www.oed.com/view/Entry/2883> [Accessed: 20 November 2020].

OED 2020b. aircraft, n. *OED Online*. Available at: <http://www.oed.com/view/Entry/4380> [Accessed: 20 November 2020].

OED 2020c. building, n. *OED Online*. Available at: <http://www.oed.com/view/Entry/24409> [Accessed: 20 November 2020].

OED 2020d. endeavour, n. *OED Online*. Available at: <http://www.oed.com/view/Entry/61901> [Accessed: 9 August 2021].

OED 2020e. happiness, n. *OED Online*. Available at: <http://www.oed.com/view/Entry/84070> [Accessed: 20 November 2020].

OED 2020f. house, n.1 and int. *OED Online*. Available at: <http://www.oed.com/view/Entry/88886> [Accessed: 20 November 2020].

OED 2020g. journalist, n. *OED Online*. Available at: <http://www.oed.com/view/Entry/101740> [Accessed: 20 November 2020].

OED 2020h. pocket, n. and adj. *OED Online*. Available at: <http://www.oed.com/view/Entry/146402> [Accessed: 20 November 2020].

OED 2020i. producer, n. *OED Online*. Available at: <http://www.oed.com/view/Entry/151981> [Accessed: 20 November 2020].

OED 2020j. result, v. *OED Online*. Available at: <http://www.oed.com/view/Entry/164062#eid25723917> [Accessed: 9 August 2021].

OED 2020k. rush, n.2, adv., and adj. *OED Online*. Available at: <http://www.oed.com/view/Entry/169027> [Accessed: 20 November 2020].

Olsen, M.B. 1994. The semantics and pragmatics of lexical aspect features. *Studies in the Linguistic Sciences*. 24, pp. 361–375.

Olsen, M.B. 1997. *A semantic and pragmatic model of lexical and grammatical aspect*. New York; London: Routledge.

Osborne, J.W. 2015. *Best practices in logistic regression*. Los Angeles: SAGE.

Oxford English Dictionary 2022. Oxford English dictionary (Online). Available at: <http://www.oed.com.abc.cardiff.ac.uk/> [Accessed: 20 November 2020].

Paradis, C. 2004. Where does metonymy stop? Senses, facets, and active zones. *Metaphor and Symbol* 19(4), pp. 245–264.

Plag, I. 1999. *Morphological productivity: structural constraints in English derivation*. De Gruyter Mouton.

Polguère, A. 2015. Lexical contextualism: the Abélard syndrome. In: Gala, N., Rapp, R., and Bel-Enguix, G. eds. *Language production, cognition, and the lexicon*. Cham: Springer International Publishing, pp. 53–73.

Prince, E. 1992. Subjects, definiteness and information status. In: Mann, W. C. and Thompson, S. A. eds. *Discourse description diverse linguistic analyses of a fund-raising text*. Amsterdam: John Benjamins Pub, pp. 295–325.

Quirk, R. 1965. Descriptive statement and serial relationship. *Language* 41(2), pp. 205–217.



Quirk, R., Greenbaum, S., Leech, G. and Svartvik, J. 1985. *A comprehensive grammar of the English language*. Harlow: Longman.

R Core Team 2018. R: A language and environment for statistical computing. R Foundation for statistical computing. Available at: <https://www.R-project.org/>.

Radden, G. and Dirven, R. 2007. *Cognitive English grammar*. Amsterdam; Philadelphia: John Benjamins Pub.

Rappaport Hovav, M. and Levin, B. 1992. -er nominals: implications for the theory of argument structure. In: Stowell, T. and Wehrli, E. eds. *Syntax and semantics, vol. XXVI: syntax and the lexicon*. New York: Academic Press, pp. 127–153.

Rasinger, S., M. 2013. *Quantitative research in linguistics: an introduction*. 2nd edition. London: Bloomsbury Academic.

Reichl, K. 1982. *Categorial Grammar and word-formation: the de-adjectival abstract noun in English*. Berlin; Boston: De Gruyter.

Richardson, J.E. 2007. *Analysing newspapers: an approach from critical discourse analysis*. Basingstoke; New York: Palgrave Macmillan.

Roberts, S.G., Torreira, F. and Levinson, S.C. 2015. The effects of processing and sequence organization on the timing of turn taking: a corpus study. *Frontiers in Psychology* 6, p. 509.

Rosch, E. 1978. Principles of Categorization. In: Rosch, E. and Barbara, B, Lloyd eds. *Cognition and categorization*. Hillsdale: Erlbaum, pp. 27–46.

Rosch, E. and Mervis, C.B. 1975. Family resemblances: studies in the internal structure of categories. *Cognitive Psychology* 7(4), pp. 573–605.

Rosch, E.H. 1973. Natural categories. *Cognitive Psychology* 4(3), pp. 328–350.

Ross, J.R. 1972. The category squish: endstation hauptwort. *Chicago Linguistic Society* 8, pp. 316–328.

Ross, J.R. 1973. Nouniness. In: Kiparsky, P. and Kenkyūjo, T. G. eds. *Three dimensions of linguistic theory*. Tokyo: TEC, pp. 137–257.

- Ross, J.R. 1974. Three Batons for Cognitive Psychology. In: Walter, B. Weimer and Palermo, D. S. eds. *Cognition and the symbolic processes*. New York: Potomac, Md : Lawrence Erlbaum Associates, pp. 63–136.
- Rothstein, S. 2008a. Telicity, atomicity and the Vendler classification of verbs. In: Rothstein, S. ed. *Theoretical and crosslinguistic approaches to the semantics of aspect*. Amsterdam: Benjamins, pp. 43–78.
- Rothstein, S. 2008b. Two puzzles for a theory of lexical aspect: semelfactives and degree achievements. In: Dölling, J., Heyde-Zybatow, T., and Schäfer, M. eds. Berlin; New York: Walter de Gruyter, pp. 175–198.
- Rothstein, S. 2010. Counting and the mass/count distinction. *Journal of Semantics* 27, pp. 343–397.
- Rothstein, S.D. 2004. *Structuring events: a study in the semantics of lexical aspect*. Malden, MA; Oxford: Blackwell.
- Roy, I. 2010. Deadjectival nominalizations and the structure of the adjective. In: Alexiadou, A. and Rathert, M. eds. *The syntax of nominalizations across languages and frameworks*. Berlin; Boston: De Gruyter Mouton, pp. 129–158.
- RStudio Team 2020. RStudio: Integrated development for R. Rstudio, PBC, Boston, MA. URL <http://www.rstudio.com/>.
- Sasse, H.-J. 2001. Scales between nouniness and verbiness. In: Haspelmath, M. ed. *Language typology and language universals: an international handbook*. Berlin; New York: Walter de Gruyter, pp. 495–509.
- Saxton, M. 2010. *Child language: acquisition and development*. Los Angeles; London: SAGE.
- Schiffrin, D. 2006. *In other words: variation in reference and narrative*. Cambridge: Cambridge University Press.
- Schmid, H.-J. 2000. *English abstract nouns as conceptual shells: from corpus to cognition*. Berlin; New York: Mouton de Gruyter.

Sedgwick, P. 2014. Multiple hypothesis testing and Bonferroni's correction. *BMJ: British Medical Journal* 349, p. g6284.

Sheskin, D., J. 2011. *Handbook of parametric and nonparametric statistical procedures*. Boca Raton, FL: Chapman and Hall/CRC Press.

SketchEngine 2017. Timestamped corpora – time annotated corpora | Sketch Engine. 8 March. Available at: <https://www.sketchengine.eu/jozef-stefan-institute-newsfeed-corpus/> [Accessed: 19 January 2022].

Smith, C.S. 1991. *The parameter of aspect*. Dordrecht: Kluwer.

Smith, C.S. 1997. *The parameter of aspect*. 2nd ed. Dordrecht: Kluwer.

Stowell, T. 1991. The alignment of arguments in adjective phrases. In: Rothstein, S. ed. *Perspectives on phrase structure: heads and licensing*. New York: Academic Press, pp. 105–135.

Strobl, C., Malley, J. and Tutz, G. 2009. An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods* 14(4), pp. 323–348.

Tagliamonte, S.A. and Baayen, R.H. 2012. Models, forests, and trees of York English: *Was/were* variation as a case study for statistical practice. *Language Variation and Change* 24(2), pp. 135–178.

Talmy, L. 1978. The relation of grammar to cognition: a synopsis. *American Journal of Computational Linguistics*, pp. 16–26.

Taverniers, M. 2003. Grammatical metaphor in SFL: a historiography of the introduction and initial study of the concept. In: Simon-Vandenberghe, A. M., Taverniers, M., and Ravelli, L. eds. *Grammatical metaphor: views from systemic functional linguistics*. Amsterdam studies in the theory and history of linguistic science. Series IV, Current issues in linguistic theory. Amsterdam; Philadelphia: Benjamins PubCo, pp. 5–34.

Taverniers, M. 2011. The syntax–semantics interface in systemic functional grammar: Halliday's interpretation of the Hjelmslevian model of stratification. *Journal of Pragmatics* 43(4), pp. 1100–1126.

- Taverniers, M. 2019. Semantics. In: Thompson, G., Bowcher, W., Fontaine, L., and Schönthal, D. eds. *The Cambridge handbook of Systemic Functional Linguistics*. Cambridge: Cambridge University Press, pp. 55–91.
- Taylor, J. 1995. Introduction: on construing the world. In: Taylor, J. and MacLaury, R. eds. *Language and the cognitive construal of the world*. Berlin; New York: Mouton de Gruyter, pp. 1–23.
- Taylor, J.R. 2002. *Cognitive grammar*. Oxford: Oxford University Press.
- Taylor, J.R. 2003. *Linguistic categorization*. Third edition. Oxford: Oxford University Press.
- Taylor, J.R. 2019. Prototype theory. In: Maienborn, C., Heusinger, K., and Portner, P. eds. *Semantics - Theories*. Berlin; Boston: Walter de Gruyter, pp. 29–56.
- Thibault, P.J. 2004. *Brain, mind, and the signifying body an ecosocial semiotic theory*. London; New York: Continuum.
- Thompson, G. 2014a. *Introducing functional grammar*. Third edition. London: Routledge.
- Thompson, G. 2014b. *Introducing functional grammar*. Third edition. London: Routledge.
- Toolan, M. 2001. *Narrative: a critical linguistic introduction*. 2nd ed. London; New York: Routledge.
- Trampus, M. and Novak, B. 2012. The internals of an aggregated web news feed.
- Valera, S. 2015. Conversion. In: Müller, P., O., Ohnheiser, I., Olsen, S., and Rainer, F. eds. *Volume 1 word formation: an international handbook of the languages of Europe*. Berlin: De Gruyter Mouton, pp. 322–339.
- Van Hout, A. 2016. Lexical and grammatical aspect. In: Lidz, J. L., Snyder, W., and Pater, J. eds. *The Oxford handbook of developmental Linguistics*. Oxford: Oxford University Press
- Van Lier, E. and Rijkhoff, J. 2013. *Flexible word classes typological studies of underspecified parts of speech*. Oxford: Oxford University Press.
- Venables, W. and Ripley, B. 2002. *Modern applied statistics with s*. New York: Springer.

- Vendler, Z. 1957. Verbs and times. *Philosophical review* 66(2), pp. 143–160.
- Vendler, Z. 1967. *Linguistics in philosophy*. Ithaca [N.Y.]: Cornell University Press.
- Verkuyl, H.J. 1972. *On the compositional nature of the aspects*. Dordrecht: Reidel.
- Vigliocco, G., Vinson, D.P., Druks, J., Barber, H. and Cappa, S.F. 2011. Nouns and verbs in the brain: A review of behavioural, electrophysiological, neuropsychological and imaging studies. *Neuroscience & Biobehavioral Reviews* 35(3), pp. 407–426.
- Vinson, D.P. and Vigliocco, G. 2002. A semantic analysis of grammatical class impairments: semantic representations of object nouns, action nouns and action verbs. *Journal of Neurolinguistics* 15(3–5), pp. 317–351.
- Vološinov, V.N. 1986. *Marxism and the Philosophy of Language*. Cambridge: Harvard University Press.
- Watanabe, S. 1985. *'Theorem of the ugly duckling', pattern recognition: human and mechanical*. New York: Wiley.
- Wickham, H. 2016. ggplot2: Elegant graphics for data analysis. Available at: <https://ggplot2.tidyverse.org>.
- Wilcox, S. 2018. *Ten lectures on cognitive linguistics and the unification of spoken and signed languages*. Leiden, The Netherlands: Brill.
- Williams, C. 2002. *Non-progressive and progressive aspect in English*. Fasano: Schena Editore.
- Williams, M. 2021. *Beginning quantitative research*. London: Sage Publications.
- Winter, B. 2020. *Statistics for linguists: an introduction using R*. New York: Routledge.
- Wittgenstein 1953. *Philosophical investigations First Edition*. Oxford: Blackwell.

## Appendix A: Study One Datasets

All Data with Repetition: [All Data with Repetition.xlsx](#)

All Data without Repetition: [All Data without Repetition.xlsx](#)

No Object Data without Repetition: [No Object Data without Repetition.xlsx](#)

## Appendix B: Study One Lexical Aspect Distribution Percentage Tables, Post-Hoc Fisher-Exact Tests and Simplified CRT Analyses

*Table AB-1. All Data without Repetition dataset situation type across word formation type.*

Situation Type	Borrowing	Compound	MDA	MDN	MDV	Other	TC	Underived
ACC	7.8%	7.9%	0%	0%	16.1%	10.9%	18.3%	3.1%
ACH	0.9%	0%	0%	0%	3.2%	0%	10.8%	1.6%
ACT	11%	1.6%	2.4%	15.4%	14%	7.8%	25.8%	5.8%
OBJ	74.9%	90.5%	48.8%	76.9%	59.1%	81.3%	40.9%	84.4%
SEM	0%	0%	0%	0%	0%	0%	1.1%	0.4%
STA	5.5%	0%	48.8%	7.8%	7.5%	0%	3.2%	4.7%

Table AB-2. Post-hoc Fisher-Exact test results on comparisons of semantic behaviour expressed by different word formation types: All Data without Repetition dataset.

Number	Comparison	p.value
1	MDV-TC	p=0.019
2	MDV-Borrowing	p=0.0035
3	MDV-MDA	<b>*p&lt;0.001</b>
4	MDV-MDN	p=0.0355
5	MDV-Compound	<b>*p&lt;0.001</b>
6	MDV-Underived	<b>*p&lt;0.001</b>
7	MDV-Other	p=0.014
8	TC-Borrowing	<b>*p&lt;0.001</b>
9	TC-MDA	<b>*p&lt;0.001</b>
10	TC-MDN	<b>*p&lt;0.001</b>
11	TC-Compound	<b>*p&lt;0.001</b>
12	TC-Underived	<b>*p&lt;0.001</b>
13	TC-Other	<b>*p&lt;0.001</b>
14	Borrowing-MDA	<b>*p&lt;0.001</b>
15	Borrowing-MDN	p=0.265
16	Borrowing-Compound	p=0.01
17	Borrowing-Underived	<b>*p=0.001</b>
18	Borrowing-Other	p=0.194
19	MDA-MDN	<b>*p&lt;0.001</b>
20	MDA-Compound	<b>*p&lt;0.001</b>
21	MDA-Underived	<b>*p&lt;0.001</b>
22	MDA-Other	<b>*p&lt;0.001</b>
23	MDN-Compound	<b>*p=0.001</b>
24	MDN-Underived	p=0.225
25	MDN-Other	p=0.0095
26	Compound-Underived	p=0.16
27	Compound-Other	p=0.229
28	Underived-Other	p=0.0435



*Table AB-3. All Data with Repetition dataset situation type across word formation type.*

Situation Type	Borrowing	Compound	MDA	MDN	MDV	Other	TC	Underived
ACC	6.7%	7.6%	0%	0%	15.6%	8.1%	16.7%	3.6%
ACH	1.1%	0%	0%	0%	2.3%	0%	10.3%	2.1%
ACT	11%	1.3%	2%	12.5%	16.4%	9.2%	25.4%	5.2%
OBJ	77%	91.1%	58%	79.2%	57%	82.8%	43.7%	86.1%
SEM	0%	0%	0%	0%	0%	0%	0.8%	0.2%
STA	4.3%	0%	40%	8.3%	8.6%	0%	4.8%	2.9%

Table AB-4. Post-hoc Fisher-Exact test results on comparisons of semantic behaviour expressed by different word formation types: All Data with Repetition dataset.

Number	Comparison	p.value
1	MDV-TC	0.0225
2	MDV-Borrowing	<b>*p&lt;0.001</b>
3	MDV-MDA	<b>*p&lt;0.001</b>
4	MDV-MDN	0.0075
5	MDV-Compound	<b>*p&lt;0.001</b>
6	MDV-Underived	<b>*p&lt;0.001</b>
7	MDV-Other	<b>*p&lt;0.001</b>
8	TC-Borrowing	<b>*p&lt;0.001</b>
9	TC-MDA	<b>*p&lt;0.001</b>
10	TC-MDN	<b>*p&lt;0.001</b>
11	TC-Compound	<b>*p&lt;0.001</b>
12	TC-Underived	<b>*p&lt;0.001</b>
13	TC-Other	<b>*p&lt;0.001</b>
14	Borrowing-MDA	<b>*p&lt;0.001</b>
15	Borrowing-MDN	0.185
16	Borrowing-Compound	0.0025
17	Borrowing-Underived	<b>*p&lt;0.001</b>
18	Borrowing-Other	0.216
19	MDA-MDN	<b>0.001*</b>
20	MDA-Compound	<b>*p&lt;0.001</b>
21	MDA-Underived	<b>*p&lt;0.001</b>
22	MDA-Other	<b>*p&lt;0.001</b>
23	MDN-Compound	<b>*p&lt;0.001</b>
24	MDN-Underived	0.052
25	MDN-Other	0.0105
26	Compound-Underived	0.096
27	Compound-Other	0.074
28	Underived-Other	0.058

*Table AB-5. No Object Data without repetition dataset situation type across word formation type.*

Situation Type	Borrowing	Compound	MDA	MDN	MDV	Other	TC	Underived
ACC	30.9%	83.3%	0%	0%	39.5%	58.3%	30.9%	20%
ACH	3.4%	0%	0%	0%	7.9%	0%	18.2%	10%
ACT	43.9%	16.6%	4.8%	66.6%	34.2%	41.7%	43.6%	37.5%
SEM	0%	0%	0%	0%	0%	0%	1.8%	2.5%
STA	21.8%	0%	95.2%	33.3%	18.4%	0%	5.5%	30%

*Table AB-6. Post-hoc Fisher-Exact test results on comparisons of semantic behaviour expressed by different word formation types: No Object Data without Repetition dataset.*

<b>Number</b>	<b>Comparison</b>	<b>p.value</b>
1	MDV-TC	0.144
2	MDV-Borrowing	0.29
3	MDV-MDA	<b>*p&lt;0.001</b>
4	MDV-MDN	0.0465
5	MDV-Compound	0.397
6	MDV-Underived	0.326
7	MDV-Other	0.354
8	TC-Borrowing	<b>*p&lt;0.001</b>
9	TC-MDA	<b>*p&lt;0.001</b>
10	TC-MDN	0.0225
11	TC-Compound	0.191
12	TC-Underived	0.011
13	TC-Other	0.313
14	Borrowing-MDA	<b>*p&lt;0.001</b>
15	Borrowing-MDN	0.154
16	Borrowing-Compound	0.106
17	Borrowing-Underived	0.0335
18	Borrowing-Other	0.14
19	MDA-MDN	<b>0.0015*</b>
20	MDA-Compound	<b>*p&lt;0.001</b>
21	MDA-Underived	<b>*p&lt;0.001</b>
22	MDA-Other	<b>*p&lt;0.001</b>
23	MDN-Compound	0.003
24	MDN-Underived	0.453
25	MDN-Other	0.0035
26	Compound-Underived	0.0765
27	Compound-Other	0.618
28	Underived-Other	0.0355

*Table AB-7. All Data without Repetition dataset situation type across count/mass status.*

Situation Type	Count	Mass
Accomplishment	8.6%	5.8%
Achievement	2.1%	0%
Activity	8.5%	16.5%
Object	80.1%	52.5%
Semelfactive	0.1%	0.3%
State	0.7%	25%

*Table AB-8. All Data with Repetition dataset situation type across count/mass status.*

Situation Type	Count	Mass
Accomplishment	6.8%	6.1%
Achievement	2%	0.4%
Activity	7.9%	17.4%
Object	82.9%	53.7%
Semelfactive	0.1%	0.2%
State	0.4%	22.2%

*Table AB-9. No Object Data without Repetition dataset situation type across count/mass status.*

Situation Type	Count	Mass
Accomplishment	43.2%	12.1%
Achievement	10.4%	0%
Activity	42.5%	34.7%
Semelfactive	0.4%	0.6%
State	3.5%	52.6%

*Table AB-10. All Data without Repetition dataset situation type across abstract/concrete status.*

Situation Type	Abstract	Concrete
Accomplishment	15.2%	2.9%
Achievement	2.9%	0.7%
Activity	18.8%	4.1%
Object	49.2%	91.7%
Semelfactive	0%	0.2%
State	13.9%	0.4%

*Table AB-11. All Data with Repetition dataset situation type across abstract/concrete status.*

Situation Type	Abstract	Concrete
Accomplishment	12.3%	2.2%
Achievement	3.2%	0.5%
Activity	18.1%	3.1%
Object	56.3%	93.8%
Semelfactive	0%	0.1%
State	10.2%	0.3%

*Table AB-12. No Object Data without Repetition dataset situation type across abstract/concrete status.*

Situation Type	Abstract	Concrete
Accomplishment	29.9%	34.6%
Achievement	5.7%	8.6%
Activity	37%	49.4%
Semelfactive	0%	2.5%
State	27.4%	4.9%

Table AB-13. All Data without Repetition dataset situation type across genre.

Situation Type	Academic Natural Science Writing	Print Advertisement Writing	(Auto)Biographical Writing	Broadsheet Newspaper Home and Foreign News Writing
Accomplishment	6.2%	4.9%	8.4%	12.2%
Achievement	2.4%	1.6%	1.1%	2.4%
Activity	7.3%	7%	11.5%	12.7%
Object	81.4%	83.2%	70.6%	67.3%
Semelfactive	0%	0%	0.2%	0%
State	2.7%	3.2%	8.1%	5.4%

Table AB-14. Post-hoc Fisher-Exact test results on comparisons of semantic behaviour expressed by different genres: All Data without Repetition dataset.

Number	Comparison	p.value
1	Academic Natural Science Writing - Print Advertisement Writing	p=0.94
2	Academic Natural Science Writing - (Auto)Biographical Writing	<b>*p&lt;0.001</b>
3	Academic Natural Science Writing - Broadsheet Newspaper Home and Foreign News Writing	<b>*p=0.0045</b>
4	Print Advertisement Writing - (Auto)Biographical Writing	<b>*p=0.007</b>
5	Print Advertisement Writing - Broadsheet Newspaper Home and Foreign News Writing	p=0.0095
6	(Auto)Biographical Writing - Broadsheet Newspaper Home and Foreign News Writing	p=0.211

*Table AB-15. All Data with Repetition dataset situation type across genre.*

Situation Type	Academic Natural Science Writing	Print Advertisement Writing	(Auto)Biographical Writing	Broadsheet Newspaper Home and Foreign News Writing
Accomplishment	5.2%	4.2%	7.3%	9.1%
Achievement	1.9%	2%	1.5%	1.9%
Activity	8.6%	6.8%	10.1%	13%
Object	82.2%	84.4%	75.1%	71.2%
Semelfactive	0%	0%	0.1%	0%
State	2.1%	2.6%	6%	4.9%

*Table AB-16. Post-hoc Fisher-Exact test results on comparisons of semantic behaviour expressed by different genres: All Data with Repetition dataset.*

Number	Comparison	p.value
1	Academic Natural Science Writing - Print Advertisement Writing	0.823
2	Academic Natural Science Writing - (Auto)Biographical Writing	<b>0.0005*</b>
3	Academic Natural Science Writing - Broadsheet Newspaper Home and Foreign News Writing	<b>0.0025*</b>
4	Print Advertisement Writing - (Auto)Biographical Writing	<b>0.0055*</b>
5	Print Advertisement Writing - Broadsheet Newspaper Home and Foreign News Writing	<b>0.002*</b>
6	(Auto)Biographical Writing - Broadsheet Newspaper Home and Foreign News Writing	0.401



*Table AB-17. No Object Data without Repetition dataset situation type across genre.*

Situation Type	Academic Natural Science Writing	Print Advertisement Writing	(Auto)Biographical Writing	Broadsheet Newspaper Home and Foreign News Writing
Accomplishment	33.3%	29%	28.7%	37.3%
Achievement	13%	9.7%	3.8%	7.5%
Activity	39.1%	41.9%	39.2%	38.8%
Semelfactive	0%	0%	0.8%	0%
State	14.5%	19.4%	27.6%	16.4%

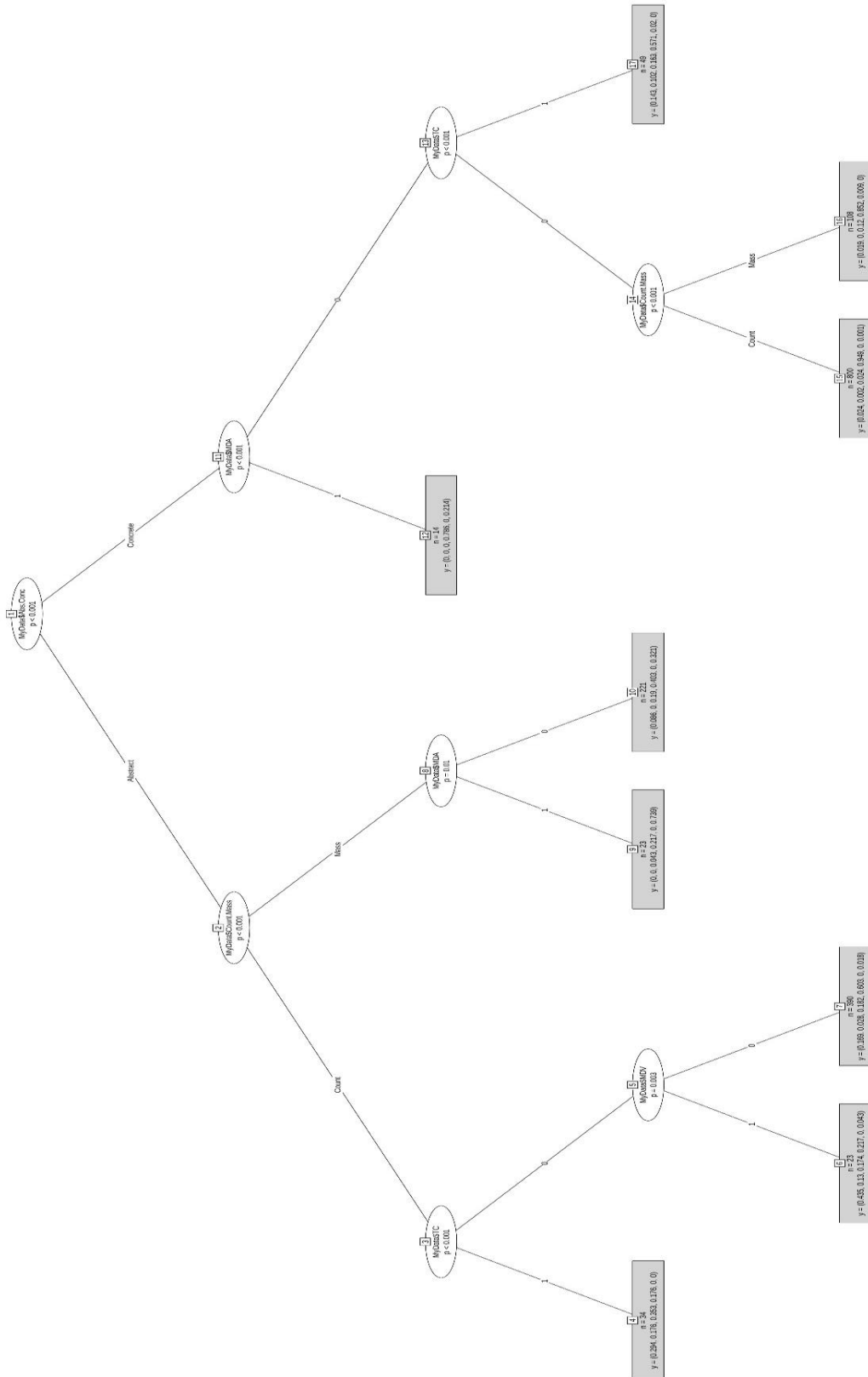


Figure AB-1. Simple decision tree visualising situation type ~ word formation type + count/mass status + abstract/concrete status + genre: All Data without Repetition dataset.

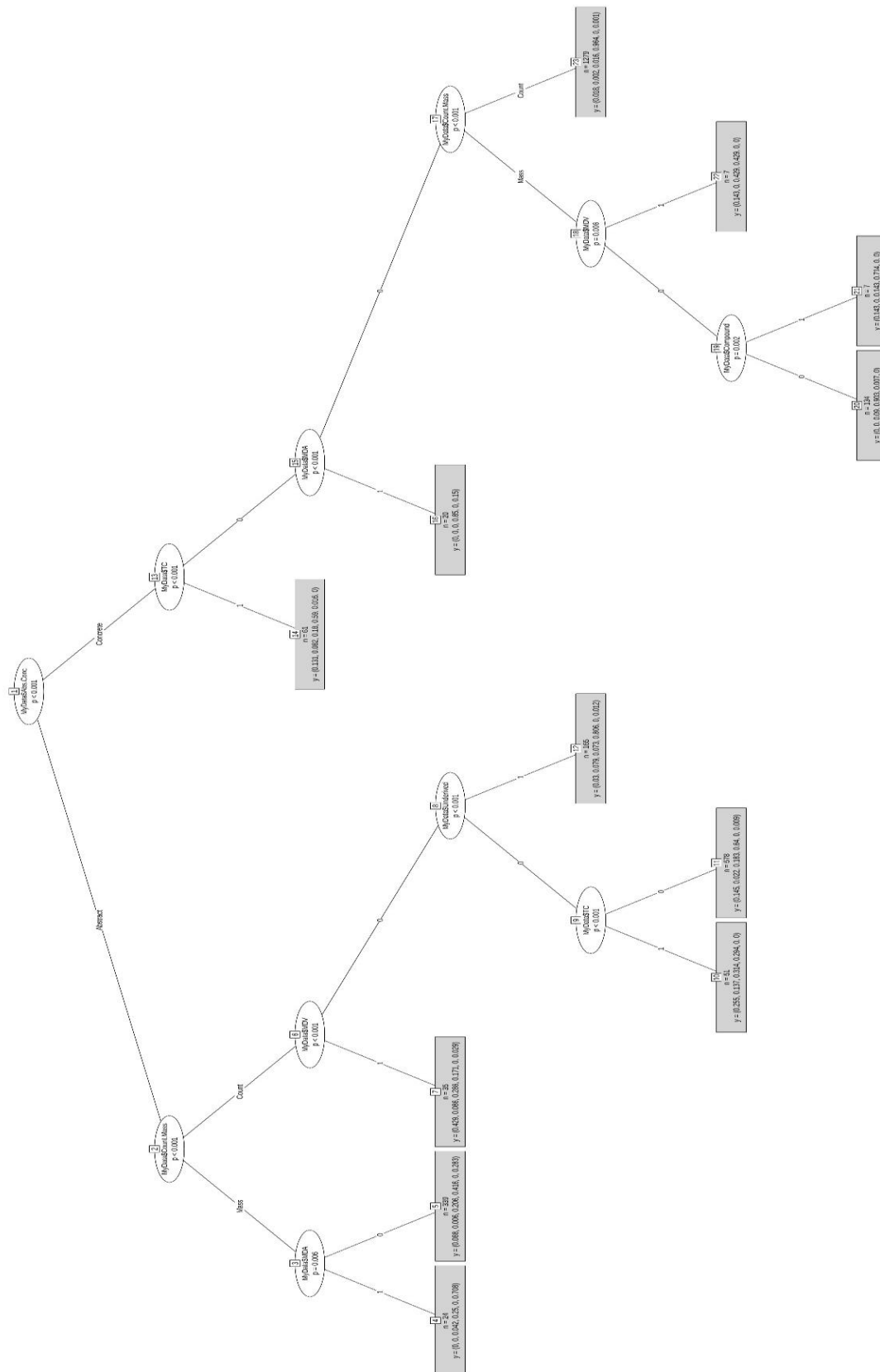


Figure AB-2. Simple decision tree visualising situation type ~ word formation type + count/mass status + abstract/concrete status + genre: All Data with Repetition dataset.

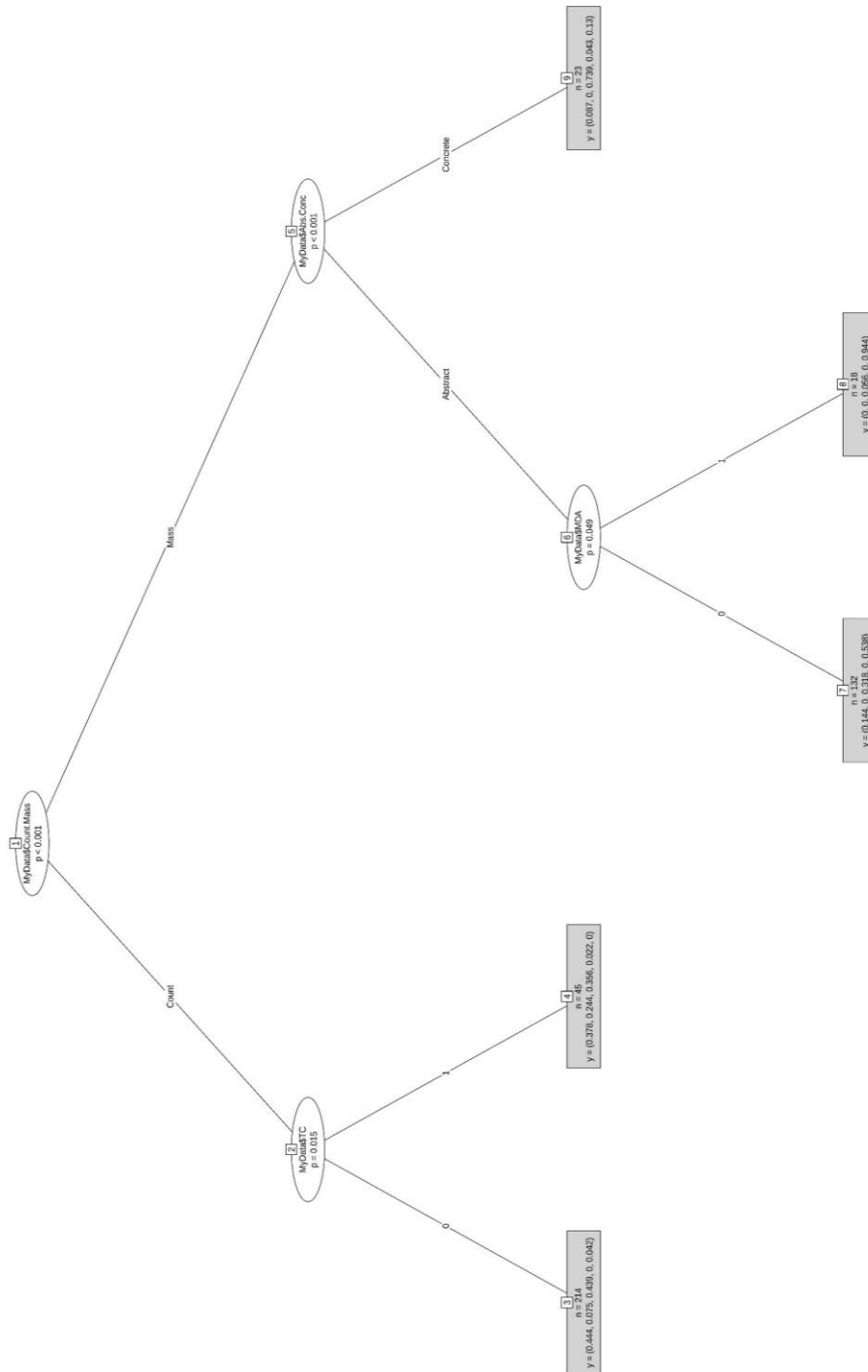


Figure AB-3. Simple decision tree visualising situation type ~ word formation type + count/mass status + abstract/concrete status + genre: No Object Data without Repetition dataset.

## Appendix C: Study Two Datasets

Game: [Game Analysis.xlsx](#)

Fire: [Fire Analysis.xlsx](#)

Stream: [Stream Analysis.xlsx](#)

Drought: [Drought Analysis.xlsx](#)

## Appendix D: R Scripts Github

Link to R scripts used in thesis: <https://github.com/alexcarr896/thesisRscripts.git>