

Disciplinary and Intra-disciplinary Developmental Variation in Shell Noun Use in Undergraduate Student Writing

by

Xiaomei Huang

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

Department of English Language and Linguistics
School of English, Drama and Creative Studies
College of Arts and Law
The University of Birmingham
September 2022

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

ABSTRACT

This study, drawing on insights from systemic Functional Linguistics and Corpus Linguistics, explores disciplinarity and intra-disciplinary developmental variation in undergraduate student writing in light of a specific linguistic aspect: the use of shell nouns (Schmid, 2000). Briefly, shell nouns are semantically abstract nouns whose meaning may only be understood by referring to their surrounding co-text. More specifically, this study aims to explore how the use of shell nouns in student writing is associated with conformity to the epistemological orientations of hard and soft science disciplinary domains. In addition, comparisons across different cohort year groups within each disciplinary domain are also carried out.

Particular attention is paid to shell nouns that occur in six grammatical patterns, consisting of four complementation patterns and two syntactic patterns headed by demonstrative *this*. These features are investigated across a sub-corpus of the British Academic Written English (BAWE) corpus (Nesi et al., 2004), representing L1 undergraduate disciplinary first-year and final-year essays from three disciplinary domains: Arts and Humanities, Natural Sciences, and Social Sciences. A corpus-based approach was adopted for the study of these linguistic features. Following this, a mixed-methods approach was then adopted through which the identified features were quantitatively and qualitatively examined at various levels of linguistic analysis. Quantitative and textual analyses revealed distinctive disciplinary flavour and some patterns of intra-disciplinary developmental variation in the use of shell nouns as observed in the BAWE data. The study shows that, a). The hard science disciplinary domain favours non-finite *to*-infinitival clauses while soft science disciplinary domain shows a preference for finite *that*- clauses as preferred shell noun complement constructions. b). In the hard science disciplinary domain, the uses of shell nouns tend to emphasize tentativeness, empirical objectivity, and scientific rationality, whereas in soft science shell-noun uses are inclined to express epistemic certainty, subjectivity

and discursiveness in the process of new knowledge-making. c). Within a given disciplinary domain, shell-noun uses are influenced by levels of study. It is suggested that the findings in the thesis may be useful for cultivating student writers' awareness of the use of language that carries disciplinary specificity.

ACKNOWLEDGEMENTS

First and foremost, my utmost gratitude and respect go to my supervisors, Dr. Paul Thompson and Dr. Garry Plappert, for their invaluable guidance throughout the different stages leading to this thesis and their encouragement over the last four years, especially during the extraordinary COVID-19 pandemic. I felt immensely proud to have exchanged many thoughts and shared countless laughs with Dr. Paul during our supervision meetings over the past four years. Dr. Garry supervised me with his academic expertise and always provided me with sound and constructive advice that helped me to further develop and improve my research. They are my lifetime mentors, not just because they enabled me to attain a degree of academic maturity, but also because showed me the personal qualities of an excellent and responsible scholar.

I would also like to thank Prof. Susan Hunston for her expertise and enlightenment. She talked with me about my research thoughts when I was in the first year of my PhD journey and kindly autographed her book for me. I would like to express my gratitude to Dr. Nicholas Groom (University of Birmingham) and Dr. Miguel-Angel Benitez-Castro (University of Zaragoza) for their invaluable assistance in reviewing my thesis and writing the referee reports required for the award of a Doctorate. Their insightful feedback and guidance were instrumental in improving the quality of my research, and I am deeply grateful for their contributions. I would like to express my sincere thanks to Dr. Matteo Fuoli, the secretary of my PhD viva, for his invaluable assistance throughout the process. His careful planning and organization helped to ensure that the viva ran smoothly and efficiently, and his attention to detail was critical in keeping everyone on schedule. Thanks are also due to my MA supervisor Dr. Nelson Flores (University of Pennsylvania) who prepared me well for PhD study. Dr. Kristina Lewis (Illinois State University) for her feedback on my MA EAP teaching practicum which has helped me cope better with my PhD journey.

My thanks also go to the staff of the department for their wholehearted support during my stay at the University of Birmingham, and to my fellow PhD students both inside and outside the department for being true friends of mine.

Last but not the least, I am immensely grateful to my beloved, open-minded, and far-sighted parents for their unconditional love, true understanding, unlimited patience, and unstinting generosities. The completion of my PhD journey would not have been possible without their support. To them all, I dedicate this thesis.

CONTENTS

Chapter 1 Introduction	1
1.1. Background to the study: EAP, EGAP and ESAP	1
1.2. Discipline and community	7
1.2.1. What is an <i>academic discipline</i> ?	8
1.2.2. The concepts of <i>community</i> and <i>discourse community</i>	11
1.3. Previous studies of disciplinary variation	13
1.4. Research questions	20
Chapter 2 Theoretical basis of the study	22
2.1. The focal concept of shell noun (Schmid, 2000)	23
2.1.1. Vocabulary 3 items (Winter, 1977; 1992)	25
2.1.2. Anaphoric nouns (Francis, 1986; 1994)	27
2.1.3. Carrier nouns (Ivanic, 1991)	28
2.1.4. Shell nouns (Schmid, 1997; 2000)	30
2.1.5. Signalling nouns (Flowerdew, 2002; 2003)	34
2.1.6. Enumerative catch-all nouns (Hinkel, 2003)	35
2.2. Shell nouns as manifestations of construals of human experience	36
2.3. Semantic typologies of shell noun	42
2.4. Grammatical features of shell nouns	45
2.5. Discourse features of shell nouns: Logico-semantic relations	56
2.6. Shell nouns and academic writing	60
2.7. Introducing the analytical framework	65
2.8. Outline of the thesis	68
Chapter 3 Methodology	69
3.1. A corpus-based approach to shell-noun uses	70
3.1.1. Corpus linguistics methodology	70
3.1.2. A corpus-based approach	71
3.2. The corpus design	73
3.2.1. Corpus selection: the BAWE corpus	74
3.2.2. Levels of study and text genre	78
3.2.3. Disciplinary domains: soft vs hard	81
3.2.4. Corpus size and type: small and specialized	85
3.3. Data collection	87
3.3.1. Data analysis software	87
3.3.2. The type of query	88
3.3.3. Semantic taxonomy of shell nouns	89
3.3.4. Procedures for data collection	93
3.4. Data analysis	96
3.4.1. Data cleaning: types of excluded items	96
3.4.2. Statistical measurements	106
3.5. Summary	109
Chapter 4 Distributions of Shell-noun Use in Disciplinary Student Writing	110

4.1. Distributions of shell nouns across three disciplinary domains	111
4.1.1. Distributions of cataphoric shell nouns	111
4.1.2. Distributions of anaphoric shell nouns	118
4.1.3. Changes of the shell-noun uses	122
4.1.4. Summary: distributions of shell nouns	123
4.2. Distributions of shell-noun lexico-grammatical patterns	124
4.2.1. Distributions of four shell-noun complement constructions	125
4.2.2. Distributions of anaphoric lexico-grammatical patterns	128
4.2.3. Changes of shell-noun syntactic patterns	129
4.2.4. Summary: distribution of shell-noun syntactic patterns	133
4.3. Distribution of types of construal instantiated by shell nouns across the three corpora	135
4.3.1. Distribution of types of construal instantiated by cataphoric shell nouns	135
4.3.2. Distribution of types of construal instantiated by anaphoric shell nouns	143
4.3.3. Changes of types of construal instantiated by shell-noun uses	147
4.3.4. Summary: distributions of types of construal	161
4.4. Distribution of the five types of transitivity process across the six SN lexico-grammatical patterns in three disciplinary domains.	165
4.5. Summary: chapter 4	168
Chapter 5 Changing Patterns of Shell-noun Use in Disciplinary Student Writing	170
5.1. Changes in AH	171
5.1.1. <i>fact</i> in the construal of <i>identifying relational</i> process	171
5.1.1.1. AH L1 writing	172
5.1.1.2. AH L3 writing	180
5.1.2. <i>example</i> in the construal of <i>identifying relational</i> process	185
5.1.2.1. AH L1 writing	187
5.1.2.2. AH L3 writing	188
5.1.2.3. Summary: shell-noun uses in the construal of <i>identifying relational</i> process.	194
5.1.3. <i>attempt</i> in the construal of <i>material</i> process	195
5.1.3.1. AH L1 writing	197
5.1.3.2. AH L3 writing	201
5.1.3.3. Summary: shell-noun uses in the construal of <i>material</i> process	206
5.2. Changes in SS	207
5.2.1. <i>fact</i> in the construal of <i>identifying relational</i> process	208
5.2.1.1. SS L1 writing	208
5.2.1.2. SS L3 writing	211
5.2.1.3. Comparison of <i>fact that</i> between the two cognate disciplinary domains .	214
5.2.1.4. Summary: shell-noun uses in the construal of identifying relational process	217
5.2.2. <i>idea, evidence, solution</i> in the construal of <i>mental</i> process	219
5.2.2.1 <i>idea</i> in creditive shell-noun use	220
5.2.2.1.1. SS L1 writing (<i>idea</i>)	222
5.2.2.1.2. SS L3 writing (<i>idea</i>)	225
5.2.2.2. <i>evidence</i> in evidential shell-noun use	226
5.2.2.2.1. SS L1 writing (<i>evidence</i>)	228
5.2.2.2.2. SS L3 writing (<i>evidence</i>)	230

5.2.2.3. <i>solution</i> in volitional shell-noun use	242
5.2.2.3.1. SS L1 writing (<i>solution</i>)	242
5.2.2.3.2. SS L3 writing (<i>solution</i>)	245
5.2.2.4. Summary: shell-noun uses in <i>mental</i> process	250
5.3. Changes in NS	251
5.3.1. <i>problem</i> in the construal of <i>attributive relational</i> process	251
5.3.1.1. NS L1 writing	251
5.3.1.2. NS L3 writing	255
5.3.1.3. Summary: shell nouns in <i>attributive relational</i> process	262
5.3.2. <i>evidence</i> in the construal of <i>mental</i> process	264
5.3.2.1. NS L1 writing	265
5.3.2.2. NS L3 writing	268
5.3.2.3. Summary: shell-noun uses in <i>mental</i> process	277
5.4. Summary: chapter 5	278
Chapter 6 Conclusion	280
6.1. Introduction	280
6.2. Summary of the main findings and answers to the research questions	283
6.3. Possible explanations for the differences in disciplinary student writing	289
6.4. Implications for EAP practitioners and curriculum designers	291
6.5. Limitations of the study	299
6.6. Directions for further research	303
6.7. Final thoughts	305
References	308
Appendices	322
Appendix 1	322
Appendix 2	325
Appendix 3	326
Appendix 4	328
Appendix 5	330
Appendix 6	331
Appendix 7	332
Appendix 8	337
Appendix 9	340
Appendix 10	343
Appendix 11	345

List of Tables

Table 2-1	39
Table 2-2	44
Table 2-3	45
Table 2-4	66
Table 3-1	78
Table 3-2	80
Table 3-3	82
Table 3-4	84
Table 3-5	88
Table 3-6	89
Table 3-7	96
Table 4-1	112
Table 4-21	136
Table 4-23	139
Table 4-24	141
Table 4-25	142
Table 4-26	143
Table 4-37	157
Table 4-38	158
Table 5-1	172
Table 5-2	172
Table 5-3	173
Table 5-4	175
Table 5-5	176
Table 5-6	176
Table 5-7	178
Table 5-8	179
Table 5-9	180
Table 5-10	181
Table 5-11	182
Table 5-12	182
Table 5-13	183
Table 5-14	187
Table 5-15	188
Table 5-16	189
Table 5-17	191
Table 5-18	196
Table 5-19	196
Table 5-20	197
Table 5-21	197
Table 5-22	198

Table 5-23	199
Table 5-24	201
Table 5-25	202
Table 5-26	204
Table 5-27	208
Table 5-28	214
Table 5-29	215
Table 5-30	220
Table 5-31	222
Table 5-32	224
Table 5-33	225
Table 5-34	234
Table 5-35	239
Table 5-36	241
Table 5-37	244
Table 5-38	246
Table 5-39	247
Table 5-40	252
Table 5-41	255
Table 5-42	259
Table 5-43	260
Table 5-44	265
Table 5-45	265
Table 5-46	266
Table 5-47	268
Table 5-48	269
Table 5-49	270

List of Figures

Figure 2.1 : Concordance lines for SN- <i>wh</i> clause	53
Figure 2.2 : Concordance lines for SN- <i>be-gerund</i>	55
Figure 2.3 : A summary of analytical framework	68
Figure 3.1 : The distribution across paper types in MICUSP corpus	75
Figure 3.2 : The distribution of disciplinary domains in the BAWE shown on the Sketch Engine online page (Kilgarriff et al., 2014)	77
Figure 3.3 : The distribution of levels of study in the BAWE shown on the Sketch Engine online page (Kilgarriff et al., 2014)	77
Figure 3.4 : The criteria options in the Sketch Engine online page	87
Figure 3.5 : The searching criteria in the Sketch Engine online page	94
Figure 3.6 : An example of the CQL for the query of shell-noun syntactic patterns	95
Figure 3.7 : Concordance lines of shell noun <i>fact</i>	95
Figure 4.1 : Distributions of five types of transitivity process across six SN lexico-grammatical patterns in AH.	166
Figure 4.2 : Distributions of five types of transitivity process across six SN lexico-grammatical patterns in NS.	167
Figure 4.3 : Distributions of five types of transitivity process across six lexico-grammatical patterns in SS.	168

List of Abbreviations

AH	Arts and Humanities
NS	Natural Sciences
SS	Social Sciences
GM	Grammatical Metaphor
SFL	Systemic Functional Linguistics
EAP	English for Academic Purposes
EGAP	English for General Academic Purposes
ESAP	English for Specific Academic Purposes
OED	Oxford English Dictionary
OALD	Oxford Advanced Learner's Dictionary
LDCE	Longman Dictionary of Contemporary English
BAWE	British Academic Written English Corpus
MCUSP	Michigan Corpus of Upper-level Student Papers
LOCNESS	Louvain Corpus of Native English Essays

Chapter 1 Introduction

1.1. Background to the study: EAP, EGAP and ESAP

The importance of communicating in written English to achieve academic success is now widely acknowledged (Hyland, 2018; Hyland and Hamp-Lyons, 2002). This is largely due to the fact that English has established a firm grip on access to universities for potential international undergraduate students (Hyland, 2018). Numerous students and academics around the world are now making concerted efforts towards gaining fluency in the ‘relatively standardized versions of academic writing in English’ (Hyland, 2018, p.383). This trend has led to a rapid development in the teaching of written English as a Second/Foreign Language in modern universities and other academic settings: the field of English for Academic Purposes (henceforth EAP). The applied nature of the EAP class and its practical role emphasize a focus on curriculum and instruction, rather than analysis and theory. In other words, although certain EAP practitioners help novice undergraduate students produce ‘standardized versions of academic writing in English’ (Hyland, 2018, p.383), they may neglect to facilitate novice undergraduate student in realizing how this standardness gradually enacts through distinctive disciplinary flavour in their writing. For example, consider the following two extracts written by L1 novice undergraduate students from the fields of biological sciences and sociology:

Disciplines	Word counts	Essay extracts
Biological Sciences	63	Extract i: <u>The animals will seem to have painful joints, retarded growth, anorexia (it will not want to eat), the loosening and maybe loss of teeth, bone deformities or easily occurring fractures, also it may drag its hind limbs.</u> This problem is very common in reptiles especially in young reptiles that develop deformities, which are then discovered by the owner too late to be rectified. (Document ID number: 6011h)
Sociology	62	Extract ii: The basic problem is that <u>statistics are collected for some other purpose and it is not possible to explore any other areas except the 'facts' already presented.</u> This limitation is obviously avoided when the sociologist is performing his or her own study and is a strong argument in favour of sociologists not relying on official statistics but seeking to generate their own. (Document ID number: 0140b)

In addition to some obvious features, such as technicality (e.g. topic of discussion), which distinguish these two extracts from each other, some subtle variations in the ways knowledge is constructed and communicated through certain linguistic features can be observed (e.g. nouns in bold). For example, there are 63 words in extract i with one noun in bold (e.g. *problem*), and 62 words in extract ii, however, with three nouns in bold (e.g. *problem*, *limitation*, *argument*). The context-specific meanings of noun phrases of this kind, headed by semantically abstract nouns, depend on the surrounding co-text (underlined). Another subtler difference is that the context-specific meaning of the noun phrase *This problem* in extract i is illustrated in the preceding segment (underlined), whereas three different noun phrases in extract ii (e.g. *the basic problem*, *This limitation*, *a strong argument*) whose context-specific meanings are actually the same, refer to the same segment (e.g. *statistics are collected ... already presented*). These subtle differences in student disciplinary writing raise interesting questions about the ways we understand and

practice EAP. On one hand, these two examples might offer evidence that how shell noun uses differ markedly across two disciplines; on the other, these examples might be just isolated examples. The point is that whether or not these differences are representative of more general trends in student disciplinary writing requires further research, which is the aim of this study.

Much debated and contested, while the notion of specificity is at the kernel of most definitions of EAP, there are debates over just how specific its purpose should be. (Hyland and Shaw, 2016, p.17). This debate is not new and perhaps more complex now as university courses become more interdisciplinary and the demands that these courses make on students become more challenging.

In the field of EAP, researchers have debated the merits and demerits of the general/specific argument. There are two positions in this debate: English for General Academic Purposes (henceforth EGAP) and English for Specific Academic Purposes (henceforth ESAP). The main theoretical justification for the former position (EGAP) emphasizes the distinction between education and training (Bodin-Galvez and Ding, 2019, p.81). EGAP is argued to be the education, while ESAP provides training (Widdowson, 1983). In this regard, education is considered a preparation for students with wide-ranging needs, and as such, it is an incremental, brick-by-brick process with a broad focus. Consequently, students need to understand a set of core linguistic forms before progressing to the acquisition of a more complex rhetorical repertoire. The basis of this idea is Bloor and Bloor's (1986) 'common core hypothesis' (Hyland, 2018, p.387), which hypothesizes that features of English are found in nearly all

varieties. By contrast, training is defined as preparing students for a restricted competence of specific tasks and thereby having a narrow focus. Huckin (2003) asserts that ESAP limits creativity and entails the reproduction of rigid conformity to conventional texts without preparing students for unpredictable future needs. In addition to this often-heard and tired argument, there are some more positive arguments for adopting EGAP. For example, Feak (2008; 2019) has argued that with the growing trend towards interdisciplinarity within tertiary education, it is the students themselves, especially post-graduate students, who are responsible for noticing disciplinary conventions rather than EAP practitioners. This broader view in turn highlights the rationale that EGAP has an interdisciplinary focus, which prepares students with the competence to gain not only the benefits from but also beyond ESAP (Feak, 2008; 2019). This perspective, to a larger or lesser extent, might invigorate the often rather emaciated arguments for the EGAP approach. Moreover, Bruce (2005, p.244), who advocates for EGAP, has provided a theoretically rigorous approach to developing an EGAP writing programme based on a cognitive discourse model. Unfortunately, perhaps due to the high demand of the theoretical and psychological efforts behind this approach (comparing to the above arguments for EGAP), ‘this cognitive genre approach to EGAP has not been as influential as it should have been’ (Bodin-Galvez and Ding, 2019, p.82).

On the other hand, by adopting a strong research orientation in EAP, research-led analysts (e.g. Thompson, 2022; Benitez-Castro, 2021; Gray, 2021, 2015a, 2015b, 2013; Viana and O’Boyle, 2021; Dong et al. 2020; Dong and Lu, 2020;

Crosthwaite and Cheung, 2019; Hyland, 2015; 2018; Anthony, 2018; Harwood, 2017; Champion, 2016) highlight the importance of placing disciplinary specificity at the heart of EAP's roles and communicative practices in particular contexts. Putting specificity into practice in the classroom often requires EAP practitioners to approach classes with a greater sense of their own expertise and to work collaboratively with subject specialists to better understand students' target discourse and course. Reasons for taking an ESAP position are discussed from two perspectives : on one hand, the teaching of disciplinary literacy skills cannot depend on subject specialists since they often lack the expertise and the desire to do so. As Lea and Street (2014) observe, most subject tutors do not have a clear picture of the role of language in their discipline. On the other, rarely do students learn language and notice disciplinarity simultaneously in an autonomous fashion. What matters in the EAP classroom is not just generic skills, such as isolated words, lexical phrases and so on, but an awareness of 'the uses of language that carry clear disciplinary values' (Hyland, 2006, p.11). By encouraging students to explore how disciplinary knowledge is constructed and communicated by means of certain linguistic features, ESAP may facilitate students to actively develop linguistic repertoires that are 'a range of literacies rather than a mere extension of general literacy to handle academic discourse'. (Bhatia, 2014, p.27).

As Hyland and Shaw (2016, p.17) argue, this debate about specificity requires EAP practitioners to 'take a stance on how they view language and learning, and to examine their courses in the light of this stance'. As an EAP practitioner, I believe that it is important for us to view these two camps of debates as 'ends of a continuum

rather than a dichotomy' (Hyland and Shaw, 2016, p.17). This is because arguments about EGAP and ESAP are both useful to a better understanding of the complexities of instructional contexts and characteristics of academic English. It is necessary to acknowledge that the relevance of these two sides of the argument regarding approaches to EAP for this thesis is that both sides have profound implications for the role of EAP practitioners. On top of that, these arguments provide insights concerning the knowledge base required to be an EAP practitioner and the types of work that EAP practitioners need to undertake to disseminate this knowledge inside and outside the classroom (Harwood, 2017, p.1).

While this thesis broadly acknowledges the idea that familiarizing students with general study skills and language formulaic patterns transferable across the academy is important, I will nevertheless continue to take a supportive stance on the part of ESAP, as it not just highlights how EAP advances as a field of inquiry and practice but also promotes EAP practitioners' awareness of approaching academic worlds by means of exploring their languages, values, genres and literacies. It is essential for EAP practitioners to remember at all times that these worlds are 'complex and evolving, conflicted and messy' (Johns, 1997, p.154).

In addition, regarding the discussion of inter-/intra-/multi-disciplinarity and so on, the truth is that only the ESAP perspective acknowledges 'the complexities of engaging in the specific literacies of the disciplines' (Hyland and Shaw, 2016, p.30) and highlights that disciplinary identity is perhaps further developed through the means of contrasting disciplinary variations rather than focusing solely on a generic

one (e.g. Bodin-Galvez and Ding, 2019; Ding and Evans, 2022). I would argue that ESAP is not restricted to a specific discipline and it is by no means defined as requiring students to mimic textual models and passively follow the decontextualized furrow laid out for them by EAP practitioners. What ESAP practitioners do is to engage students as active learners in understanding the ways in which knowledge is constructed and communicated through certain linguistic features and thus may prepare them for future inter-/intra-/multi-disciplinary assignments, as well as other challenging situations rather than simply describing academic language and then teaching it directly. As an old Chinese saying goes, ‘Give a man a fish and you feed him for a day. However, teach a man to fish and you feed him for a lifetime’. ESAP produces an agenda of academic literacy education concerned with raising students’ awareness of disciplinarity, rather than providing ‘an overarching set of value-free rules and technical skills’ (Hyland, 2018, p.392). Briefly, only by acknowledging the inseparability of disciplinary knowledge and its discourse can teachers successfully assume the role of EAP practitioners and help their students consistently achieve academic success.

1.2. Discipline and community

This section defines *academic discipline* (Trowler et al., 2012) and explains why this definition is used in the present study.

1.2.1. What is an *academic discipline*?

It is apparent that central to the teaching of academic writing of all ESAP practices is the understanding of what an academic discipline is. Although people may claim that as a common enough label, used to describe and distinguish academic knowledge, scholarship topics as well as institutional structures and so on, and that they know perfectly well what a discipline is, however, when examined in any detail, the notion can disappear into fuzziness, because the definition of *academic discipline* contains complications. As Becher and Trowler (2001, p.65) observe, ‘there is no single method of enquiry, no standard verification procedure, no definitive set of concepts that uniquely characterize each particular discipline’, and so the definition of an *academic discipline* is never straightforward. This is because disciplines have been seen in numerous ways: one way of looking at disciplines is through Kuhn’s (2011) identification focusing on their paradigms status: they have clearly established paradigms or they are at a looser, pre-paradigm stage; Biglan (1973) and Donald (1990) identify them according to faculty perceptions and Kolb (1981) draws on learning-style differences to distinguish ‘hard’ from ‘soft’ and ‘applied’ from ‘pure’ knowledge fields; Berliner (2002) distinguishes ‘hard’ (e.g. social science) from ‘easy-to-do’ disciplines (e.g. physics, chemistry, geology) in terms of the researchers’ ability to understand the phenomenon they study, make predictions and control the experiments they conduct. Others (e.g. Gilbert, 1995; Gergen and Thatchenkery, 1996) argue from post-modern perspectives that the death of disciplines has been a result of the fragmentation of academic life.

Clearly, it is crucial to treat the concept of *academic discipline* with caution. This is not only because new disciplines emerge due to interactions between existing ones, but also due to the fact that cultural and geographic factors vary among disciplines across different education systems, political ideologies, levels of economic development (Podgórecki, 1997). It is important to recognize that whilst there is a certain degree of cohesiveness and internal agreement about disciplines, an *academic discipline* is ‘as much determined by social power and members’ categorizations as it is by epistemological categories’ (Hyland, 2012, p.25). Trowler et al. (2012), writing from a social practice perspective, provide an adequate starting point of the definition of academic discipline. According to the authors, *academic discipline* has been well-explained as:

reservoirs of knowledge resources shaping regularised behavioural practices, sets of discourses, ways of thinking, procedures, emotional responses and motivations. These provide structured dispositions for disciplinary practitioners who reshape them in different practice clusters into localised repertoires. While alternative recurrent practices may be in competition within a single discipline, there is common background knowledge about key figures, conflicts and achievements. Disciplines take organizational form, have internal hierarchies and bestow power differentially, conferring advantage and disadvantage (Trowler et al., 2012, p. 9).

The most remarkable merit of this definition is the fact that it critically held that

there are division and conflict within most disciplines but also a degree of commonality (Trowler et al., 2012). It is important to note the vulnerability of the idea of disciplinary homogeneity in the 21st century, which views the boundaries of disciplines as fixed and rigid existences. In fact, instead of uniformity, we are more likely to find fluid and permeable entities that are impossible to describe with precision within disciplines.

However, academic disciplines are by no means mere fantasy; they have a very real existence for those who work and study in the world of scholarship. For instance, individuals who gain access to university, complete assignments, read journals, attend lectures, conferences and take exams all engage with the realities of these disciplines. It is in disciplines rather than particular physical sites that new knowledge occurs, because ‘disciplines have ‘real’ epistemological characteristics, that knowledge structures do condition practices in quite real ways’ (Trowler et al., 2012, p.246). These communications and interactions bring discourses, practices, texts as well as academics and so on together into a common rhetorical locale in which members find that they are similar to each other (Hyland, 2012; Durrant, 2017).

In this respect, this thesis highlights the view that members of different disciplines represent themselves and communicate with their readers in distinctive ways. Whilst academic disciplines may be ‘diffuse structures with unclear boundaries’ (Hyland and Jiang, 2018a, p.27), acquiring new knowledge in a specialized field significantly contributes to achieving professionalism. As such, academic disciplines are ‘not just sources of knowledge but also foundations of a professional identity and

the bases for shared communicative practices' (Hyland and Jiang, 2018a, p.28). In a nutshell, an *academic discipline* is a reservoir of knowledge resources at the interface of academic communications and commonplace workings where individual ideas and community expectations meet to reveal their distinct epistemologies.

1.2.2. The concepts of *community* and *discourse community*

In the field of applied linguistics, the distinctive patterns of discourses cohering around the concept of *community* have been confirmed by research (Swales, 2014a). The view of placing community at the heart of written discourse draws on ideas such as *communicative competence* in applied linguistics, *situated learning* in education and *social constructionism* in the social sciences (Hyland and Bondi, 2006, p.19). Specifically, the concept of *discourse community* draws on two concepts: *speech community* and *interpretative community*. *Speech community* (Hymes, 1972) refers to people who recognize their language uses as different from other language users, such as American English and Australian English. By contrast, *interpretative community* (Fish, 1980) highlights the social derivation of text interpretation, primarily literary texts, as it refers to a group of people who have similar ways of reading text. *Discourse community*, on the other hand, is predominately characterized by its communicative needs of a broadly agreed set of common public goals and its mechanisms of intercommunication among its members (Swales, 2014a). In a *discourse community*, members not only actively share goals but also communicate with each other to pursue these goals. Additionally, the communication and discussion

of *discourse community* is typically characterized by the use and analysis of written language, primarily within academic fields.

The concept of *discourse community* thus allows us to understand how meaning is produced in interaction and help us to identify how writers' rhetorical choices depend on purposes, setting and audience (see e.g. Bruffee, 1986; Starfield, 2001). To put it crudely and no doubt a shade polemically: a biologist is a biologist because he or she communicates like one and the same is true for the engineers, historians and linguistics. It should be mentioned that the emphasis on what is 'shared' by members of a discourse community has led to critics viewing the concept of *discourse community* as too static and structuralist (e.g. Canagarajah, 2013; Prior, 2013). These critiques caution us to bear in mind that while discourse constitutes the community's knowledge and practices, the community members are not necessarily composed of a uniform and undifferentiated mass. For example, heterogeneous individuals ranging from undergraduate student neophytes, postgraduate PhD students, professional researchers may participate in the community's activities with the same texts and genres with different questions and purposes. Whilst several issues in relation to discourse communities have not been well-defined, the concept of *discourse community* has proved significantly fruitful for the study of ESAP writing (e.g. Flowerdew, 2000; Woodward-Kron, 2004; Basturkmen et al., 2014; Zhang et al, 2020; Hu and Perez, 2022; Jwa, 2022).

In the linguistic investigation of ESAP writing, by applying the term of *discourse community* in the concept of *academic discipline*, the researcher may not just have a

more rounded and socially informed theory of texts and contexts but also a solid framework for conceptualizing the expectations, epistemic conventions and practices across disciplinary communities (Hyland and Bondi, 2006, p.20). Considerable efforts in elaborating the differences in the discoursal construction, production and communication of knowledge across disciplines have been made by researchers. I now turn to look at some of this research of how disciplinary specificity is anchored around sets of norms that regulate epistemic values and particularly linguistic choices.

1.3. Previous studies of disciplinary variation

Disciplinary variation has attracted considerable attention over the decades, with the majority of research taking place in the field of EAP based on corpus research (Thompson and Hunston, 2019, p.26). From the 1980s onwards, the literature of disciplinary variation in the early years of EAP tended to focus on register features but has more recently moved to look more at patterning, rhetorical moves, stance and at a discoursal or genre level. In this field, Hyland's (2000) *Disciplinary discourses: Social interactions in academic writing* is a key book of the study of disciplinary variation. In his analysis, Hyland (2000) investigated a range of text genres (research articles, book reviews, scientific letters, abstracts and textbook chapters) in eight disciplines (molecular biology, magnetic physics, mechanical engineering, engineering, philosophy, sociology, marketing and applied linguistics). Hyland's (2000, p.33) study is influential in revealing how writers persuade readers of their ideas by framing their messages in 'ways which appeal to appropriate relationship

with community recognized relationships', namely, epistemological conventions of a discipline. Whilst Hyland (2000) conducts reliable research by comparing different disciplines (e.g. biology and physics) in that book, he provides no discussion of how much variation is within any given disciplinary grouping (such as engineering).

Inspired by Hyland's (2000) study, many other studies have been conducted with the aim of identifying considerable differences in these epistemological conventions across disciplines by comparing more variables in disciplinary knowledge construction. Moore (2002) investigates the extent to which disciplinary knowledge is attributed to individual scholars, schools of thoughts, conventional wisdom, or a non-attributed canonical form across three disciplinary textbooks (e.g. sociology, economics and physics). His study indicates that whilst sociology and economics are typically grouped together under the social sciences, economics textbooks are more akin to those of physics than those of sociology. The latter consists of more references to social actors and processes. Moore (2002) speculates that this finding might be a result of the fact that sociology has historically failed to achieve a clear paradigm in terms of the degree of agreement among its members about how the world is seen and how subject matters are to be tackled. Alternatively, it could be due to the fact that sociology lacks clearly defined boundaries for research issues to address and is characterized by a relatively uncircumscribed sets of problems.

The issue of foregrounding a particular topic is also an important variable in disciplinary knowledge construction. In science and engineering abstracts, for example, Hyland (2000) finds that scientific writers frequently present their research

as a valuable contribution to solving real-world issues and thus lay stress on the novelty of their studies, while engineering writers emphasize the utility of their research primarily to the industrial world. By contrast, introductions in the soft fields, such as marketing, applied linguistics and sociology are more likely to be associated with unresolved disciplinary relevant problems. In the same vein, Samraj (2002) finds that biology research articles differ from wildlife behavior research articles, as the elements of persuasion and promotion are more strongly presented in the former while the latter tend to justify the research being reported.

Writing style also plays an important role in disciplinary variation in terms of knowledge construction. Hartley et al. (2004) observe clarity across three disciplines (e.g. science, social science, arts and humanities) by examining sentence lengths and Flesch Reading Ease scores, and surprisingly, they find that science texts are more readable (even though may contain more passives) than those parallel texts in the social sciences and arts fields. Hyland and Tse (2007), for instance, examine writers' sub-technical lexis and find that so-called universal items from the Academic Word List vary enormously across disciplines in terms of range, frequency, collocation and meaning. This thus undermines the assumption that there is a single core vocabulary needed for academic study.

The different use of lexical bundles is another parameter that corresponds to disciplinary variation in knowledge creation. Gray's (2015a) survey investigates the lexical distribution of different word classes, the structural complexity, and 70 lexical and grammatical features in eleven disciplines. Omidian et al. (2018) focus on the

lexical bundles in research article abstracts in six disciplines (e.g. applied linguistics, marketing, sociology, physics, biology and mechanical engineering). Both of these studies suggest that hard science writers focus on the description of data and methods, which is taken as a reflection of the epistemological orientations that guide research practice in the hard sciences. By contrast, writers in the soft fields emphasize the research purpose and implications of their study, which are readily understood by their target readership.

The frequency and the use of citation are also crucially significant in situating research and creating new knowledge. Citation is a primary way in which writers construct new knowledge by situating the latest findings in the acknowledgment of previous established facts or by challenging those facts to demonstrate novelty. The tension between originality and rhetorical accommodation contributes to the collaborative construction of new knowledge (e.g. Hyland, 2011; Thompson and Tribble, 2001). In an investigation of citation practice between soft and hard disciplines (e.g. applied linguistics and medical), Hu and Wang (2014) find similar levels of citation-based dialogic engagement between the applied linguistics and medical research articles. This finding is inconsistent with the results reported in previous studies (e.g. Hyland, 2011; Thompson and Tribble, 2001) that identified a higher citation density in soft knowledge disciplines than in hard ones. Despite the discrepancy between Hu and Wang (2014) and previous studies, it is the difference in citation forms across disciplines that reveals a typical soft vs hard divide. There is a greater dialogic expansion in the soft discipline (e.g. applied linguistics), such as

integral citations, and a lower frequency of dialogically contractive citations, such as non-integral citations. By contrast, hard discipline (e.g. medicine) is characterized by a high incidence of non-integral citations. More recently, in a study of 360 articles in four disciplines over 50 years (e.g. applied linguistics, marketing, Sociology, Biology, electronic Engineering), Hyland and Jiang (2019b) find that while citations in hard fields (e.g. biology and electric engineering) have shot up, soft fields (e.g. applied linguistics and sociology) have seen the largest rise over the years.

Different from the above nine studies about disciplinary variation, Durrant's (2017) study does not assume that disciplinary categories exist outside of the analysis but explores disciplinary variation by analyzing the use of four-word sequences across all writers in the corpus. Based on the quantitative overlap in 4-gram use across 285 authors, four main disciplinary clusters (e.g. science, humanities, technology and social sciences) and a primary distinction between soft (humanities/social sciences) and hard (science/engineering) disciplines emerge in Durrant's (2017) analysis, with two groupings (life science/commerce) lying in the intermediate point between these two. In the same vein, Thompson and Hunston (2019) extend Durrant's (2017) method that does not take interdisciplinary groupings for granted but emerge from an analysis of citation statistics.

Such observed cross-disciplinary differences found in these different variables can be ascribed to the main epistemologies underlying the soft and hard disciplines. Soft disciplines have been primarily influenced by considerable anti-positivist and anti-foundationalist epistemologies (Baert, 2005). By contrast, hard disciplines have

been greatly influenced by a positivist epistemology that assumes the legitimacy of a truth independent of human subjectivity and governed by universal laws of causality (Cohen et al., 2002). At the risk of over simplification, one reason for this disciplinary variation is that new knowledge in the hard science tends to be developed in a more linear way and characterized by highly specialized features than that of humanities and social sciences, in which new knowledge tends to emerge from an existing state of knowledge (e.g. Hyland, 2008; Omidian et al., 2021). These are just around ten disciplinary variation studies selected at random from a large number.

The point that I wish to make is that the corpus-based studies of disciplinary variation have tended to focus chiefly on the rhetorical functions. There are two approaches to research on rhetorical function: the ‘function-first’ approach, where the researcher identifies a rhetorical function first and then asks what forms are used to carry out that function (e.g. Zou and Hyland, 2022; Jiang and Hyland, 2022), and the ‘form-first’ approach, where the researcher identifies a lexico-grammatical form first, and then examines what functions that form carries out. The present study takes the latter ‘form-first’ approach. Moreover, while it has been established that there is epistemological variation between disciplines in terms of linguistic variables, almost no studies, with few exceptions (e.g. Thompson, 2005; Mbodj, 2021), have systemically applied linguistically oriented methods to reveal internal differences and patterns of variation within a given disciplinary domain.

This thesis addresses these gaps by exploring the rhetorical functions of what Schmid (2000) called shell nouns and by mapping them onto a model of grammatical

metaphor, which serves to reconstruct human experience into knowledge through the disciplinary specific construal of different transitivity processes. As Thompson (2005) argues, the existence of variations exist in approach and epistemology within disciplines necessitates that tertiary writing and literacy teaching be equipped with sensitivity towards the contexts in which writers develop their texts. The findings of such research would undoubtedly be beneficial to disciplinary writing instruction in the field of EAP.

This study takes a step in that direction by aiming to explore not only disciplinary specificity but also variation within a given disciplinary domain. The approach this thesis employs involves using corpus analysis and close reading techniques, working with a representative corpus of disciplinary student writings: British Academic Written English (BAWE) (Nesi and Gardner, 2012). By analyzing the shell noun frequencies, syntactic patterns where shell nouns occur, and the lexicalizations of shell nouns, this thesis investigates whether the use of shell nouns is an empirical dimension of distinctive disciplinary trends. If such trends exist, it examines whether these differences are consistent across various disciplinary domains and how these trends differ across levels of study. This thesis also considers some pedagogical implications and provides suggestions for the EAP/ESAP writing classroom. The following chapter expounds on the theoretical basis and presents a detailed structure of the thesis.

1.4. Research questions

As the review of previous research in this chapter has shown, it is now widely accepted that ‘differences in fields of knowledge are reflected in differences in linguistic form’ and ‘differences in linguistic form signify differences in fields of knowledge’ (Becher, 1987, p. 261). This thesis aims to build on this previous work by exploring disciplinary-specific epistemological tendencies in shell-noun usage in student academic writing, and how each distinctive disciplinary trend (if any) varies across levels of study. Based on the focal shell noun concept, the framework of grammatical metaphor, and considering the dearth of research, the overarching question addressed in this thesis is as follows:

What patterns of disciplinary and intra-disciplinary variation can be found in the use of shell nouns as manifestations of knowledge construals in first-year and third-year essay writing produced by L1 students?

The focus of this study is the use of shell nouns, which is ‘interpreted together with their shell-content complexes’ and this co-interpretation is often ‘triggered by lexico-grammatical patterns that can link shell nouns to their contents and the semantic relations underlying them’ (Schmid, 2000, p.21). Therefore, the research question is addressed from three perspectives: 1). the distribution features of shell nouns; 2). the distribution of six shell-noun syntactic patterns; 3). the distribution

characteristics of shell-noun semantic types and the lexicalizations of shell nouns. To guide the analysis of the data, the following research questions are proposed accordingly:

Question 1: What is the characteristic distribution of the shell nouns, the six shell-noun syntactic patterns, and the types of disciplinary-specific transitivity process across the three disciplinary domains ?

Question 2: How, if at all, do the three disciplinary domains vary in their construals of disciplinary knowledge according to the distribution of grammatical metaphors manifested by shell-noun constructions across the three corpora?

Question 3. What differences, if any, are evident in the uses of shell nouns across levels of study within each disciplinary domain? Have these differences been consistent across the three disciplinary domains?

Chapter 2 Theoretical basis of the study

The noun *problem* in bold in example (2.1) illustrates what Schmid (2000, p.13) regards as a ‘shell noun’.

- (2.1) The **problem** is that there are so many people unaware of how to carry this out and some that aren't willing to, that we cannot totally ensure food safety. (disciplinary domain: natural science, text: 6004c, level of study: year 1)

In example (2.1) above, *problem* is a shell noun, which semantically encapsulates the information encoded in the *that*-clause *there are so many people unaware of how to carry this out and some that aren't willing to, that we cannot totally ensure food safety*. As such, shell nouns are a group of semantically abstract nouns of this kind, whose context-specific meaning may only be fully grasped by referring to the surrounding co-texts. Their lexical meanings and references are contextually modulated in such specifics which not only endow the discourse with cohesion but also reflect writers’ epistemology in the construal of world common experience into disciplinary specific knowledge.

This thesis is concerned with the metadiscursive function of shell nouns and the uses of shell noun as manifestations of disciplinary knowledge construals. In chapter 2, I explore the literature on the notions of varied sub-types of shell nouns (e.g. vocabulary 3 items, anaphoric nouns, carrier nouns and signalling nouns as well as enumerative catch-all nouns) to lay the groundworks for the main focal concept (i.e. shell nouns) that drives the entire study: that we can infer whether or not a linguistic

choice (i.e. the use of shell nouns) is disciplinary distinctive, and to what degree, by looking at how student writers in different disciplinary domains and different levels of study reconstrue human experience into disciplinary knowledge when they use shell nouns.

The following sections first explicate why this thesis adopts the notion of shell noun (Schmid, 2000) as the focal concept (section 2.1) and then suggest the idea that views shell nouns as instances of grammatical metaphor serving to reconstruct human experience into disciplinary knowledge by means of different transitivity process construals (section 2.2). Literatures of semantic typologies, grammatical features and discourse features of shell nouns are reviewed in section 2.3, section 2.4, and section 2.5 respectively in order to suggest how we might investigate disciplinary variation and reveal distinctive disciplinary flavor in undergraduate student writing through exploring their uses of shell nouns. Section 2.6 follows that, in which a research gap will be presented by reviewing the literature on shell noun uses and academic writing. In section 2.7, I introduce an analytical framework for exploring shell-noun uses. Section 2.8 illustrates the outline of the thesis.

2.1. The focal concept of shell noun (Schmid, 2000)

The uses of shell nouns have attracted considerable attention in the literature, although with different names. In chronological order, for Halliday and Hasan (1976) they are ‘*general nouns*’, for Ivanič (1991), ‘*carrier nouns*’, for Francis (1986, 1994), ‘*anaphoric nouns*’ or ‘*retrospective and advance labels*’, for Hinkel (2001), ‘*cohesive*

nouns', for Flowerdew (2002; 2003) '*signalling nouns*' Less centrally, they are also named as '*deverbal nouns*' (Akimoto, 1990), '*enumerable nouns*' (Tadros, 1994) and '*metalanguage nouns*' (Winter, 1977; 1992). Although there are a considerable number of studies that provide different descriptions of these semantically unspecific abstract nouns by using different terminologies, their descriptions, for the most part, are often subsidiary to a more general concern with discourse structure or a partial account of shell-like units. Through a comparison of three main studies, this chapter provides justifications for the decision to focus on the concept of 'shell nouns' (Schmid, 2000), as it offers a more comprehensive encompassment of other types of definition in function and in concept.

The three studies to be presented below are *vocabulary 3* (Winter, 1977; 1992), *anaphoric nouns* (Francis, 1986) and *carrier nouns* (Ivanic, 1991). Before moving on, it is worth noting Halliday and Hasan's (1976) remarks on *cohesion and coherence*. Both of these notions refer to the connectedness of spoken or written discourse and text. The major difference between these two concepts is that *coherence* may not be linguistically-encoded since it is also related to pragmatics. In contrast, *cohesion* focuses on linguistic form and is a feature of textuality that depends on 'relations of meaning that exist within the text' (Halliday and Hasan, 1976, p.4). As Halliday and Hasan (1976) suggest, due to the importance of written discourse analysis, the concept of *cohesion* is of paramount importance in investigating how actual writing is performed. Accordingly, the metadiscursive function performed by shell nouns investigated in this thesis is firmly rooted in the notion of *cohesion*. In the present

study of shell noun usage, it is suggested that *cohesion* extends primarily in three ways: the organization of cohesive discourse, the signalling of textual patterns and the construction of discourse by recovering their meanings in the text where they occur (Halliday and Hasan, 1976; Halliday and Matthiessen, 2014).

2.1.1. Vocabulary 3 items (Winter, 1977; 1992)

Winter (1977) argues that the discourse organizing function depends on three types of clause-relational words: Vocabulary 1, Vocabulary 2 and Vocabulary 3. Vocabulary 1 and 2 comprise closed-class items such as sentence connectors and conjuncts (e.g. *whereas*, *however*), while vocabulary 3 consists of nouns, verbs and adjectives. The two main features of vocabulary 3 items are that they are characterized by cataphoric signposting function and they are vocabularies of clause relations and text relations. For the former feature, vocabulary 3 items often find ‘their lexical realizations in the following clause or group of clauses’ (Winter, 1977, p.7) and thus function as ‘anticipators for the next part of their paragraphs’ (Winter, 1977, p.9). For the latter feature, by noting that many vocabulary 3 items have analogues among grammatical words (e.g. the word *cause* has a grammatical analogue in the subordinator *because*), Winter (1977) sees vocabulary 3 items as the lexical end of a continuum of signals for expressing logical relations in English. In other words, vocabulary 3 items (Winter, 1977) can inform the reader of the logical function the noun item is representing within a textual pattern, as shown in the examples (2.2) and (2.3):

- (2.2) By accepting the European material culture, the natives were thus impelled to accept European abstract culture, especially the European religion. The **result** was that their own spiritual beliefs were subverted as they abandoned their implements for those of the white man. (Arts and Humanities, 0029n, L3)
- (2.3) A **problem** however is that the assumptions of the leniency bias could be explained by the fact that most of the data was collected using students,... (Natural Sciences, 0014e, L3)

The noun *result* in (2.2) can indicate that the passage, *by accepting the...* and *The result was that...* has a Cause/Consequence clause relationship. The noun *problem* (2.3) can be an indication that the paragraph is a Problem segment within a Problem-Solution text pattern. A list of 104 potential instances of vocabulary 3 items, comprising 60 nouns and 44 adjectives and verbs is provided by Winter (1977).

However, Winter (1977) prioritizes logical coherence relations but fails to recognize the fact that ‘the logic of *ideas* and *locutions*’ (Halliday and Matthiessen, 2014) can be readily accommodated alongside this logical coherence in the treatment of clause and text relations. In fact, even in his more recent study, Winter (1992) uses the term ‘metalinguage noun’ (Winter, 1992, p.140) to refer to nouns that ‘talk about the nature of the clause or sentence as a message in the text itself and do not refer to concrete things in the outside world’ (Winter, 1992, p.133). However, he only sporadically includes verbal and mental activity nouns. To recapitulate, the difference between metalinguage nouns (Winter, 1992) and vocabulary 3 items (Winter, 1977) is that the latter specifies the clause relation holding between two clauses, while the former provides a language for talking about how the text is organized.

2.1.2. Anaphoric nouns (Francis, 1986; 1994)

While Winter's (1977; 1992) studies tended to focus on the vocabulary of clause relations and text relations, Francis's (1986; 1994) studies have tended to do the opposite by focusing on metalanguage nouns to the detriment of nouns having to do with logical coherence relations.

Turning now to the anaphoric nouns, anaphoric nouns are nouns 'which talk about a stretch of discourse as a linguistic act, labelling it as, say, an *argument*, a *point* or a *statement*' (Francis, 1994, p.83). Furthermore, anaphoric nouns provide the readers with signposts to guide them through the discourse in that they are 'primarily interactive organizational signals with signposting function for the organization of arguments' (Francis, 1986, p.3). Anaphoric nouns often occur at a 'major division of discourse' (Francis, 1986, p.38), since the 'evaluations' (Francis, 1986, p.49) conveyed by the noun comprise 'a new topic or an aspect of the new topic' (Francis, 1986, p.38). This is shown in example (2.4):

- (2.4) In reasoning that everyone has an equal entitlement to society's natural resources it reduces the problem of having to work out and favour a particular type of the 'good life' which as was mentioned earlier has become symptomatic of our society. This **advantage** is what Loek Groot has termed the neutrality postulate and is essential when trying to decide upon what is fair and what is not as Richard Arneson points out... (Social Sciences, 0244k, L3)

By referring to a stretch of preceding discourse, the clause initial *This advantage* encapsulates the old information and starts a new focus of the discourse as to how *this advantage* is going to affect a certain decision-making procedure. In addition, in

contrast to the cataphoric function of vocabulary 3 items (see section 2.1.1), as indicated by their name, ‘anaphoric nouns’ (Francis, 1986; 1994) mainly refer backward. However, they can, at times, concurrently refer to the succeeding discourse. Francis’s (1986) study provides a list of 234 units. All the same, as acknowledged by the author herself, such a list is far from comprehensive, in that ‘any noun which can be used metadiscursively can be function as an A-noun within a discourse’ (Francis, 1986, p.7).

2.1.3. Carrier nouns (Ivanic, 1991)

Carrier nouns are countable abstract nouns that ‘carry a specific meaning within their context in addition to their dictionary meaning’ (Ivanic, 1991, p.95). Abstract nouns, such as *example*, *advantage*, *question* are all said to have a meaning which ‘remains constant’ and an additional ‘variable meaning’ (Ivanic, 1991, p.96). Based on Halliday’s Systemic Functional Grammar (SFG), the term ‘carrier’ resembles the Carrier Participant in relational process clauses, which indicates that the ‘variable meaning’ is associated with the carrier nouns’ feature of ‘not subject specific’ (Ivanic, 1991, p.96).

While Francis’s (1986; 1994) work in particular is to be credited with drawing attention to intersentential realization (i.e. cross-clause anaphora), Ivanic’s (1991) study concerns not just the intersentential function of these nouns, but also the local intrasentential function (i.e. in-clause cataphora) within Vendler’s (1968) container sentences:

Nominalization *is* N

N is Nominalization

In these container sentences, the meaning of a container noun (N) is expressed in a complement (e.g. Nominalization) in the form of a *that*-clause, *to*-infinitive clause, a *wh*-question clause or a deverbal noun. For instance, in example (2.5) below, the carrier noun *aim* is lexically specified by the *to*-infinitive clause acting as subject complement and this lexicalization is realized within the same clause.

- (2.5) In terms of M4L, the **aim** is to combine all these 'memories' into an easily searchable base. (Natural Sciences, 6139a, L1)

The carrier nouns' potential of referring is closely related to their nominal properties and thus carrier nouns can 'take different positions in the information structure of the clause' (Ivanic, 1991, p.107). More importantly, Ivanic (1991) argues that the functional role of carrier nouns is to signal or indicate a segment topic where a new topic comes into discussion or is ended, shown in the following example (2.6):

- (2.6) Squaring numbers involving two places of decimals is a tedious matter, This **difficulty** can be circumvented by using.... (Ivanic, 1991, p.99).

As shown in the sentence above, the carrier noun *difficulty* signals the organization of the local discourse (e.g. difficulty-solution pattern) by indicating a segment in which a new topic begins (e.g. how to circumvent this difficulty). To put it briefly, the main feature of carrier nouns is that they have a co-text dependent meaning that is expressed in a nominalization and plays local discourse signalling or textual organizing role in a discourse. Additionally, although providing abundant textual examples to describe the conceptualization of carrier nouns, Ivanic's (1991)

study provides a list of only 33 examples which is far from a complete list of instances.

2.1.4. Shell nouns (Schmid, 1997; 2000)

In the references reviewed above, vocabulary 3 items (Winter, 1977; 1992) and anaphoric nouns (Francis, 1986; 1994) are concerned with cataphoric and anaphoric intersentential function respectively, while carrier nouns (Ivanic, 1991) focus on local discourse organizing functions, such as the intrasentential function. These three descriptions of semantically unspecific abstract nouns discussed so far are the starting point of a comprehensive functional description of ‘shell nouns’ (Schmid, 1977; 2000).

The term ‘*Shell Noun*’ was first introduced by Schmid (1997) to refer to ‘an open-ended, functionally defined class of abstract nouns that have, to varying degrees, the potential for being used as conceptual shells for complex, proposition-like pieces of information’ (Schmid, 2000, p.4). Schmid (2000) emphasizes that shell nouns can only be understood from a functional perspective. Therefore shell nouns as such do not exhibit any inherent formal properties, which is different from other prototypical nouns such as carrier nouns (Ivanic, 1991) (countable abstract nouns) (see section 2.1.3). In fact, it is shell nouns’ uses in context that primarily endow them with ‘shell-nounhood’ (Schmid, 2000, p.13). In this respect, three essential functions that ascribed to shell nouns are also pointed out. They are: the semantic function of characterizing, the cognitive function of temporary concept formation, and the textual

function of linking.

By ‘Characterization’, it is meant that the shell nouns have the ability to name a stretch of discourse in a certain way, such as *fact* or *problem* (Schmid, 2000, p.5). ‘Temporary concept-forming’ refers to the idea that shell nouns have the potential ‘to pack different kinds of experience into one single neatly bounded conceptual entity of a thing-like quality’ (Schmid, 2000, pp.16-17). This temporarily formed concept varies according to the surrounding discourse. ‘Linking’ rests largely on the roles played by the shell nouns as anaphoric textual linkers.

Perhaps more importantly, it is argued that the functions of shell nouns are co-interpreted with their shell-content complexes. As Schmid (2000, p.21) points out, certain syntactic patterns can ‘link shell nouns to their contents and the semantic relations underlying them’, and such a co-interpretation is essentially triggered by the syntactic patterns where the shell noun occurs. Specifically, Schmid (2000) discusses the the potential for each of the four syntactic patterns is associated with any of the three functions aforementioned.

The semantic effect of characterization is most noticeable in *th-be-SN* (*th* refers to a deictic mainly *this*; *be* refers to a *be* verb; *SN* refers to a shell noun) and *SN-be-cl* (*cl* refers to a clause) patterns, as shown in example (2.7):

(2.7) For a while there I was thinking, you know, I’m gonna write pop songs, dammit. And that was a big **mistake**. (Schmid, 2000, p.329)

Furthermore, the concept-forming function is strongly associated with *SN-cl* and

SN-be-cl syntactic patterns. These syntactic patterns can exert a strong rhetorical effect, since shell nouns in this syntactic pattern can act as anchors of writers' attitudinal expressions, as topic initiators in clauses, as anaphoric linkers and signposts (Schmid, 2000, p.369). Examples (2.8) and (2.9) are two cases in point:

(2.8) Their **aim** is to meet President Saddam in Baghdad (Schmid, 2000, p.68)

(2.9) His **decision** to stand for them at the general election on June 6... (Schmid, 2000 p.330)

In (2.8), the underlined appositive *to*-clause represents an event, and it is condensed into a single temporary discourse entity: *aim*, and this can be replaced by shell nouns such as *plan*, *task*, *challenge* as discourse unfolds. In (2.9), *decision* is a piece of new information, and its meaning is partitioned as a conceptual entity in an appositive clause as if a piece of given information. These examples show some aspects of how **SN-cl** and **SN-be-cl** can function with a strong rhetorical effect.

The function of linking is mainly associated with *th-SN* and *th-be-SN* syntactic patterns. This is not to say that the linking roles are absent from the other two syntactic patterns, such as **SN-cl** and **SN-be-cl**, but that they are formed 'by means of the grammatical structure of clauses' (Schmid, 2000, p.339). Speaking of which, there are two main types of this linking function. The first one is that shell nouns are featured as anaphoric signals that summarize the preceding discourse and act as a marker of topic change. The second is that shell nouns act as cataphoric signposts to organize the internal structure of texts, which is helpful in guiding readers through the writer's arguments (Schmid, 2000, p.350).

To recapitulate, in terms of metadiscursive function, whilst acknowledging the limitation of this simplified discussion, the merit of the concept of shell noun manifests itself through the following points: shell nouns' concept-forming function includes the functions of nominalizations of carrier nouns (Ivanic, 1991); shell nouns' characterization function involves the role of topic change marker performed by anaphoric nouns (Francis, 1986; 1994); shell nouns' anaphoric linking functions encompass the discourse-organizing function of anaphoric nouns (Francis, 1986) and carrier nouns (Ivanic, 1991); shell nouns' cataphoric signposting function incorporates the cataphoric marker role played by vocabulary 3 (Winter, 1977; 1992). Therefore, it is clear that the conception of shell noun (Schmid, 2000) provides the researcher with a robust theoretical basis for the analysis of the uses of shell nouns in this study.

Furthermore, it is worth complementing the focal concept of shell nouns (Schmid, 2000) with a pedagogical concern. This concern arises from my own teaching experiences in which the most troublesome problem that my students, especially L2 students, experienced is the difficulty related to nominal cohesion in their disciplinary writing. In addition, Sarte and Gnevsheva's (2022) study about English as a Second Language (henceforth ESL) written essays reveals that noun phrasal complexity differs significantly between groups with lowest proficiency and the highest proficiency level. On top of that, as Benitez-Castro and Thompson (2015, p.401) urge, in the realm of EAP research, there is 'a clear need for more detailed research on shell-noun use in novice, semi-expert and expert disciplinary writing'. Therefore, in order to shed light on the implications for the EAP teaching of disciplinary writing,

the section below reviews the concepts of shell nouns from a pedagogical perspective.

2.1.5. Signalling nouns (Flowerdew, 2002; 2003)

Starting from a pedagogical grammar perspective, Flowerdew (2002) synthesizes seven key points for the definition of ‘Signalling nouns’ by summarizing the shell noun literature from Halliday and Hasan (1976) up to Winter (1992). These seven points are listed as follows:

- 1) The meaning of signalling nouns is only made clear by their linguistic (or non-linguistic) context.
- 2) Lexical signalling items are mainly nouns, but may also be verbs and adjectives.
- 3) The meaning of signalling nouns may be realized in three ways: across clauses, within the clause and outside the text.
- 4) Across-clause realization may be anaphoric or cataphoric
- 5) When meaning is realized across clause the lexical signal may refer to whole sections of text, not just an individual clause.
- 6) The evaluative function of signalling nouns may be introduced through the choice of noun or the pre- or post-modifiers.
- 7) When the semantic meaning of a signalling noun is specified within the clause, reference may be either general background of knowledge (exophoric reference), or the specific knowledge shared by the interlocutors (homophoric reference) (Flowerdew, 2002, pp.152-154).

A year later, Flowerdew (2003) attempts to provide a more pedagogically appropriate and comprehensive definition of signalling nouns. He provides a pedagogically-oriented description of signalling nouns by arguing that this term is applied to describe ‘any abstract noun whose meaning can be specified by reference to its context’ (Flowerdew, 2003, p.331). Perhaps the most useful aspect of these definitions is that it provides a better identification of signalling nouns’ essential features from a perspective of raising learners’ awareness of their persuasiveness in academic language. Besides, whilst acknowledging the limitation of the range of items, Flowerdew (2003) identifies 166 (written) and 112 (spoken) signalling nouns in his study. To conclude, what is important in Flowerdew’s (2002; 2003) studies is the fact that signalling nouns play a major role in the organization of scientific discourse, and thus the learners should be guided to awareness of these signalling nouns through hands-on study and analysis of authentic concordances.

2.1.6. Enumerative catch-all nouns (Hinkel, 2003)

Hinkel’s (2003) definition of shell nouns in her study of teaching academic ESL writing is mainly driven by their enumerative functions. Hinkel (2003, p.135) describes the category of enumerative catch-all nouns as lexically simple nouns referring to a stretch of discourse or individual nouns. Their ‘enumerative’ meaning allows them to refer to preceding or succeeding textual points. Besides, the function of ‘catch-all’ meaning refers to their encapsulating function. She provides a list of 63 common enumerative catch-all nouns and claims that 34 out of these 63 nouns are

highly prevalent in written discourse (Hinkel, 2003, p.284). It is suggested that teaching enumerative catch-all nouns explicitly is beneficial for preparing L2 students in the learning of academic writing, especially in the ‘highly nominalized written academic registers’ (Hinkel, 2003, p.136).

The usefulness of these aforementioned studies discussed so far lies in better defining the features of these nouns in the context of EAP teaching and learning. However, their focus is predominately on the study of L2 students, overlooking L1 student writing. Little is known about the ways L1 students use shell nouns in their disciplinary writings. In this respect, this study attempts to tackle this literature dearth in order to extend the existing knowledge about how shell nouns are used in disciplinary student writing and thus to provide some pedagogical implications for the EAP researchers and practitioners as well as curriculum designers.

2.2. Shell nouns as manifestations of construals of human experience

The present study is concerned with the question of how disciplinary variation is characterized by the distinctive disciplinary knowledge construal patterns manifested by the choice of shell nouns as meaning-making devices. Accordingly, this study proposes an approach fundamentally rooted in Systemic Functional Linguistics (SFL) (Halliday and Matthiessen, 2014). That is, shell nouns are viewed as manifestations of grammatical metaphor that arise from the nominal re-mappings between the semantic and grammatical categories (e.g. Dong et al, 2020; Fang and Dong, 2021). This can be

illustrated by example (2.10) from the Natural Sciences (NS) corpus:

(2.10') Despite his inaccuracy, he assumes the gradient remains constant, in keeping with scientific procedure.

(2.10) Despite his inaccuracy, he does state his **assumption** that the gradient remains constant, in keeping with scientific procedure. (NS, 6207a, L1)

Formally, The noun *assumption* is considered as a shell noun since it is complemented by a finite *that*-clause the gradient remains constant, in keeping with scientific procedure. This **SN-that**-clause grammatical construction is transformed from a congruent SVO construction, as shown in example (2.10') he assumes the gradient remains constant, in keeping with scientific procedure. The difference between the congruent counterpart (2.10') and (2.10) is that the latter involves a rank shift of categories from a clause to a nominal group. From a systemic functional semantic perspective, the congruent form *assume* represents a mental process. By contrast, the metaphoric form *his assumption that...* encrypts the construal of the mental process as an entity playing the role of *Verbiage* of a matrix clause of a verbal process (or symbolic process) *he does state his assumption that the gradient remains constant, in keeping with scientific procedure*. In the transitivity configuration, the choice of the shell-noun re-constructs a mental process into a verbal process: Possessive pronoun+Thing+Complement.

More importantly, the notion of 'semantic junction' (Halliday and Matthiessen, 1999, p.244) refers to a compound semantic junction between two choices of elemental semantic categories 'Process' (e.g. original verb *assume*) and 'Entity' (noun

assumption), which is triggered by such lexico-grammatical shifts. In light of meaning as choices, the transformation in form brings about the transference in meaning in grammatical metaphor. To put it simply, metaphorical forms create, profile and add new meanings to the same human experience that is encoded in the common congruent forms.

The construal of the uncommon-sense worldview of disciplinary knowledge can be comprehensively revealed through the lens of grammatical metaphor, due to its notion that construal is firmly placed in the linguistic encrypting of human cognition (Halliday and Matthiessen, 1999, pp.250-271). This thesis examines disciplinary student writing to investigate a set of special nominalizations, shell nouns, as grammatical metaphors in meaning-making and knowledge-production, aiming to contribute to the anatomy of discipline-specific epistemology and to shed some light on the path to successful development of epistemology in relation to different disciplinary clusters. This purpose is particularly pertinent in 21st century academic and education settings, where the interplay of disciplinarity is rapidly developing, and the feature of disciplinarity is more challenging to understand than ever before. (e.g. Thompson and Hunston, 2019).

Given the fact that metaphoric forms encode the same experience as congruent forms but also create and add new dimensions to their common base, this essential idea shared by grammatical metaphor and cognitive metaphor firmly places construal within the linguistic encoding of human cognition. In particular, objectification is claimed to be a general human cognitive mechanism, and thus nominalised

constructions have been paramount in providing the typical resources for grammatical metaphor, enabling the writer/speaker to re-transform world experience in the nominal mode (Halliday and Matthiessen, 1999, p.250). The semantic model of grammatical metaphor provided by Halliday and Matthiessen (1999) suggests 13 types of nominalizations, as shown in Table 2.1 below:

Table 2-1

Table 2.1: Domains of elemental metaphors (Halliday and Matthiessen, 1999, p.254).						
Congruent			Metaphorical			
		→	circumstance→	process→	quality→	→thing
Quality	→	<i>unstable</i>				1. <i>instability</i>
process	→	<i>absorb</i>			3. <i>absorptive</i>	2. <i>absorption</i>
circumstance	→	<i>Instead of: on the surface</i>		6. <i>replaces</i>	5. <i>alternative ;superficial</i>	4. <i>replacement</i>
relator for/because	→	[<i>b, for/because a</i>] so [<i>a, so b</i>]	10. <i>because of as a result</i>	9. <i>causes, proves ensues, follows from</i>	8. <i>causal; consequent</i>	7. <i>cause, proof,result</i>
∅				12. <i>occurs; imposes does, has</i>		11. <i>phenomenon ,fact</i>
thing, circumstance	→	<i>drive[be safe] decided[today]</i>	13.expansion of thing<in environment of 1 or2> <i>driver [safety] driver's [safety] [safety of the driver] today's [decision] [decision of today]</i>			

A divergence which makes type 11 different from other types is the absence of explicit transformation. In other words, the congruent form and metaphorical form in type 11 is the same. These special abstract nouns are viewed as being derived from conversion (Quirk, 2010) or zero derivation (Biber et al., 1999), but they share similar grammatical behaviour with the nominalizations derived from verbs or adjectives; noteworthy examples include *idea, fact, problem*. In other words, this set of small

lexical items is possessed of ‘clause-like semantics’ (Flowerdew and Forest, 2015, p.28) and a distinguishing feature of the ‘context-dependent use’ (Schmid, 2000, p.31).

Flowerdew and Forest (2015, p.17) suggest that ‘while they are not nominalizations in a traditional sense (they are nominalizations of not verbs and adjectives), they are nominalizations of logical coherence and relations’, and this point is demonstrated in the following example (2.11) from the Social Sciences (SS) corpus:

(2.11) Another **problem** with the concept of housing class is that it assumes collective solidarity, or at least awareness between those in an objective empirical position. (SS, 0179g, L1)

(2.11’) It assumes collective solidarity, or at least awareness between those in an objective empirical position.

The shell noun phrase, *Another problem*, which is not a nominalization, exhibits the NP complementation pattern in relation to the clause structure of hyper-theme. Meanings presented in the congruent form (as in (2.11’)) are distilled and compressed by this noun phrase. Furthermore, it enables a significant step moving from a congruent representation of content to a metaphorical encapsulation of the writer’s argument. Perhaps more importantly, it should be noted that such a shell noun phrase constitutes not just a retrospective but also a prospective link, thus providing an abstract encapsulation of the parallelism between the preceding segments and suggesting ways forward (Simon-Vandenberg et al., 2003, p.58). In addition to this logical coherence relation that is successfully established by the writer, it also highlights the writer’s nuanced interpretation of the proposition (underlined) as problematic and troublesome. Therefore, although shell nouns are not all

nominalizations in a traditional sense, such as *problem*, *fact* and *evidence*, they can also be viewed as nominalizations of a different part of the grammar, in that they are nominalizations of logical coherence relations rather than of process and qualities.

Furthermore, in view of ‘clause as representations’ (Halliday and Matthiessen, 2014), while the congruent form (2.11’) is a clause construing a mental process, the SVC clause *Another problem is that* encodes an identifying relational process, in which *problem* acts as an entity playing Identifier/Value, and *that* clause as Identified/Token. Accordingly, it is reasonable to justify the proposal that shell nouns are not just plain trans-categorizations or simple alternative grammatical realizations, but act as compound semantic choices, morphologically or syntactically functioning in the guise of an entity in order to re-construe different attributes, processes, and qualities, and facilitate the writers’ learning of building the knowledge of the world and self. As Schmid (2007, p.313) claims, shell-noun constructions, especially complement constructions, often carry ‘a meaning that goes beyond the sum of the constituents’. Shell nouns in these noun complement constructions have a prominent position through topicalization and thus, have the potential to be identified, classified, modified and qualified, to express the writer’s stance and serve as a useful resource for evaluation. Correspondingly, the objectification of human metacognitive ability and the epistemological aspects may thus be manifested by the uses of a shell noun. This aspect is shown in the following example (2.12):

(2.12) For this reason there is much evidence to support the **idea** that emotion and cognition are closely connected. (NS, 0016a, L1)

In fact, several alternatives such as *belief*, *view*, *understanding*, *notion* and so on can encapsulate the proposition *emotion and cognition are closely connected* in example (2.12). The writer might foreground a less subjective and argumentative assessment towards the proposition through his/her linguistic choice of characterizing it by using shell noun *idea*.

In an effort to uncover the disciplinary flavour from an SFL perspective, a key is to understand how semantic meaning as choice is encrypted in the patterns of construals of various transitivity processes by which world experience is transformed into disciplinary knowledge (Halliday and Matthiessen, 1999, p.510). A fine-grained categorization of shell-noun semantics that is suitable for such an understanding should be proposed.

2.3. Semantic typologies of shell noun

Whilst there is a range of classifications of the semantic types of shell noun with different scope available in the literature, two categories constantly appear in these classifications (see section 2.1.1 and section 2.1.2). One category is comprised of shell nouns construing the mental processes and mental products (e.g. *idea*, *theory*, *consideration*, *assumption*). The description of these nouns has been variously provided with different names. They are ‘*Cognition nouns*’ (Francis, 1986, pp.14-15); ‘*Mental-process nouns*’ (Francis, 1994, pp.92-93); ‘*Belief nouns*’ (Francis et al., 1998, pp.109-110); ‘*Nouns referring to something thought or believed*’ (Hunston and

Francis, 2000, p.187); or ‘*Mental nouns*’ (Schmid, 2000, pp.184-230). Another category comprises nouns construing verbal or linguistic meanings. This type of nouns is comprised of the nominalizations of verbs representing illocutionary speech acts, such as *suggestion* (*suggest*); *argument* (*argue*); *statement* (*state*) and nouns construing the propositional and informative contents, such as *news*, *myth*, *rumor* and so on.

However, as Schmid (2000, p.136) argues, ‘whereas sub-classes of linguistic shell-noun uses can be distinguished fairly ... there is no clear boundary to the class of mental shell nouns’. Briefly, it is the transitional area of a cline stretching from purely illocutionary nouns (e.g. *claim*) to purely mental nouns (e.g. *idea*) that ambiguous nouns might occur. By contrast, while the spectrum of verbal and mental shell nouns is generally uniform, there are shell nouns that do not fit into either of these two categorizations. These heterogeneous shell nouns are labeled by different names. In a chronological order, they are ‘*ownerless nouns*’ (Francis, 1986, pp.17-18); ‘*non-metalinguistic labels*’ (Francis, 1994, p.89); ‘*nouns that do not fit in any of the other groups*’ (Hunston and Francis, 2000, pp.187-188) and ‘*factual, modal, eventive circumstantial shell nouns*’ (Schmid, 2000, pp.91-279).

In terms of the ‘*ownerless nouns*’ (Francis, 1986) and ‘*non-metalinguistic labels*’ (Francis, 1994), Francis et al. (1998) provide more sub-categorizations. They are fine-grained classifications, such as desire (e.g. *purpose*, *wish*, *goal*), request (e.g. *appeal*) and so on. With regard to Schmid’s (2000) classification, factual shell nouns are those that construe facts and states of affairs, such as *thing*, *reason*, *evidence*,

difference). Modal shell nouns are used to encapsulate the likelihood or necessity of a piece of information based on the speakers' assessments, such as *possibility*, *probability* and *obligation*. Moreover, eventive shell nouns are used for the labelling of what Lyons (1977) terms as second-order entities, such as *event*, *change*, *attempt* while the circumstantial shell nouns refers to 'situations, times, locations, manners of doing things and conditions of doing things', such as *position*, *method*, *time* (Schmid, 2000, p. 275).

The difference between Francis et al.'s (1998) categorizations and Schmid's (2000) classification is that the latter provides a more general categorization of linguistic and modal categories, which subsumes some fine-grained labels of nouns conveying verbal (for example, 'request nouns', such as *appeal*) and modal meanings (for example, 'ability nouns', such as *capability*) respectively. Therefore, a typology of six main semantic features and a total of 25 sub-categorizations is provided by Schmid (2000). Table 2.2 outlines the different semantic classifications of shell nouns referred to in this section.

Table 2-2

Table 2.2: The different classifications of shell-noun semantic types in comparison.				
Francis (1986)	Francis (1994)	Francis et al. (1998)	Hunston and Francis (2000)	Schmid (2000)
Illocutionary/verbal activity nouns	Illocutionary/language activity	Suggestion, answer.	Nouns referring to something written or spoken.	Linguistic
Cognition nouns	Mental process nouns	Belief, happiness, desire and so on.	Nouns referring to something thought or believed.	Mental
Text nouns	Text nouns	Sign, reason, and so on.		Factual
		Possibility, ability and so on.	Nouns which do not fit into any categories.	Modal

Ownerless nouns	Non-metalinguistic nouns	Attempt, place, way and so on.	Eventive
			Circumstantial

It is clear that the most comprehensive semantic classification of shell nouns up to the present is provided by Schmid (2000). Table 2.3 below gives an overview of Schmid's (2000) semantic categorization of shell nouns. The complete list of types of shell nouns is comprised of a total of 25 sub-classes and 5 transitivity processes construed by shell nouns is provided in table 3.6 (see section 3.3.3 in chapter of methodology).

Table 2-3

Table 2.3: Semantic classes of shell nouns in Schmid (2000)

Classes	Examples
Factual	<i>fact, reason, problem</i>
Mental	<i>assumption, idea, aim</i>
Modal	<i>possibility, right, freedom</i>
Linguistic	<i>news, statement, question</i>
Eventive	<i>attempt, option, trouble</i>
Circumstantial	<i>time, circumstance, approach</i>

2.4. Grammatical features of shell nouns

While there is no single agreed-upon set of criteria for the identification of shell nouns, some predefined syntactic patterns or common grammatical features have been noted as typical ways for identifying shell nouns. With the exception of Vendler (1968), shell noun examples are retrieved automatically from these patterns in many large corpora studies. Structurally, shell nouns tend to occur in two types of syntactic pattern: 1). shell nouns tend to occur with the proximal demonstrative *this*, but not with distal demonstrative *that*, to form definite noun phrases.. 2). shell nouns tend to occur in the form of SN-complement clause, taking a post nominal *that-*, *to-*, and *wh-*

or gerund complement clause 3). shell nouns tend to occur in the subject position in Vendler's (1968) 'container sentence' frame in the form of SN-*be*-Nominalization (see section 2.1.3). Compared to the first grammatical feature, where the intersentential lexicalization of shell noun phrases tends to be highlighted, the emphasis of the latter two is often laid on the encapsulation within the same sentence. In the references reviewed below, we will discuss each of these features in turn.

The observation that shell nouns frequently follow demonstratives, such as *this way*, *this problem*, has been acknowledged by a number of corpus studies (e.g. Francis, 1994, p.85; Ivanic, 1991, p.111) although Hoey (1997) questioned the importance of this pattern to the word class by pointing out that nominal groups containing *such*, *another* also label a previous stretch of text. In fact, substantial evidence has shown that the pattern is indeed a common grammatical feature of shell nouns. For example, Charles's (2003) study shows that the *this*-SN structure is important for disciplinary academic writing, as it constitutes a valuable resource for textual organization and the construction of convincing arguments. Furthermore, Flowerdew and Forest (2015) provide corpus evidence from colligations that endorse Charles's (2003) view that this structure is typical of shell nouns and is an important structure in academic writing. However, while the *this*-SN feature is an important aspect of shell noun grammatical features, it may not function well as a discriminator of whether an item should or should not be counted as a shell noun, in that the *this*-SN pattern is, in fact, ubiquitous in other constructions, such as *this book* (Schmid, 2000, pp.40-42). With this said, the data retrieved (e.g. the identification of shell nouns) via this pattern requires

context-sensitive analysis and careful handling, with attention given to the wider context (see below in section 3.4.1).

As Flowerdew and Forest (2015, p.14) argue, the SN-complement clause is ‘one of the most dependable phrase level characteristics of shell nouns’. The complementation construction (e.g. SN-Complement), including three prototypical syntactical patterns where shell nouns are found to occur, is applied to the retrieval of examples in a number of studies. (e.g. Schmid, 2000; Flowerdew and Forest, 2015; Jiang and Hyland, 2015; Liu and Deng, 2017; Dong et al. 2020; Fang and Dong, 2021). Specifically, these three prototypical syntactic patterns are: SN-*that*-clause; SN-*to*-clause; SN-*wh*-clause. In Schmid’s (2000) research, this structure is labeled as ‘SN-cl’, where cl (e.g. the complement clause) provides the specifics of the SN (e.g. shell noun), and uses it along with Vendler’s ‘container sentence’ (see below) as a frame for the identification of shell nouns. Examples of these structures found in this study include (2.13) through (2.15):

- (2.13) The **idea** that humankind was split into different races and that certain races, usually white British, were superior was strong at the time, and reinforced to justify slavery and imperialism. (SS, 0001a)
- (2.14) In his **decision** to depict the moment when the people's representatives swore to die if France were not made free, David uses the Tennis Court Oath...(AH, 0252r)
- (2.15) This concept of justice as fairness is qualified by Rawls in a **situation** where everyone is under a so-called veil of ignorance in the original position. (SS, 0244k)

In fact, it is worth noting that, as Hunston and Francis’s (2000) study shows, the SN-*that*-clause is a primary identifying feature of shell nouns as a class. However,

wh- complement clauses are relatively rare in all registers, including academic prose, news, conversation and so on (Biber et al., 1999, p.647), in comparison to the other two primary types of noun complement clause (e.g. *that*-clause; *to*-clause). In the realm of corpus linguistics, there is evidence that a restricted version of the **SN**-complement pattern tends to be the most reliable and frequent syntactic pattern in which shell nouns occur. Accounts of the adoption of a restricted version of **SN**-complement as shell noun syntactic criteria can be found in many corpus studies of shell nouns, such as Sinclair et al. (1990), Francis et al. (1998), Hunston and Francis (2000); Charles (2007), Jiang and Hyland (2015), Liu and Deng (2017), Fang and Dong (2021). The first three studies are corpus-driven grammars, while the latter four studies are corpus-based investigation of shell nouns. We will discuss each of these studies in turn.

Sinclair et al.'s (1990) study of English grammar derives linguistic description from an analysis of the 20 million-word Birmingham Collection of English Texts. Both intra- and inter-sentential syntactic patterns of shell nouns are investigated with a balanced consideration in this first-ever corpus-driven research. In terms of the former, their discussion of shell nouns is based on the description of 'shell nouns' relative to the treatment of the **SN-*that*** and **SN-*to*** infinitive clauses. Nouns that are followed by *to*-infinitives tend to derive from verbs or adjectives, such as *fail to*→*failure to*, *able to*→*ability to* (Sinclair et al., 1990, p.134). Nouns that occur in *that*-clause are mainly related to reporting verbs, such as *feel that*→*feeling that*, *state that*→*statement that* (Sinclair et al., 1990, p.338). Non-derived nouns, such as *fact*,

advantage, problem denoting facts and beliefs are also followed by *that* clauses. A decade after Sinclair et al. (1990), in their grammar of nouns and adjective patterns, Francis et al. (1998) describe the phraseology associated to nouns and adjectives. Their research is driven by the 350 million-word BoE corpus, and the patterns retrieved are accompanied by semantic groups of units.

The underlying hypothesis of their research is that words with similar meaning tend to share the same pattern, implying that meaning and form are inseparable. To confirm this assumption, they narrow down the range of items by focusing on a typical shell-noun pattern: *SN-that*. In Francis et al. (1998), nouns that exhibit the ‘*SN-that*’ syntactic pattern are categorized into sets based on their semantic similarity (see section 2.3):

Nouns referring to ‘something that is written or spoken’, such as *suggestion, argument, accusation, promise* and so on.

Nouns referring to ‘beliefs, ideas, wishes, or thought processes’, such as *belief, awareness, realization* and so on.

Nouns referring to ‘emotions’, such as *amazement, gratitude, pleasure* and so on.

Nouns referring to ‘signs or evidence that something is the case’, such as *sign, evidence, clue, proof* and so on.

Nouns referring to the degrees of likelihood, such as *chance, hope, odds, probability* and so on.

Nouns with other meanings, a category made up of ‘abstract nouns which need the sort of definition or expansion that is provided by the *that*-clause’, such as *fact, problem, reason, advantage* and so on (Francis et al., 1998, p.108).

In 2000, Hunston and Francis further developed the discussion on shell nouns, building on their earlier work in Francis et al. (1998). The main difference between the two studies is that Hunston and Francis excluded two groups of **SN-*that*** nouns, namely ‘emotions’ and ‘likelihood’, from the shell nouns class. The categorization of shell nouns in Hunston and Francis (2000) is outlined below (see section 2.3):

Nouns referring to something that is written or spoken.

Nouns referring to something that is thought or believed.

Nouns referring to something that is similarly need lexicalization in a *that*-clause but do not fit into either of the above groups.

(Hunston and Francis, 2000, pp.186-187).

In contrast to previous studies that primarily focused on ‘corpus-driven’ methodologies, Charles (2007) applies a corpus-based approach to investigate the use of shell nouns by student writers in political science and material science. The study examines the **SN-*that*** syntactic pattern, like Francis et al. (1998), categorizes the nouns into semantic groups. The findings reveal the political science student writers

tend to use nouns from the *argument* and *idea* groups (as shown in (2.16) and (2.17)), whereas the *evidence* groups (as in (2.18)) is most frequently used in the material corpus. Additionally, the study shows a strong preference for unattributed nouns in *that*-clauses, enabling the construction of an ostensibly ‘objective’ stance (as in (2.19)) (Charles, 2007, pp.207-209).

(2.16) Where Waltz is original is not so much in the **claim** that the rigors of international political life produce a tendency toward interstate homogeneity, but rather in his attempt to...(pol1) (Charles, 2007, p.214)

(2.17) This disjointedness led to trouble on the ground, as it produced the **belief** that UNPROFOR was an ally of the Muslims and an enemy of the Serbs...(pol7) (Charles, 2007, p. 209)

(2.18) Despite the limited data set this provides convincing **evidence** that growth was halted by quenching and then re-initiated. (mat7) (Charles, 2007, p.209)

(2.19) However, there is a **possibility** that quenching will cause stresses in the crystal...(mat6) (Charles, 2007, p.207)

In the same vein, Jiang and Hyland (2015) examine the use of what they term ‘stance noun’ exclusively in the *SN-that* and *SN-to* syntactic patterns. By analyzing the frequencies, forms and functions of these structures in a 1.7 million-word research article corpus, they develop a functional classification of stance nouns. The study categorizes nouns exhibiting the Noun-Complements into the following classifications:

Nouns referring to **entities** do so by ‘orienting to objects, events, discourse or aspects of cognition’, such as *report*, *process*, *evidence*, *argument*, *belief*, *idea*.

Nouns relating to **attributes** concern judgments and evaluations of quality, status and formation of entities, such as *advantage*, *difficulty*, *method* and *possibility*.

Nouns relating to **relation** concern ‘how a writer understand the connection or relationship to information in a proposition’, such as *reason*, *result* and *difference*

(Jiang and Hyland, 2015, pp.7-9).

Overall, Jiang and Hyland’s (2015) research finds that nouns occurring in the Noun-Complements of **SN-that** and **SN-to** patterns tend to function as shell nouns, as they exhibit ‘a strong preference for abstraction over concretizable or objective entities’ (p.10). Additionally, the study shows that the occurrences of **SN-Complements** are more frequent in soft fields than in hard fields, and these constructions have a strong preference for shell nouns that refer to attributes in hard fields, while soft fields exhibit an even distribution of shell nouns referring to entities and attributes.

Liu and Deng (2017), like Charles (2007) and Jiang and Hylnad (2015) use a corpus-based methodology to investigate a specific type of noun complement clause. However, their study focuses specifically on **SN-be-that** clauses. The study examines popular science and professional articles and finds that the majority of shell nouns in the **SN-be-that** construction are factual nouns. Additionally, the construction has a preference for nouns of epistemic certainty in popular science writing compared to professional science writing. More recently, Fang and Dong (2021) discuss the use of

shell nouns in spoken and written registers, focusing exclusively on *SN-that*-clauses. Their observations suggest that shell noun occurrences increase with the level of formality; in other words, ‘the higher level of formality, the more frequent the use of *SN-that*’ (p. 229).

On the other hand, evidence from the data in hand (see Figure 2 below) also suggests that a particular type of syntactic pattern: *SN-wh*-clause is neither frequent nor reliable, for example, the majority of nouns occur in this type are nouns referring to first-order entities, such as names, places, as shown in the concordance lines in Figure 2.1 below:

① Essay • Politic...	"objective" report is reflected in the article in The	Guardian where guardian-N where-R	the approach was more narrative and descriptive
① Essay • Politic...	I system that ensures that every vote counts like	Malta where malta-NP where-R	every important political decision is going to be at
① Essay • Politic...	. </s><s> This is evidently displayed in the 1956	elections when election-N when-R	party identification was the key to the vote. </s></s>
① Essay • Politic...	er in which they deal with Muslims, which in turn	conditions how condition-N how-R	Muslims think of and deal with non-Muslims". </s>
① Essay • Politic...	e the controversies over the "headscarf affair" in	France where france-NP where-R	Muslims girls were suspended from school for we
① Essay • Politic...	iples illustrate that ethnic conflicts may empower	women when woman-N when-R	traditional structures of domination are shaken, e
① Essay • Politic...	ler system. Real victimisation becomes symbolic	victimisation when victimisation-N when-R	women are targeted on ethnic grounds rather tha
① Essay • Politic...	811. </s><s> Lastly, it is indeed difficult to find a	war where war-N where-R	women's groups have not "braved ethnic terror at
① Essay • Politic...	only groups that established links across former	Yugoslavia when yugoslavia-NP when-R	every other civil society network collapsed during

Figure 2.1: Concordance lines for *SN-wh* clause

Taking these aspects into consideration, it was thus decided that *SN-wh*-clause and its sub-types should be jettisoned in the present study. In terms of noun-complement (*SN-cl*) structure, this study restricts the range of items by relying on two types of noun-complement construction: *SN-that*-clause and *SN-to*-infinitive-clause.

Another major structure used to identify shell nouns is Vendler’s (1968)

‘container sentence frame (see section 2.1.3 and 2.1.4). Because ‘the container sentence is perhaps the best syntactic discriminator of shell noun membership’ (Flowerdew and Forest, 2015, p.18), Ivanic (1991) uses it in her study of carrier nouns, and Schmid (2000) adapts it to his corpus-based query techniques. In this pattern, the shell noun appears in the subject position, and the linking verb *be* is followed by a nominalization that functions as the subject complement. From a SFL perspective (Halliday and Matthiessen, 2014), the nominalization is an example of a relational identifying clause in which the subject and complement are equated and presented as having a shared identity (see section 2.2.). This pattern is a useful complement to the **SN-*that*** pattern discussed above for identifying shell nouns. This is because, as Ivanic (1991) notes, the term ‘nominalization’ is meant in a broad sense, which can include a *that*-clause, a *to*-clause, a *wh*-clause or a deverbal noun or gerund. The following examples of this pattern are taken from Ivanic’s study, with the first two examples also occurring in the **SN-*that*** clause (see examples (2.20) through (2.24))).

- (2.20) Jackson's **claim** is that this view denies or fails to adequately define the subjective features of experience. (AH, 0294c)
- (2.21) The Modernist **belief** is that language cannot be trusted as it is unable to directly describe things only hinting at a world seen in the mind. (AH, 3108e)
- (2.22) The **idea** is to allow the patient to associate alcohol with the nausea and reduce or prevent excessive drinking. (NS, 0011b)
- (2.23) The crucial **question** is why the Nazi were able to go so much further than the Democratic government in America. (SS, 0408f)
- (2.24) His **example** is crawling becoming walking - a stage of motor development. (NS, 0190d)

The same reason previously discussed (see above) and also the corpus theoretical and manual approach adopted by this thesis (see chapter 3) require modifications to tailor Vendler's 'container sentence' frame to the nature of the data at hand. However, as a shell noun discriminator, *SN-be-wh*, *SN-be-deverbal* and *SN-be-gerund* patterns are excluded from this study due to their inappropriateness and unreliability, as can be seen from the concordance lines obtained for the search of *SN-be-gerund*, as shown in Figure 2.2 below. Additionally, based on the theoretical framework (see section 2.2), the lexicalization of a given shell noun must be clause-like, which must include at least one process in the form of a finite verb. Therefore, *SN-be-gerund* and *SN-be-deverbal* are discarded in this study for these reasons.

① Essay • Busines... an Resource Management. </s><s> Today	businesses are investigating business-N be-V investigate-V	deeper and deeper into the world of Humar
① Essay • Busines... rvice industry. </s><s> Human Resources	Management is playing management-N be-V play-V	a defining role in trying to unlock the secret
① Essay • Sociolo... as evidently increased, but the prosecution	rate is dropping rate-N be-V drop-V	as is the conviction rate. </s> Problem abou
① Essay • Sociolo... a woman's responsibility and yet although	women are working woman-N be-V work-V	a full day with increasing demands, women
① Essay • Philoso... mortal account of how the world works, the	goddess is emphasising goddess-N be-V emphasise-V	that the Way of Truth is the only true interpr
① Essay • Philoso... / work also appeals to the context in which	Parmenides was writing parmenides-NP be-V write-V	. </s><s> He claims that the Way of Seemir
① Essay • Philoso... the characters can travel in that time if the	play is striving play-N be-V strive-V	for verisimilitude. </s><s> Though this is fre
① Essay • Philoso... utes, must necessarily exist. </s><s> Here	Descartes is using descartes-NP be-V use-V	'exists' as though it were a first order predic
① Essay • other ðwan completely romantices the fact that a	rape is occuring rape-N be-V occure-V	e.g. </s><s> Yeats' poem. </s><s> This cor
① Essay • other e changes in her body. </s><s> In this way	Carter is writing carter-NP be-V write-V	her own version of a fairytale, placing a ferr
① Essay • other ctim to male dominance. </s><s> Perhaps	Carter is criticising carter-NP be-V criticise-V	the system but finds that she can offer no a

Figure 2.2: Concordance lines for *SN-be-gerund*

Admittedly, the major limitation of restricting these patterns to the key grammatical characteristics of shell nouns is that this action may seem too restrictive. However, it should be reiterated that the purpose of this study is not to provide a linguistics description of shell nouns but to identify differences in their usage in

constructing and communicating knowledge. Therefore, with respect to Vendler's (1968) 'container sentence' frame (Noun-*be*-nominalization), it seems reasonable to narrow down the range of potential shell noun items by focusing on SN-*be-that* and SN-*be-to* syntactic criteria.

In conclusion, this study attempts to identify patterns of shell noun usage that may reveal disciplinary variation in terms of knowledge construction by examining the deployment of six grammatical patterns: *this*-SN, *this-be*-SN, SN-*that*-clause, SN-*to*-infinitive-clause, SN-*be-that*-clause, SN-*be-to*-infinitive (see section 2.2).

2.5. Discourse features of shell nouns: Logico-semantic relations

In addition to grammatical and semantic features, discourse features are crucially relevant to the characterization and identification of shell nouns. If we return back to Winter's (1977) conception of Vocabulary 3 items (see section 2.1.1), these nouns are lexical signals of coherence relations in discourse. In his accounts of coherence in text, Winter (1977) highlights the importance of a small system that includes the *basic clause relations* and *basic text structures*. The former includes *matching relations* and *logical sequences*. *Matching relations* include relations such as general-specific, exemplification, encapsulation, comparison, contrast and incompatibility while the *logical sequences* include relations such as cause-effect, assessment-basis, means-purpose, and condition-consequence. The latter includes *situation-evaluation*, such as the *problem-solution* structure (e.g. Hoey and Winter, 1983; Flowerdew, 2003)

and *hypothetical-real* structures, such as denials and corrections, expectations and counter-expectations and so on. It is important to note that the distinctiveness of Winter's (1977) vocabulary 3 items lies in the idea that these abstract nouns function as signposts in text, indicating how particular clauses are meant to be interpreted in relation to other clauses (*basic clause relations*) and explaining how long stretches of text are meant to be interpreted in relation to each other (*basic text structures*). However, Winter's (1977, 1992) predominant focus was on the nouns of logical coherence, such as *fact* and *problem*, with sporadic inclusion of mental and verbal nouns. In contrast, Francis's (1986; 1994) study was primarily based on metalanguage nouns, including sub-classes of cognitive nouns such as *idea*, *conception* and verbal activity nouns such as *argument*, *claim* and so on, with peripheral mention of logical coherence nouns. This study thus intends to supplement these previous studies by providing a theoretical justification for an integrated description of shell nouns that accommodates the inclusion of mental and verbal activity nouns alongside logical coherence nouns on an equal footing, namely the system of Logico-semantic relations (Halliday and Matthiessen, 2014)

In systemic functional grammar (Halliday and Matthiessen, 2014), the Logico-semantic relations include two primary types of coherence relations: *projection* and *expansion*.

Projection accounts for the linguistic representation of thoughts and feelings and sets up a semiotic system for relationships such as direct and indirect speech, reports and paraphrases. It refers to 'the logico-semantic relationship whereby a clause comes

to function not as a direct representation of (non-linguistic) experience but as a representation of a (linguistic) representation' (Halliday and Matthiessen, 2014, p.441). In Francis's sense, projection pertains to the text about text, talk about talk and the labelling of a stretch of discourse as a particular kind of linguistic or semiotic act. Specifically, it refers to the representations of ideas (as in (2.25)) and locutions (as in (2.26)).

- (2.25) The concept of mood congruity refers to the **idea** that "emotionally toned information is learned best when there is correspondence between its affective value and the learner's current mood state" (Eysenck & Keane, 2000). (NS, 0016a)
- (2.26) One major criticism of the behaviouristic approach addresses the **claim** that behaviour can be explained without reference to mental activity. (NS, 0011b)

Expansion accounts for various kinds of comparisons and contrasts, temporal, spatial and casual relationships in discourse, among others. It construes a nexus between two clauses that are logico-semantically joined and comprises three sub-types of relations between clauses. One clause either *elaborates* on the meaning of another through identification, specification or clarification (as in (2.27)), *extends* the meaning of another by providing pure addition, addition with an adversative feature or alternative (as in (2.28)) or *enhances* the meaning of another by referring to circumstantial features such as manner, cause and condition (as in (2.29)).

- (2.27) The **problem** is that more recent research by Kagan et al (1980) found that so long as day care is well equipped and staffed, there is no difference in a child's intellectual and emotional development. (NS, 0304a)

- (2.28) This technique was developed by Joseph Wolpe and involves understanding what provokes least and most anxiety, learning relaxation techniques, then associating relaxation with the least provoking stimuli. With this **method**, patients are allowed to gradually overcome their phobias. (NS, 0017b)
- (2.29) This can be interpreted in a **way** that females may have more experience with being sexually harassed than males, and as a result, have schemas which reflect this. (NS, 0014e)

The notions of *projection* and *expansion* are useful in understanding the functioning of shell nouns since they provide insight into the writers' decision-making process. For example, they allow for an interpretation of example (2.25) *The **problem** is that more recent research by Kagan et al (1980) found that so long as day care is well equipped and staffed, there is no difference in a child's intellectual and emotional development* as a case of expansion, where the logico-semantic relation is one of elaboration. This linguistic decision can be understood as the writer's strategic rhetorical choice, since the writer could have chosen projection by using a shell noun such as *claim* or *argument*. In this case, presenting information as 'a figure of being' rather than 'a figure of sensing' (Halliday and Matthiessen, 1999, pp.144-147) may have been more convincing and effective for the writer. The choice of the shell noun (e.g. *problem*) is also significant, as the writer could have opted for shell nouns that play a role in coherence relations between clauses, such as *fact*: *The fact is that more recent research....* This suggests that there is a possibility that the writer might be either concerned about its 'causation-related' function at a discourse level (e.g. Flowerdew, 2003) or its 'evaluation-related' function at a local sentence level (e.g. Schmid, 2000; McEnery and Kifle, 2002) rather than mere elaboration.

Moreover, as discussed earlier, the majority of shell noun taxonomies in the literature take a position on the relative importance of verbal and mental nouns compared to logical relations nouns. While Schmid (2000) provides the most thorough treatment of shell nouns to date, including all types equally, he overlooks the relationship between cognitive nouns, linguistic nouns and nouns expressing logical relations in text. This is because Schmid's (2000) analysis of what he terms 'prime', 'good' and 'less good' shell nouns is primarily based on their abstractness. The present study benefits from Halliday's account of logico-semantic relations, which provides a useful linguistic framework that can address the over-reliance on syntactic tests as evidence of shell noun status and establish a firm theoretical foundation that is sensitive to discourse analysis (see below in section 2.8).

2.6. Shell nouns and academic writing

Becher and Trowler (2001) propose a standard typology for the classification of academic knowledge that distinguishes between 'hard' and soft' disciplinary domains (see section 1.2). Typically, knowledge in hard disciplines is quantitatively oriented and tends to be characterized by steady and cumulative development, where new knowledge derives linearly from an existing body of knowledge. Soft knowledge, on the other hand, is qualitatively oriented, and new knowledge tends to derive discursively from the combination and recombination of existing work and results (Becher, 1989, p.13; Becher and Trowler, 2001, p.39).

Academic writing has spawned major investigatory efforts in both research and

practice in the linguistic inquiry of disciplinary variation (see section 1.3). Prominent studies on this topic include citations in PhD theses (Thompson and Tribble, 2001), evaluative *that* in research article abstracts (Charles, 2007), multi-word expressions in research articles (Omidian et al., 2018), existential *there* in research articles (Jiang and Hyland, 2020), metadiscourse in research articles (Hyland and Jiang, 2018a), and sentence-level lexical features in manuscript reviews (Samraj, 2016). Findings from this body of research have shown that disciplinary variation arises from distinct epistemological values and the conformity to academic community conventions (e.g. Omidian et al., 2018). Moreover, such conformity to norms can manifest itself in the rhetorical choices and linguistic resources used in the discursal construction of knowledge (e.g. Omidian et al., 2018; Hyland and Bondi, 2006). However, while the study of disciplinary variation has a strong foundation in written English research, most of the literature in this area has primarily focused on professional writing (e.g. Flowerdew and Forest 2015; Jiang and Hyland, 2018) and has little to say about shell nouns in disciplinary student writing.

This being the case, valuable insights into disciplinary shell-noun uses may be gleaned from five studies on student writing: Flowerdew (2008), Caldwell (2009), Sing (2013), Thompson (2014), Benitez-Castro (2021).

Flowerdew's (2008) corpus study is based on two corpora, one of professional writing and the other of undergraduate writing. She investigates how the Problem-Solution discourse pattern is signalled differently through the use of 'inscribed and evoking items' (Flowerdew, 2008, p.55), including shell nouns. The

study finds that professional writing uses a wide range of inscribed and evoking items while students writing heavily relies on a restricted set of inscribed items. The most notable finding in Flowerdew's (2008) study is the association between the shell noun *problem* and the reason-result pattern. Example (2.30) below illustrates this association, where the reason segment is *works at the tunnel portal* and the result is *a noise problem*.

(2.30)works at the tunnel portal will create a noise **problem**. (Flowerdew, 2008, p.58)

Whilst Flowerdew (2008) focuses on a specific discourse pattern, Caldwell (2009) emphasizes how the use of shell nouns in professional writing, as well as in native and non-native student writing, can contribute to lexical vagueness. By analyzing the distribution of a range of shell-noun patterns and the lexicalizations of shell nouns across the three corpora, Caldwell (2009, p.177) suggests that it is the 'particular ways of using abstract nouns' that give rise to vagueness in academic writing, rather than the mere 'occurrence of abstract nouns'. It is worth noting that whilst the uses of shell nouns in novice academic writing may be associated with lexical vagueness, this is not necessarily the case in professional academic writing.

Although both Flowerdew's (2008) and Caldwell's (2009) analyses consider professional writing, Sing's (2013) study specifically focuses on L1 (native) and L2 (non-native) student writing across three disciplines (e.g. economy, business and sociology), omitting professional writing. By examining only the use of two shell-noun syntactic patterns: **SN-cl** and **SN-be-cl**, Sing's (2013) study suggests an

overall preference for these two shell-noun syntactic patterns. However, there is a greater tendency to use three of the four SN-cl and SN-be-cl variants in L1 student writing, whereas only the SN-be-to variant is more common in L2 student writing. Thus, it is argued that L1 student writing features a wider range of shell noun types in these four syntactic variants, whereas L2 student writing exhibits a limited range of shell-noun types that occur repeatedly in a specific pattern, such as fixed noun-phrases like *the purpose of this paper is to* (Sing, 2013, p.420).

Thompson's (2014) study examines the associations between paragraph structure (e.g. paragraph-initial, paragraph-medial, paragraph-final) and specific words and grammatical or P-frames (e.g. *one of the...*) in essays using an automated analysis of the BAWE corpus. The study demonstrates that shell nouns are strongly linked to paragraph initial sentences (Thompson, 2014, p.364), and certain typical shell-noun P-frames (e.g. *one of the problems*) are similarly primed for paragraph-initial positions. Furthermore, the paragraph-initial shell nouns and frames are favored by third-year essays compared to first-year essays (Thompson, 2014, p.366). Third-year writing is considered to be more matured and improved as it incorporates both retrospective and prospective functions provided by hyper-themes in the organization of professional writing.

It is noteworthy that Benitez-Castro's (2021) study, unlike previous studies, examines specific shell noun uses (*problem* and *way*) across disciplines (sociology, business and engineering) rather than analyzing the overall frequency of shell nouns or their syntactic patterns. This approach provides a more nuanced understanding of

how shell nouns are used in disciplinary writing and how they are associated with specific discourse features and epistemological values. The study highlights the importance of considering specific shell noun uses in relation to disciplinary context, as well as the potential of shell nouns to convey disciplinary knowledge and expertise. Specifically, the uses of shell noun *problem* in both engineering and business writing show a similarity in the exophoric realizations while sociology writing differs markedly from engineering writing. Since the emphasis of *way*'s shell-noun uses is often laid on explanation and specification of inner workings of certain techniques and procedures, engineering writing shows a less compressed style. Business and sociology writings are closer in the shell-noun uses of *way*.

From all these studies reviewed in this section, although shell nouns 'are often treated as belonging to allegedly all-purpose lexical repositories that remain unaffected by disciplinary ways of meaning' (Benitez-Castro, 2021, p.146), it appears that there are considerable disciplinary associations between shell-noun uses and the practices of meaning-making. While these analyses have provided a very detailed picture of differences between L1 and L2 writing or between professional and novice writing, far less is currently known about disciplinary variation in L1 students writing. Such disciplinary variation in student writing clearly needs further investigation in terms of rhetorical purpose and epistemological orientations, as they may bring to light the pedagogical implications for EAP/ESAP practitioners and curriculum designers.

This study is a response to the specific need of substantiating existing knowledge

about how shell noun functions in reconstructing human experiences into disciplinary knowledge. The present research utilizes the concept of grammatical metaphor in Systemic Functional Linguistics. I expound on this premise fully in section 2, where I provide an overview of grammatical metaphor theory, describe the proposal of treating the Schmid's (2000) shell nouns as instances of transformations of human experience into disciplinary knowledge through different transitivity processes, and develop the semantic typologies. I characterize the grammatical and discourse features (see sections through 2.2 through 2.5) used for exploring the use of shell nouns as manifestations of disciplinary knowledge construals supporting linguistic variations across the three disciplinary domains.

2.7. Introducing the analytical framework

The existing bulk of studies on shell nouns have primarily been conducted from a lexical or phraseological perspective, focusing on describing a particular word class (e.g. service list of shell nouns) or producing a coherent taxonomy of a shell nouns (e.g. lexical semantic groupings of shell nouns) (e.g. Francis et al., 1998; Hunston and Francis, 2000; Schmid, 2000; Flowerdew, 2003). However, this proposed model aims to substantiate existing knowledge about how shell nouns function in reconstructing human experiences into disciplinary knowledge. This model is based on the grammatical metaphor theory in systemic functional linguistics (see sections through 2.2 through 2.5), which offers a coherent theoretical framework for addressing the diverse types of shell nouns within a unified account, and for analyzing their semantic

features. It is important to note that shell nouns are not limited to nouns derived from verbs, such as *belief*, *argument* or adjectives such as *instability*, but also include nouns associate with mental activities, such as *idea*, and logical relations in text, such as *problem*, *fact* (see section 2.2).

To provide a comprehensive account of how shell nouns reflect disciplinary preferences in the construction of knowledge in different domains, it is necessary to view them not just from a syntactic (see section 2.4) or cognitive perspective (see section 2.3), but also from the standpoint of discourse (see section 2.5). This is because shell nouns serve as useful nominal resources for the organization of discourse (e.g. Winter, 1977; 1992; Flowerdew and Forest, 2015). Therefore, a primary semantic criterion for identifying shell noun status is that the lexical realization of a shell noun must construe the semantics of a clause in the form of a full clause (e.g. the construal of process), which can be applied to the major types of shell noun realizations without controversy. For instance, all instances of *basis* in the following examples are shell nouns (see section 3.3.1 data cleaning).

Table 2-4

Table: 2.4 Major types of shell noun realizations in this study

<i>this</i> -SN	(2.31)	<u>...the quality of products is directly influenced in nine areas, in what he calls the 9Ms of 'Markets, Money, Management, Men, Motivation, Materials, Machines and Mechanization'</u> . Using this basis I have constructed a fishbone diagram below which can be used to delve into the problems at EHL.(SS, 0202K)
<i>this-be</i> -SN (across-clause)	(2.32)	<u>Political issues by themselves may consist of definite and objective facts and figures, but the interpretation and presentation of political issues will always be a product of the writer's analysis and perspective.</u> This is the basis for the political issue to be discussed.(SS, 0137d)
SN-Clause (within clause)	(2.33)	This essay rests on a fundamental basis that <u>there is an intricate link between the abstract world of theory and the real world of</u>

policy. (SS, 0137g)

SN-*be*-Clause (within-clause) (2.34) As mentioned earlier, the very **basis** of the right to silence is that it is a basic constitutional right of the individual not to have to answer questions and for no adverse inferences to be drawn through their refusal. (SS, 0411c)

The proposed model for analyzing shell nouns is based on four main aspects: lexical, semantic, grammatical and discourse features. These aspects are summarized and defined in Figure 2.3. The model is both deductive and inductive. On one hand, it applies predefined syntactic patterns and previously proposed taxonomy of semantic categories of shell nouns, such as Schmid's (2000) work on syntactic patterns and taxonomy of semantic categories, and Halliday and Matthiessen's (2014) work on knowledge construals and discourse features. On the other hand, it is inductive as it uses authentic data to revise and modify certain aspects of Schmid's taxonomy of semantic categories in order to better support EAP teaching and research. Overall, the proposed analytical model is analytically empirical, as it considers different perspectives and uses authentic data both reflect on existing studies and produce a revised semantic classification of shell nouns.

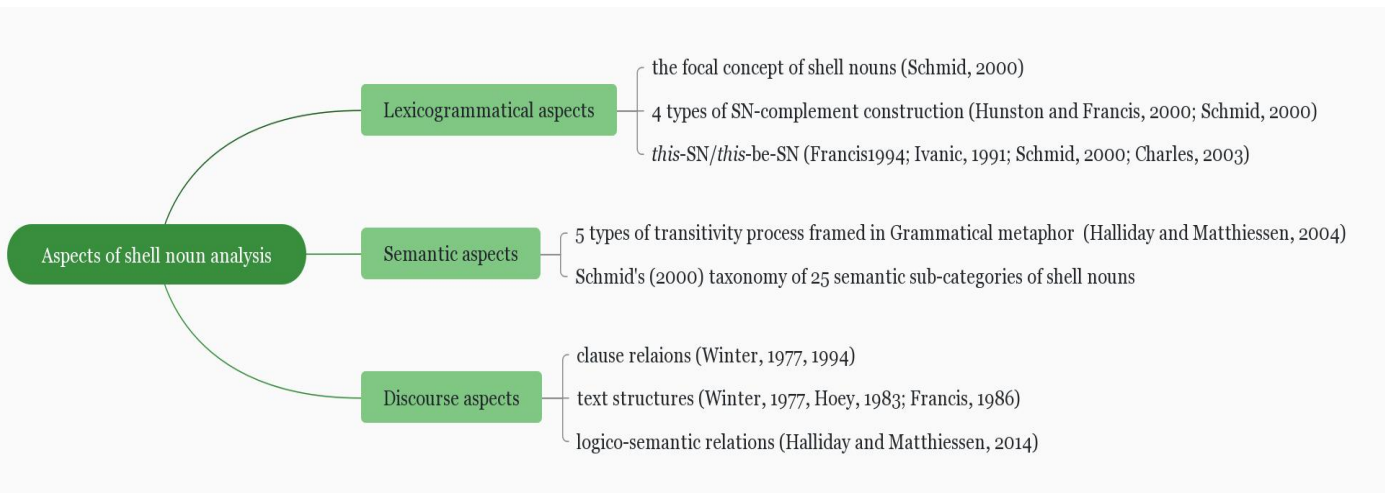


Figure 2.3: A summary of analytical framework

2.8. Outline of the thesis

This section outlines the structure of this thesis and how it will address the research questions. The subsequent chapters are organized as follows: Chapter 3 describes the corpus design and methodology used to address the research questions. Chapter 4 presents the quantitative analysis of the distributions of shell nouns, shell-noun syntactic patterns, and the types of different transitivity processes in order to answer RQ 1. Chapter 5 examines the lexicalizations of shell nouns to address RQ 2. The analysis of rhetorical features and differences in epistemological orientations will be addressed transversely across chapter 4 and 5 to answer the RQ 3. Finally, chapter 6 will present general conclusions, avenues for further research, and pedagogical implications for EAP/ESAP teaching and learning.

Chapter 3 Methodology

This chapter describes the methodology used to address the research questions outlined in the previous chapter. Section 3.1 first explains why a corpus-based approach was chosen for this thesis. Section 3.2 then details the decisions made regarding corpus selection, levels selection, text genre selection, and disciplinary domains selection. Next, in section 3.3, the data collection procedures, instruments used in the study, and the type of query, semantic taxonomy of shell nouns are outlined. section 3.4 explains the decision made regarding data cleaning and statistical measurements. finally, a summary of the methodology is presented in section 3.5.

3.1. A corpus-based approach to shell-noun uses

This section explains how research questions can be answered by triangulating the automated analysis of data inspired by a corpus-based approach with the exhaustive and context-sensitive manual inspection and interpretation of texts based upon their uses of shell nouns.

3.1.1. Corpus linguistics methodology

This thesis utilizes a corpus-linguistic methodology to analyze the uses of shell nouns in student disciplinary writing. A corpus is defined as ‘a collection of naturally occurring examples of language’ (Hunston, 2002a, p.2), which has been collected and stored electronically for linguistically motivated research and study. The use of corpus linguistics allows the researcher to access numerous authentic examples and a considerable amount of information, but it is the researcher who must interpret them (Hunston, 2002a). This thesis employs corpus linguistics for both automated quantitative and manually qualitative analyses. The use of automated pattern-driven investigation in large datasets retrieved by computer enables faster, more accurate, and more consistent processing of a large amount of data than manual analysis. As Barnbrook (1996, p.11-12) notes, this makes the use of a computer indispensable to this research, as it allows for data to be ‘manipulated, sorted and formatted without human bias for the purpose of the analysis’. While the empirical dimension of corpus linguistics strengthens the reliability of analytical results, this methodology treats text as product rather than an unfolding discourse as process. This is because ‘the

computer can only cope with the material products of what people do when they use language. It can only analyse the textual traces of the processes whereby meaning is achieved'. (Widdowson, 2000, p.4).

It is important to note that these statistically-motivated data in this study can only be viewed as a description of writing quality in terms of relevant linguistic features. The corpus linguistics methodology utilized in this study is also evident in the manual and context-sensitive analysis of every concordance line. This is because a close reading of individual corpus texts can provide more specific and important information about the ways in which certain shell nouns are used differently in each of the corpora. As noted by several linguists, including Carter (2004) and O'Keeffe et al. (2007), corpus linguistics methodology is more reliable when the automatic corpus analytic techniques are complemented with a more fine-grained qualitative investigation of the study of language. Such reliability has been demonstrated in previous studies, such as those by Schmid (2000), Hinkel (2003), Flowerdew (2015), and will be employed in this research.

Specifically, in this thesis, distributions of shell nouns and their syntactic patterns are mainly identified with quantitative statistical analysis, while the interpretation of shell-noun uses to perform the metadiscursive and disciplinary knowledge construal functions is conducted with manual qualitative analysis.

3.1.2. A corpus-based approach

In the field of corpus linguistics methodology, the main contention revolves

around the distinction between ‘corpus-driven’ and ‘corpus-based’ approaches (Flowerdew, 2006). The former is a fully a posteriori approach, in which ‘a theoretical statement can only be formulated in the presence of corpus evidence and is fully accountable to it’ (Tognini-Bonelli, 2001, p.11). Proponents of this approach view corpus linguistics as a theory that unveils previously unknown aspects of language (Flowerdew, 2006, p.174), and may challenge well-established theoretical positions (Tognini-Bonelli, 2001, p.48). By contrast, linguists in the corpus-based camp regard corpus linguistics as a methodology rather than a theory (Biber et al, 1998; Conrad, 2002; McEnery and Hardie, 2012). They perceive and analyse the corpus through models and descriptions of language and filter the data accordingly. Corpus evidence is seen as an extra bonus rather than a determining factor regarding the analysis. Researchers use the corpus primarily to expound, test, or exemplify theories and provide corroborations to the existing description of language, making ‘needed changes where corpus data does not fit’ (e.g. Tognini-Bonelli, 2001, p.66).

This present study is best described as corpus-based because it is deductive in nature and in line with the goals of corpus-based investigations. The aim is not simply to report quantitative findings, but to explore ‘the importance of these findings for learning about patterns of language use’ (Biber et al., 1998, p.5).

This study aims to explore the epistemological variations across and within the disciplinary boundaries reflected by written academic English. In applied linguistic studies, corpus-based methodology has been used to understand disciplinary differences because it offers an empirical dimension that strengthens available

interpretations and contrasts with other methods that can produce partial and prescriptive findings, such as impressionistic methods of text analysis. These methods do not check researchers' judgments, which can lead to the effects of whim and caprice. Corpus-based methodology differs from observation methods, such as keystroke logging analysis, which observe how writers compose texts by recording what they do (e.g. key-presses, duration of the key-strokes, pauses) on a computer or digital devices. Additionally, corpus-based methodology differs from elicitation methods (e.g. questionnaires and interviews) and introspection methods (e.g. think-aloud protocols) (Hyland, 2015, p.291), which rely on writers' and readers' self-reports to understand their thoughts and perspectives on how they use texts. These self-reports may not correspond with actual behaviour.

More importantly, a corpus-based approach to academic writing is particularly useful because it offers 'insights into disciplinary practices that help to explain the mechanisms by which knowledge is socially constructed through language' (Hyland, 2015, p.292). This approach provides solid support for understanding how the conformity of epistemological orientation characterizing a disciplinary domain is reflected in the ways in which disciplinary knowledge is constructed and communicated through the use of shell nouns in academic discourse.

3.2. The corpus design

This section explains the key decisions about the corpus design in this study.

3.2.1. Corpus selection: the BAWE corpus

After deciding to use corpus-based approach, three possible corpora were investigated to determine which would offer the most representative and balanced coverage of disciplinary student writing. These were the Louvain Corpus of Native English Essays (LOCNESS) (Granger, 1998), the Michigan Corpus of Upper-level Student Papers (MICUSP) (2009), and the corpus of British Academic Written English (BAWE) (Nesi and Gardner, 2012). Although LOCNESS is a well-known L1 student writing corpus with a total number of 264,095 words of university students' essays, its compilation is heavily based on American students' written English essays, containing 168,400 words (nearly 64% of the total university students essays). Therefore, the LOCNESS corpus was deemed unsuitable for the purpose of this study. The focus of this study is the distinctive ways in which disciplinary knowledge is constructed and communicated through the use of shell nouns, rather than the the American cultural background impacts on students' writing. Furthermore, the LOCNESS essays were not written for module assessment (the writers were asked to write an essay that they would not receive any marks for it), and it is difficult to know how motivated the writers were. The primary purpose of the LOCNESS corpus design, according to Alsop and Nesi (2009, p.72), is to 'monitor non-native-speaker errors and the processes of language acquisition, rather than the development of academic literacy skills and disciplinary knowledge acquisition'. Additionally, since LOCNESS was completed in the latter part of the 20th century, it may be less representative of student writing produced in the 21st century's classrooms.

In contrast to the somewhat outdated LOCNESS corpus, the Michigan Corpus of Upper-level Student Papers (MICUSP) was developed in the late 2009s. It consists of 829 A-graded papers written by University of Michigan students in their final undergraduate year or in their first three years of graduate education (Römer and Swales, 2010, p.249). These papers were collected from six disciplines across the four academic divisions of the university (ibid.). While the MICUSP corpus was designed to mirror the features of disciplinary student writing, it is not representative in terms of text type, as it is heavily skewed towards one type of paper: reports, which make up 44% of the corpus, with a total of 364 texts. Figure 3.1 illustrates the distribution (in percentages) of paper types in the MICUSP corpus (<https://elicorpora.info/main>):

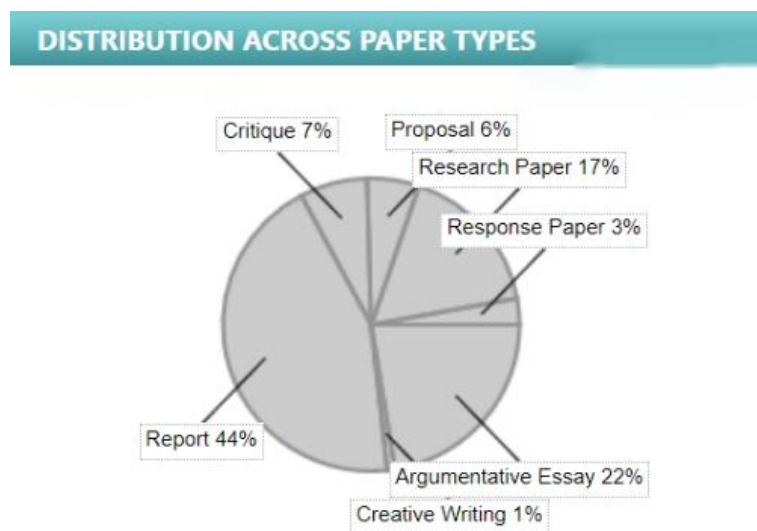


Figure 3.1: The distribution across paper types in MICUSP corpus

Furthermore, the majority of papers in the MICUSP corpus are written by senior undergraduate students (52%) and first to third year graduate students (42%), making

the corpus more suited to EAP practitioners, such as EAP writing instructors and graduate-level students, rather than potential international undergraduate students who are non-native speakers of English as an additional language.

In contrast, the BAWE corpus is a small but representative corpus of student-written English collected from four universities in England between 2004 and 2007. Unlike the LOCNESS corpus, which lacks clear information about student backgrounds, or the MICUSP corpus, which has an uneven distribution of genres, the BAWE corpus is designed to support the consistent identification and coherent description of students writing across disciplines, genres and academic levels (Alsop and Nesi, 2009, p.72). Additionally, the BAWE corpus's contextual annotation allows for the corroboration of findings from small qualitative studies, the triangulation of data, and the provision of strong quantitative insights into the use of grammar, lexis, and discourse patterns in disciplinary texts produced by student writers (e.g. Alsop and Nesi, 2009, p.75; Nesi et al., 2004, p.441). Considering these factors, the BAWE corpus was deemed the most consistent, balanced, and representative corpus for this study, and thus was chosen for analysis in this thesis.

The BAWE corpus is a collection of 2,761 high-quality assessed proficient student assignments, with marks equivalent to 'distinction' or 'merit' in the British system, corresponding respectively to American 'A' and 'B' grades in the U.S. system. This corpus comprises 6.5 million words and is evenly distributed across 35 disciplines from four broad discipline domains (arts and humanities, social sciences, life sciences and physical sciences) and four levels of study (first-, second-, final-year

of undergraduate and master). While the number of texts in each discipline and level of study may not be perfectly balanced, the BAWE corpus ‘provides a reasonable sampling across the discipline domains and the years’ (Thompson 2014, p.353). Figure 3.2 and Figure 3.3 present information about the disciplinary domains and levels of study in the BAWE corpus, respectively.

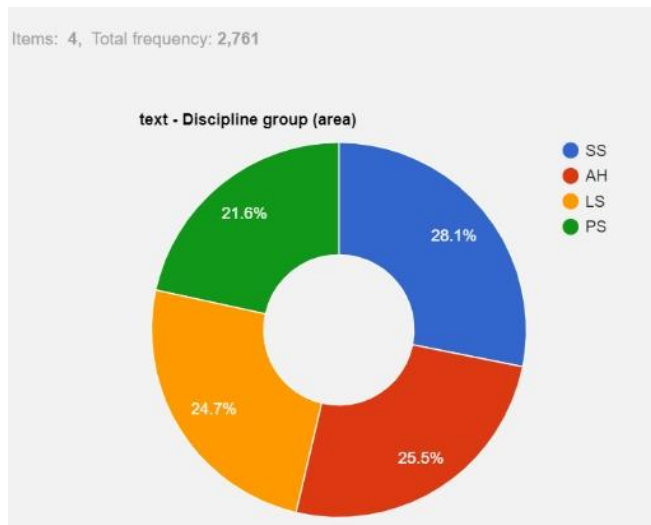


Figure 3.2: The distribution of disciplinary domains in the BAWE shown on the Sketch Engine online page (Kilgarriff et al., 2014)

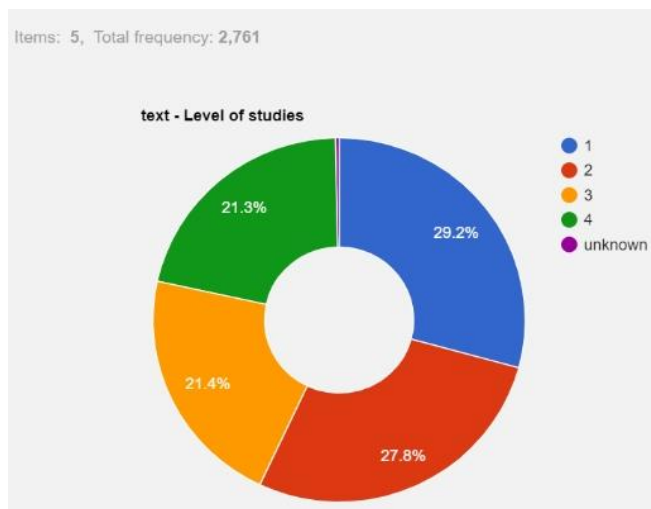


Figure 3.3: The distribution of levels of study in the BAWE shown on the Sketch Engine online page (Kilgarriff et al., 2014)

The corpus used in this thesis is divided into six sub-corpora consisting of academic essays from various disciplinary domains. The selection of texts for the corpus and its sub-corpora was guided by two main considerations. Firstly, a decision was made regarding the levels of study and text genres to be included in the study. Secondly, a decision was made concerning the disciplinary domains to be represented in the corpus.

3.2.2. Levels of study and text genre

The BAWE corpus consist of data that is evenly distributed across 35 disciplines from four broad disciplinary domains and four levels of study. The majority of the papers in the corpus were written by ‘native’ or ‘near-native’ speakers of English, with L1 English writers accounting for two thirds of the papers (1,948 papers in total). The remaining 804 papers were written by non-native speaker students, and there were nine papers of unknown origin. Table 3.1, presented below, illustrates the distribution of these non-native assignments across the levels of study and disciplinary domains:

Table 3-1

Table 3.1: The distribution of non-native texts in the BAWE corpus by level and disciplinary domains

Disciplinary domain	Level 1	Level 2	Level 3	Level 4	Total
Arts and Humanities	15	25	44	30	114
Life Sciences	24	31	32	101	188
Physical Sciences	43	31	45	47	166
Social Science	82	55	61	138	336
Total	164	142	182	316	804

It is clear that nearly 40% (316) of the non-native texts were contributed by students at level 4. This said, the majority of international students in British

universities tend to be international postgraduate students who might be post-experienced or department transferred (e.g. moving into MBA from a degree of finance or politics) and these were excluded from the study. That is, the texts produced by level 4 students were discarded. This is because that, once again, the purpose of this study is to explore the disciplinary specificity reflected in the ways in which knowledge is constructed and communicated through the use of shell nouns in L1 student writing, and so the consistency of writers disciplinary backgrounds is an important factor in the consideration of the selection of the text data. In addition, the aim of this study is not so much to track the developmental trajectories in student writing, as to provide insightful understanding of the broad soft vs. hard disciplinary divide and unveil significant variations and differences within each disciplinary domain, so that material and syllabus can reflect such differences and EAP/ESAP practitioners can facilitate L2 students' academic writing efficiently. Thus, first-year and third-year texts might give a more marked degree of comparison and the texts produced by second-year students were excluded. Taking these aforementioned factors into consideration, it was, thus, decided to use the texts produced by level 1 and level 3 students for the study in this thesis.

To examine the impact and the interaction only between disciplinary domains, the imbalanced distribution of genre across discipline groups needs to be taken into consideration. For instance, there are 128 case studies in NS but 0 in AH, and some less extreme but more typical distributions can also be found, as shown in table 3.2 below:

Table 3-2

Table 3.2: Distribution of genres by discipline groups (Heuboeck et al., 2010, p.7)					
	Arts and Humanities	Life Sciences	Physical Sciences	Social Sciences	Total
case study	0	91	37	66	194
critique	48	84	76	114	322
design specification	1	2	87	3	93
empathy writing	4	19	9	3	35
essay	602	127	65	444	1238
exercise	14	33	49	18	114
explanation	9	117	65	23	214
literature review	7	14	4	10	35
methodology recount	18	158	170	16	362
narrative recount	10	25	21	19	75
problem question	0	2	6	32	40
proposal	2	26	19	29	76
research report	9	22	16	14	61
Total	724	720	624	791	2859

If we look back to the table 3.2, essays are found in all 24 disciplines across the four disciplinary domains represented by 50 or more assignments in the corpus. This fact endorses not just Warburton's (2020, p.7) view that 'whatever you study, at some point you will be asked to write an essay' but also Thoreau's (2013, p.29) claim that 'if you know about two essay genres, you will have a good foundation for tertiary study'. More importantly, bearing in mind the purpose of this study is to investigate disciplinary variation, with a particular focus on the text genre of the essay. Unlike other text genres, such as exercise, which aim to provide practice in key skills or develop an understanding of professional practice, the essay genre serves the purpose of not only developing students' ability to produce coherent individual arguments but also demonstrating their understanding of shared disciplinary knowledge and critical thinking skills. The ways in which new knowledge is constructed and communicated are reflected in the essay genre, making it a crucial component of developing a

comprehensive understanding of a particular discipline (see, e.g. Nesi and Gardner, 2012, pp.37-39). given the factors mentioned above, it was decided to use essays as focus of this thesis. It is possible to criticize this decision on the grounds that a restricted genre such as essays cannot lead to a conclusion that is applicable to the uses of shell nouns in other types of student writing. while this criticism is reasonable, it is important to note that essays offer the most even distribution and representativeness of disciplinary specificity among the genres of text that were tested. Therefore, essays are better suited than any other text genres in the BAWE corpus for studying the research questions outlined in this thesis.

3.2.3. Disciplinary domains: soft vs hard

As regards the disciplines involved, decisions were made according to the type of epistemological properties that the disciplines involved might show. Thus, it was necessary to consider several disciplinary taxonomies to describe the nature of each disciplines and, at the same time, to explore different epistemological ways in which these disciplines might overlap when conceptualizing new knowledge.

Based on Biglan's (1973) model, Becher and Trowler (2001) use the metaphor of tribes and territories to describe academic disciplines. The tribal aspect involves the ways in which academics engage with their subject matter while the territorial part associates with the formation of disciplinary cultures (Becher and Trowler, 2001, p.23). It is noted that the former (disciplinary tribes) is an important factor in the formulation of the latter (disciplinary territories). Accordingly, four types of

disciplinary domain can be identified by a system in which both the epistemological properties of knowledge fields and social characteristics of research group were taken into consideration. Specifically, they are ‘hard-pure’ (e.g. physics, chemistry), ‘soft-pure (history, anthropology)’, ‘hard-applied (medicine, engineering)’ and ‘soft-applied’ (education, law) (Becher and Trowler, 2001, p.35).

It is important to acknowledge that the disciplinary boundaries ‘are constantly shifting and are sometimes poorly demarcated’, and that there are ‘numerous gaps and overlaps in their patterns of coverage of knowledge domains’ (Becher and Trowler, 2001, p.64). Nevertheless, this system has identified features that characterize the four disciplinary domains, as summarized in Table 3.3 below:

Table 3-3

Table 3.3: Features of the four disciplinary domain (Becher and Trowler, 2001, p.36)				
Disciplinary domain	Disciplinary concern	Knowledge construction	Research results	Characteristics
Hard-pure	Universals and quantities	Cumulative; atomistic (crystalline/tree-like)	Discovery and explanation	Impersonal and value-free
Soft-pure	Particular cases	Reiterative; holistic (organic/river-like)	Interpretation and understanding	Personal and value-laden
Hard-applied	Pragmatic know-how and the proficiency of the functional practice	Mastery of the physical environment; applies heuristic approaches	Products and techniques	Purposive and functional
Soft-applied	Utilitarian know-how and the enhancement of the professional practice	Applies case studies	Protocols and procedures	Purposive and functional

While Becher and Trowler’s metaphor for describing academic disciplines has been influential and widely used, it is important to recognize the complex and contradictory nature of disciplines. As Trowler (2013, p.5) points out, ‘discipline does

not have essential characteristics which are all necessarily present in every instance and each individual discipline has no essential core characteristic either'. The study of the academic disciplines, therefore, needs to be situated within a broader context that appreciates the heterogeneity and dynamism within disciplines.

Critics (e.g. Martin, 2011; Manathunga and Brew, 2012) have argued that moving beyond land-based metaphors such as territories towards watery metaphors such as oceans can illuminate the connotation and fluidity of knowledge domains. Furthermore, the tribes-and-territories metaphor can imply an imperialistic connotation, leading Kagan (2009) to propose referring to academic communities as three cultures: the culture of humanities, the culture of natural sciences, and the culture of social sciences.

Despite the limitations of Becher and Trowler's (2001) taxonomy, which was proposed over two decades years ago, it remains a valid approach for exploring disciplinary differences from a linguistic perspective, especially in the field of applied linguistics. For example, Thompson and Hunston's (2019, p.22) corpus-linguistic observations on the 'preferred ways' of talking about disciplines and disciplinarity, as well as the chief ways in which people conceive of disciplines, demonstrate that '*discipline* predominantly indicates a grouping, a training, or a way of building knowledge that either has attributes or is developed'.

Although some scholars firmly believe that any attempt to define scientism is bound to fail, as the methods and approaches used in one discipline (e.g. astronomy) often bear little resemblance to those used in other disciplines (e.g. computer science)

(e.g. Bauer, 1992), this argument overlooks the possibility that certain higher-order epistemic commonalities cut across cognate disciplines. In this respect, Thompson and Hunston's (2019) observations further support Becher and Trowler's (2001) view on the overarching epistemological properties of cognate disciplines, namely soft-science and hard-science fields. Therefore, the disciplinary taxonomies proposed by Becher and Trowler (2001) can be useful for investigating disciplinary variations in this thesis.

Moreover, since the aim of this work is to compare hard and soft disciplinary domains and explore the variations within the broad sense of academic field or branch of knowledge, the focus has been placed on exploring the differences between those disciplinary domains rather than the variations across specific disciplines. For this purpose, this study adopts the classification of four disciplinary domains identified by Nesi and Gardner (2012): Arts and Humanities; Life Sciences; Physical Sciences and Social Sciences. To ensure research efficiency and the clarity of comparison, the study groups texts from Life and Physical Sciences together to represent a hard knowledge field: Natural Sciences (henceforth NS), as has been done in other recent studies (e.g. Biber et al., 2016). Table 3.4 below summarizes the information on the corpora used in this study.

Table 3-4

Table 3.4: Corpora for this thesis

Genre	Disciplinary domain	Level	Number of texts	Number of words	Average length of text
Essay	Arts and Humanities (AH)	1	198	413,663	2,089
		3	102	331,745	3,252
	Subtotal	300	745,408	2,484	
	Social Science (SS)	1	77	157,065	2,039
		3	70	229,817	3283

	Subtotal	147	386,882	2,632
Natural Sciences	1	55	107,705	1,958
(NS)	3	34	84,356	2,481
	Subtotal	89	192,061	2158
Total		536	1,324,351	2,471

3.2.4. Corpus size and type: small and specialized

Based on Table 3.4, the size of the three disciplinary corpora in this thesis is considered small, as any written corpus under five million words can be described as small (Flowerdew, 2004). Additionally, Koester (2010) suggests that there are many small corpora much smaller than that size, so a commonly agreed-upon threshold for small corpora is up to 250,000 words (Flowerdew, 2004). Therefore, the majority of the sub-corpora can also be classified as small corpora.

There are two reasons why this thesis uses the term *specialized* to describe the corpora used in this study. Firstly, it is related to the investigation of ‘a particular type of language’ (Hunston, 2002a, p.14), which in this case is the investigation of disciplinary written language. Secondly, it relates to the specific type of texts used, which is ‘student academic writing’ in this case. Investigating specialized corpora may provide the researcher with more representative linguistic features of a particular background, making it a suitable choice for this study.

The use of small specialized corpora in this study is guided by a specific set of goals that differ from those of most recent work on shell nouns (e.g. Hunston and Francis, 2000; Schmid, 2000; Benitez-Castro, 2014). In their book, which mainly focuses on pattern grammar, Hunston and Francis (2000) aim to emphasize how word patterns can reveal shared semantic features of words and thus help to narrow down

the scope of unrecognized word classes. Shell nouns are only briefly mentioned as a part of a larger argument about the lexico-grammar of English. Schmid (2000) deals with the description of a particular word class and the coherent taxonomy of them based on evidence from corpora. Benitez-Castro's (2014) goal is to describe what is counted as shell nouns, revolving around nine variables. In both of the previous cases (Hunston and Francis, 2000; Schmid, 2000), corpora play an important role in pushing towards narrower definitions of what is counted as an open-class shell noun item: both studies make use of relatively large corpora due to the precedence given to the formalized query procedure. Although Benitez-Castro's (2014) corpus is smaller, he is less interested in making use of specialized corpora and more interested in investigating English general corpora with an even distribution of genres and modes.

The smaller in size and more specialized size of the corpus in this study allows for a more intensive, discourse-oriented approach to investigating how shell nouns are used in constructing and communicating disciplinary knowledge. This is because a small and specialized corpus 'allows a much closer link between the corpora and the contexts in which the texts in the corpus were produced' (Koester, 2010, p.67).

After defining the major scope, size and type of the corpora, attention is now turned to the collection and analysis of the data. Before discussing the analysis of the data in the next section, it is necessary to point out some limitations here at the outset. The first limitation concerns the different size of the three corpora, including the text numbers and average length. Another limitation is the diverse modules within each of the three corpora. It is acknowledged that differences in modules may have important

impacts on the shell-nouns, particularly in terms of which nouns occur with what frequency or in what syntactic patterns they appear. However, since it is impossible to control these limitations simultaneously, such factors will be taken into consideration when interpreting the data.

3.3. Data collection

This section explains the choice of data analysis software and explains how it is used for the data collection for the purposes of this study.

3.3.1 Data analysis software

Having defined the major scope of the corpora, attention is now turned to the selection of the text analysis software for the three research questions of this study. Sketch Engine software is chosen for two reasons: 1). the corpora were part-of-speech tagged, reaching up to ‘a satisfactory accuracy of 98% and the mistakes are typically only limited to phenomena of less interest such as misspelt words, rare usage or interjections’ (Kilgarriff et al., 2014). 2). Sketch Engine software is highly flexible in changing criteria for the sub-corpora. The screenshot below shows the criteria options for the building of different sub-corpora.

The screenshot displays the 'Filter context' interface in Sketch Engine. At the top, there is a 'Filter context' dropdown menu. Below it, the 'Text types (4)' section is expanded, showing a grid of filter options. Each option is in a white box with a red 'GO' button at the bottom. The options are: Grade, Discipline, Course title, Level of studies, Text genre, Module title, Discipline group (area), and First language of author. There are also 'expand all' and 'collapse all' links in the top right corner of the filter area.

Figure 3.4: The criteria options in the Sketch Engine online page

The following section (see section 3.3.2) explains the functions of Sketch Engine and how it is applied for the data collection in this research.

3.3.2. The type of query

The present study addresses the need to supplement previous studies about how shell nouns and their syntactic constructions act as a lexico-grammatical unified whole to construe disciplinary knowledge. Therefore, the Corpus Query Language (henceforth CQL) is utilized to examine the frequency and diversity of shell nouns across disciplines. According to Kilgarriff et al. (2014, p.12), CQL sets conditions for tokens (words) and thus captures the occurrences of the uses of shell nouns. In Table 3.5 below, the specific CQL for each shell-noun grammatical pattern is presented.

Table 3-5

Table 3.5: CQL for each shell-noun syntactic pattern

Function	Pattern	Sub-pattern	CQL
Cataphoric	SN+cl	SN-cl (to infinitive)	[tag="N.*"][word="to"&tag="T.*"]
		SN-cl (that-clause)	[tag="N.*"][word="that"&tag="C.*"]
	SN+be+cl	SN-be-cl (that-clause)	[tag="N.*"][lemma="be"][word="that"&tag="C.*"]
		SN-be-cl (to infinitive)	[tag="N.*"][lemma="be"][word="to"&tag="T.*"]
Anaphoric	Th-SN		[lemma="this"][tag="N.*"]
	th-be-SN		[lemma="this"][lemma="be"&tag="V.*"][tag="N.*"] [lemma="this"][lemma="be"&tag="V.*"][tag="AT.*"][tag="N.*"]

It is to be noted that despite the strengths of the corpus-based approach presented here, this thesis retains the convention of labelling the lemma with the singular

(unmarked) form of shell nouns. This is because ‘the singular form is much more common’ in academic writing (Flowerdew and Forest, 2015, p.87), from those that occur only in the singular (e.g. *evidence, basis*) to those that are simply much more often singular than plural (e.g. *view, intention*). Needless to say, the inclusion of the analysis of the plural (marked) form of shell nouns could have provided more well-reasoned interpretations and evidence for the conclusions made in this thesis. Unfortunately, those shell nouns that occur in the plural form were not taken into consideration mainly due to the nature of the data at hand, and methodological restrictions.

3.3.3. Semantic taxonomy of shell nouns

The decision to use Schmid’s (2000) semantic taxonomy is grounded on the fact that, as Benitez-Castro (2014, p.119) affirms, ‘it is the most thorough semantic classification of shell nouns to date’. The semantic classification in Schmid (2000) distinguished six major classes and a total of twenty five sub-classes. Based on the description of the type of human experience characterized by each sub-class, Table 3.6 below summarizes these types and the sub-class of shell nouns.

Table 3-6

Table 3.6: Transitivity process types construed by Shell nouns mainly after Schmid’s (2000) semantic categorization.					
Semantic classification in Schmid (2000)	Sub-class	Type of transitivity process construed	Human experience related	Shell noun example	
Factual	Neutral	Identifying relational process	Abstract state of affairs and neutral facts.	<i>fact, thing</i>	
	Partitive		Aspects, features of state of affairs and facts.	<i>example, feature,</i>	

	Comparative	Attributive relational process	Comparison of state of affairs and facts.	<i>difference, similarity</i>
	Attitudinal		Attitudes toward state of affairs and facts.	<i>problem, advantage,</i>
	Causal			
Mental	Evidential	Mental process	An observable state of affairs or fact is conceived of as a reason for mental state.	<i>evidence, sign</i>
	Conceptual		Shelling events and abstract relations as ideas	<i>idea, theory</i>
	Creditive		Psychological states of abstract relations (rather than events)	<i>belief, assumption</i>
	Dubitative		Psychological state of doubting.	<i>doubt, question</i>
	Emotive Volitional		Psychological state of emotion. Psychological state of wanting that future events, especially future activities take place.	<i>fear, regret wish, aim, solution</i>
Modal	Epistemic	Attributive relational process	judgments about the possibility, probability or certainty that something is or is not the case.	<i>possibility, chance</i>
	Deontic		The permission, needs and obligation of acts performed by morally responsible agents resulting from the propositional content.	<i>right, freedom</i>
	Dynamic		What can possibly, probably or necessarily happen under certain circumstances.	<i>ability, power</i>
Linguistic	Propositional	Verbal process	The propositional contents of utterances with a simultaneous characterization from a metalinguistic categories, mainly as text types of some sort.	<i>news, message</i>
	Assertive		Illocutionary speech acts of stating states of affairs and other abstract relations or contents of acts of reporting events	<i>statement, declaration</i>
	Rogative		Illocutionary speech acts of question	<i>question, query</i>
	Directive		Illocutionary speech acts of getting the original hearer (including the speaker) to do something	<i>order, command</i>
	Commissive		Illocutionary speech act of committing oneself to a future activity.	<i>promise, offer</i>
	Expressive		Illocutionary speech act of stating states of affairs or events that has caused feelings.	<i>complaint, grievance</i>
Eventive	General	Material process	Event with a duration in the physical world	<i>event, change</i>
	Specific		Event with a intended purpose,	<i>attempt, option</i>

		goal or aim; habitual event; comparative event.
	Attitudinal	Attitudes towards certain <i>trouble,</i> events <i>problem</i>
Circumstantial	General Specific	

It should be noted that the sub-class of causal and the class of circumstantial shell nouns do not construe any type of transitivity process. This is mainly because these shell nouns construe logical semantic relations rather than transitivity process by encapsulating information that is ‘typically expressed by adverbial phrases’ (Schmid, 2000, p.276). Additionally, these shell nouns present peripheral information about additional non-focal human experience, such as setting, time or manner. As a result, shell nouns that draw on these sub-classes are excluded from this study.

Although Schmid’s (2000) fine-grained semantic taxonomy of shell-noun meanings is thorough and comprehensive, which contrasts to the ‘all-embracing nature of other classifications’ (Benitez-Castro, 2014, p.259) provided by previous studies on shell nouns (e.g. Francis, 1986; Hunston and Francis, 2000), it should be pointed out that it is an important but not solely reliable measure of the precise categorization of shell nouns. To fit existing frameworks to the nature of data at hand, some modifications are required in this study.

Although the semantic categorization tags adopted in this study are in the main Schmid’s (2000), the approach adopted in this contextualized analysis of data required two types of decisions to be made: 1). shell nouns which were left out in Schmid’s (2000) list had to be accommodated in existing or modified categories (see examples

(3.1) and (3.2) below). 2). Shell nouns included in Schmid's (2000) list but their semantic meanings were partially covered were accommodated in existing or modified categories (as in (3.4)).

(3.1) While there is a **sense** that the debate has been bogged down by personal animosities and the lack of fresh interpretations, it is still of interest today... (AH, 0019d)

(3.2) While the need to substantiate his social critiques and ideas with biblical references is perhaps traditional, in the **sense** that he acknowledged the authority of the Bible and used it as a legitimising force for his ideas. (AH, 0019e)

Specifically, both uses of the shell noun *sense* in examples (3.1) and (3.2) suggest a creditive mental reading. That is, writers include 'a characterization of the psychological state which they want to attribute to EXPERIENCER in their shells for IDEAS' and 'the subject of these psychological states are abstract relations' (Schmid, 2000, p.195). It is worth mentioning here that this mental creditive tag is moderately modified by the additional 'feeling' in example (3.1) and 'understanding' in example (3.2), according to their semantic affinity with nouns like *belief* and *comprehension* respectively.

The shell noun *failure* in Schmid (2000, pp.252-254) is only categorized as a dynamic modal use as such uses are often pre-modified by possessive determiners, genitive phrases or post-modified by *of*-prepositional phrases to indicate their subject-oriented attributes (as in (3.3)).

(3.3) it is my opinion that his **failure** to differentiate between 'Ranter' myth and 'Ranters' has been a stumbling block in this debate. (AH, 0019d)

- (3.4) **Failure** to ensure full economic recovery, along with badly planned New Deal policies, also meant that American businesses fared poorly. (AH, 0302a)

However, the focus in example (3.4) is not on the subject-oriented attribute of dynamic modality, but on the ‘unsuccessful attempt’, which is eventive. The semantic affinity of *failure* in example (3.4) is with nouns like *catastrophe* and *debacle*. Therefore, the code of this shell noun, *failure* is ‘eventive attitudinal’. It is important to note that in the semantic categorization of shell noun instances, only by close reading of the co-text surrounding discourse segments, assisted by dictionary and corpus evidence, can one identify unaccounted senses of shell nouns that are overlooked in Schmid’s (2000) automated corpus study. In other words, it is worth emphasizing the importance of close reading of the extended context to provide reliable evidence and solid findings for the semantic categorization of shell-noun instances.

3.3.4. Procedures for data collection

In this comparative study, the data collection falls into three steps according to the two variables: disciplinary domains and levels of study. At the Sketch Engine online page, from top to bottom the following procedures are followed: 1). Set the searching criteria in the text types. They are: Level of studies (level 1 and level 3), each of the three disciplinary domains (Arts and Humanities, Life Sciences and Social Sciences), Text genre (essay), First language of author (English). A screenshot of the searching criteria is given below in Figure 3.5:

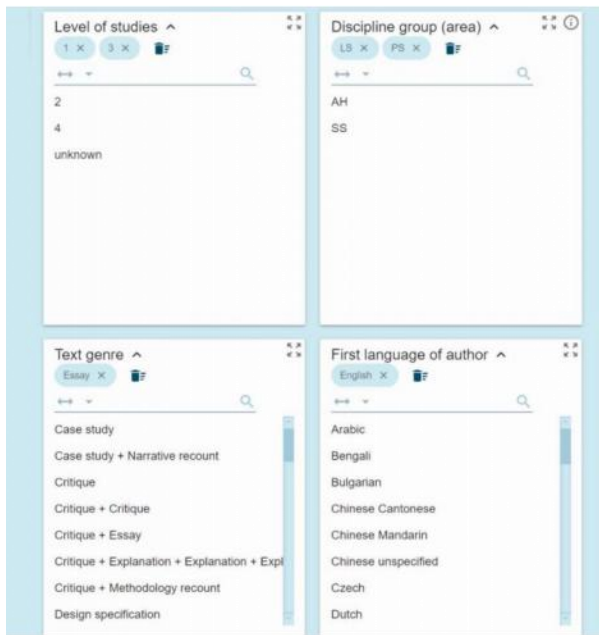


Figure 3.5: The searching criteria in the Sketch Engine online page

2) As noted earlier (see section 2.4 and 3.3.2), this study takes account of six shell-noun syntactic patterns in each of the three disciplinary domains: *SN-that*, *SN-to*, *SN-be-that*, *SN-be-to*; *th-SN*, *th-be-SN*. On the Sketch Engine online page, the CQL (see tTble 3.5 in section 3.3.2) was entered in the Advanced Query Type to retrieve occurrences of all potential shell nouns in each of the six syntactic patterns. Figure 3.6 below shows the CQL query for the syntactic pattern *SN-that*:

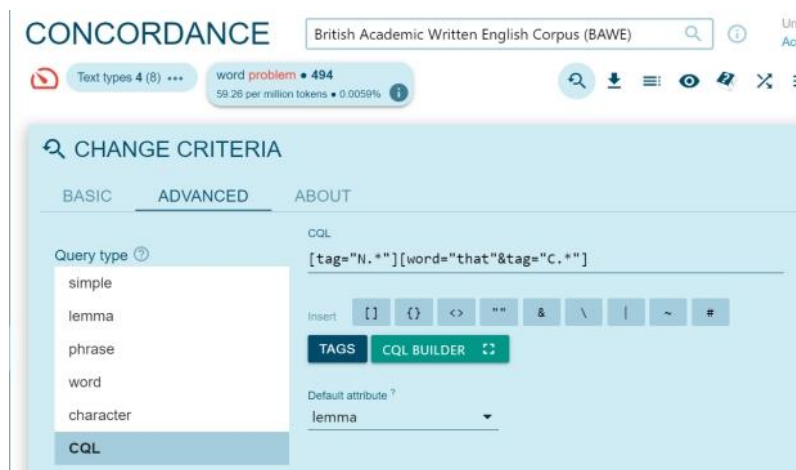


Figure 3.6: An example of the CQL for the query of shell-noun syntactic patterns

3) One primary research tool used for data collection in this study is the Key Word In Context (KWIC) feature of Sketch Engine. KWIC provides all occurrences of potential shell nouns in concordance lines. In this thesis, non-shell noun uses are sorted out manually examining their contexts (see section 3.3.1). Figure 3.7 below shows the surrounding co-text of the node noun (e.g. *fact*).



Figure 3.7: Concordance lines of shell noun *fact*

4). Concordance lines are retrieved to identify shell noun items and also their syntactic patterns. These lines are copied from Sketch Engine's online page to Microsoft Excel spreadsheets. This thesis takes two variables, disciplinary domains and levels of study, and these two variables are displayed separately in each column of the Microsoft Excel spreadsheet. There are six columns in total, from left to right:

- i). the sequence number for each concordance line, ii). the second column specifies the shell noun items and is accompanied by additional text needed for further interpretation of the shell noun. iii). followed by two columns that specify the disciplinary domain and levels of study respectively. v). and the last three columns

specify the corresponding semantic tags and the types of transitivity process of shell nouns

After these procedures, a list of shell-noun uses in each of the six syntactic patterns in the three corpora is generated. The next step is to calculate the raw frequency rates of shell nouns and their syntactic patterns. Table 3.7 below is an example of a fully coded shell noun item:

Table 3-7

Table 3.7: A complete coding of a shell noun in the syntactic pattern of <i>SN-be-that</i>					
No.	Concordance line	Disciplinary domain	Levels of study	Semantic tags	Transitivity process type
38	However, one of the objections to the Ontological argument was that <u>'exist' is not a predicate.</u>	AH (Arts and Humanities)	L1 (first-year)	Lin-as-argu (Linguistic-assertiv e-argumentative)	V (verbal process)

3.4. Data analysis

This section summarizes certain types of non shell-noun uses, providing an explanation that some concordances returned by the corpus not necessarily represent shell noun uses. Besides, the choice of statistical measurements is explained.

3.4.1. Data cleaning: types of excluded items

Since there is no guarantee that concordances returned by the corpus accurately reflect shell-noun uses, it is necessary to manually examine and analyze the context of every concordance line to determine whether or not to include it as an instance of a shell-noun use. Corpus studies have shown that moving beyond the limited scope of

concordance lines to the extended context is helpful and necessary for ensuring that ‘intensive textual analysis and analytical decisions make sense’ (Mahlberg, 2005, p.58). By clicking on the ‘Key Word in Context’ (KWIC) feature in Sketch Engine, researchers can access to the extended context for any concordance (see section 3.3.4). The data was cleaned by eliminating unwanted hits produced mainly by twelve types of interference (Non-shell instances):

i). One is when the word that does not function as a complementizer. This kind of interference applied to some thorny cases in the query of potential shell nouns in the syntactic pattern of **SN-that**, such as *problem, fact, etc.*, This is shown in the examples (3.5) and (3.6) retrieved from SS sub-corpora below:

(3.5) On the other hand, official statistics can be used as ammunition as numbers that will draw attention to a problem that needs addressing. (SS, 0140b, L1)

(3.6) The Amnesty Committee's approach to the notion of 'political motive' as applied in Amy Biehl's case often seems to have the effect of protecting perpetrators of gross human rights abuse from criminal and civil liability, a fact that is, admittedly, often difficult to reconcile oneself with. (SS, 0186, L3)

ii). The other type of interference occurred when the node word is a nominal phrase, as shown in examples below:

(3.7) The primary **aim** of the study was to deepen the understanding of fish-whale interactions - the strongest advocates for a resumed hunt are fisherman concerned that whales are taking too many fish. (NS, 6035d, L3)

(3.8) The primary **reason** why the English did not know how to cultivate the land was that they were unwilling to ask the Native Americans for assistance. (AH, 0029m, L3)

In examples (3.7) and (3.8), *study* and *land* before *to*-infinitive clause and *be-that* clause respectively appear to qualify as shell nouns, but, in fact it is *aim* and *reason* can be recognizable as shell nouns as they are semantically equivalent to the content coded by the *to*-infinitive clause and *be-that* clause.

iii). The third type of excluded item is a potential shell noun acting as premodifier of another head noun. It is not the query lemma, but the following noun that encapsulates preceding or subsequent discourse segments. This is typically apparent with anaphoric shell-noun uses, when the query lemma acts as a restriction of the following noun's meaning. This is shown in example (3.9) below, in which the returned query is in boldface.

(3.9) This illustrates that in most cases a majority rule social decision scheme can be implemented, but it needs to have a sub scheme allowing for this **leniency** bias. (NS, 0014e, L3)

As shown in example (3.9), nominal premodifier *leniency* is similar to an adjective without any reference, while the nominal prepositional complement (e.g. *this leniency bias*) is the referring expression, and thereby only the noun *bias* is viewed as shell noun.

iv). In regard to *it*-cleft sentences and existential *there* constructions, nouns in *it*-cleft sentences are not counted as shell nouns. This is because the clause in *it*-cleft sentence represents the 'notional subjects of the clauses' (Schmid, 2000, p.24) rather than the shell content. Manual verification was necessary in all these cases (as in (3.10)) and (3.11)).

(3.10) ...it is perhaps a good **idea** to weigh the arguments for and against such a political system.
(SS, 0140d)

(3.11) ...it is my **opinion** that his failure to differentiate between 'Ranter' myth and 'Ranters' has been a stumbling block in this debate. (A,H, 0019d, L3)

If we look back to (3.11), for example, the content of the clause *his failure to differentiate between 'Ranter' myth and 'Ranters' has been a stumbling block in this debate* is in fact a post-nominal appositive clause or what Schmid (2000, p.24) terms as an 'extraposition of a clause subject'. However, the noun in an existential *there + verb 'be' + noun + that* is considered as an instance of shell-noun use, as the noun in this construction forms a weak conceptualization boundary between the noun and the referent, as shown in example (3.12) below:

(3.12) However, there is **evidence** that autistic children do suffer certain lexical deficits through their lexis often seeming immature due to their use of neologisms and errors in the production of words and phrases seeming to mirror that of younger children. (AH, 6206d, L3)

It can be seen from example (3.12) that the shell noun *evidence* and the shell content in the *that*-clause together form a post nominal appositive relation, thus *evidence* is considered as a shell-noun instance.

v). The fifth type is the potential noun is a concrete, visible and tangible first-order entity. This is apparent with the syntactic pattern *th-SN*, where the meaning of **SN** is that of first-order concrete object. Examples (3.13) and (3.14) below represent a first-order non-shell use.

(3.13) The Gods of the earth and sea Sought thro' Nature to find this **Tree** , But their...(AH, 0013c, L1)

(3.14) Yet across the gulf of space, minds that are to our minds immeasurably superior to ours, regarded this **earth** with envious eyes, and slowly and surely drew their plans against us. (AH, 0024h, L3)

vi). When the potential noun is a first-order concrete entity, a shell noun use does not apply. This is shown in examples (3.15) and (3.16) below:

(3.15) The more recent Right Wing Authoritarianism (RWA) scale (Altemeyer, 1996) claims that it has fixed the problem of acquiescence. This **scale** uses cognitive theory to account for authoritarianism instead of the psychodynamic approach that is used in the f-scale. (NS, 0020b, L1)

(3.16) Even when the company ceased trading in 1834, this dividend still had to be paid. The Indians obviously could not afford this **tax** so a loan was created. (AH, 0030a, L1)

vii). When a potential noun refers to a first-order entity that is primarily associated with the measurement of something (e.g. the length of time), it is excluded from the category of shell-noun uses. Two specific nouns that deserve special mention are *time* and *period*. According to Schmid (2000, p.276), ‘circumstantial shell nouns do not shell circumstances but events’. Therefore, the noun *time*, which refers to ‘the thing that is measured in minutes, hours, days, years etc. Using clocks’ (LDCE), is not considered a shell-noun use (as in (3.17) and (3.18)) unless it represents ‘an occasion when something happens or someone does something’ (LDCE). However, such instances are rare in this study.

- (3.17) One example of this is servants in the Victorian era. Many statistics from this **time** were produced by those in positions of power, for example mistresses... (SS, 0001b, L1)
- (3.18) However post-1730's their role changed, and they became increasingly involved in routine births, especially with the upper and middle classes. What was seen as 'a problem birth' was redefined, for example prior to this **time** breech births were regarded as normal...(SS, 0001c, L1)

With regard to *Period*, it would be discarded when the meaning is that of 'a particular length of time with a beginning and an end' (LDCE). The meaning of *period* in example (3.19) is close to that of *time* in examples (3.17) and (3.18), in this case, it encapsulates a quantifiable amount of time. Therefore, it is a first-order non-shell use, which is excluded from the analysis.

- (3.19) Between 1620 and 1642 about 25,000 people migrated to New England, and by 1640 the colony had a total population of 40,000. After this **period**, though, migration was not an ongoing phenomenon...(AH, 0029I, L3)

viii). When the potential noun is a second-order entity just acting as a support in discourse context, it is excluded from the analysis of shell-noun use. Based on the argument that shell-noun uses of abstract nouns have context-dependent meaning, instance like (3.20) below is thus discarded.

- (3.20) Villages expanded into towns and cities of ten or twenty thousand inhabitants. This **overpopulation** led to poor living standards, over-extended planting of wheat and a fall of productivity as the goodness was drained from the soil. (AH, 0013a, L1)

The meaning of *overpopulation* in (3.20) is that of 'there are too many people living in a particular place' (LDCE). In this respect, it is possible to eliminate the

potential noun *overpopulation* without any meaning changes (e.g. *This lead to poor living standards...*). In other words, the second-order noun *overpopulation* here just acts as a support in this discourse context, which enhances the speed of processing the preceding single large noun-based discourse entity. Accordingly, it is excluded from the shell-noun uses, as it gives rise to a fixed term use rather than shell use.

ix). The ninth type is the case where the node word is an unwanted wrong item. This is because, when considered in context, the automated CQL query may return potential shell nouns that reveal themselves as unwanted wrong lexical items. In terms of the cataphoric shell-noun syntactic pattern such as **SN-to-infinitival** clause, the automated query may return a node word followed by adjectives, which is by no means a shell-noun use, such as example (3.21). With regard to anaphoric shell-noun syntactic pattern such as *th-SN*, demonstrative *this* functions as pronouns rather than determiners followed by unwanted wrong noun items (e.g. names). The excluded instances are evident in example (3.22) below:

(3.21) Johnston and Ramsted (1983) only found seriation **task** to be deficit with the SLI children involved performing approximately two years below their chronological age (AH, 6206e, L3)

(3.22) In doing this **Huntley** and Costanzo are surely promoting a Bayesian method of processing where evaluation of the evidence is updated with each new piece of evidence. (NS, 0014e, L3)

In example (3.22), although the assumed shell-noun use (i.e. *...this Huntley...*) meets the syntactic construction of *th-SN*, *this* functions as a pronoun rather than a determiner and *Huntley* is just a part of a naming expression that is not referential.

x). Sometimes, the occurrence of a determiner *this* is no guarantee of shell-noun uses, a shell analysis excludes the use of the noun that requires no reliance on a segment larger than a clause for its interpretation. This is because a crucial distinction between shell nouns and other abstract nouns is that the recovery of meaning of a potential shell noun requires a segment larger than a clause. On the other hand, other abstract nouns can often be associated with a specific reference to a noun item, such as *man or place*. According to Mahlberg (2005, p.8), ‘general nouns refer back not only to previous noun phrases, but also to longer stretches of texts’. In other words, possible links between general nouns and preceding stretches of text provide the basis for establishing parallels between general nouns and other sub-classes of nouns. Example (3.23) and (3.24) are two cases in point:

- (3.23) Castles and Kosack have argued that capitalism requires a 'reserve army of labour' to ensure the successful continuation of the economic production along capitalist values, and that ethnic minorities provide this **army** (SS, 0001a, L1)
- (3.24) Overall, therefore, he cannot be considered as the 'father of scientific history' because he did not fulfil both of the criteria that would have earned him this **title**. (AH, 0029f, L3)

Example (3.23), which refers to *this army*, is excluded from the category of shell-noun use because its referent (e.g. *reserve army of labour*), is a noun phrase that is less than a clause in length and has somewhat metaphorical or idiomatic meaning, rather than a context-dependent meaning. Similarly, in example (3.24), the phrase *this title* is not an example of shell-noun use because its referent is smaller than a clause (e.g. *the father of scientific history*). In other words, these examples are not

considered to be shell-noun uses because they are more closely related to deictics, which do not entail any context-dependent metadiscursive meaning.

xi). The use of the potential shell noun is generic and ‘self-contained’ (Ivanic, 1991, p.110), inasmuch as the recovery of its meaning does not require any contextual information (e.g. Martin, 1992, p.103). Examples (3.25) is a case in point.

(3.25) In this **essay** I will look at what racism is, and how it is defined in contemporary society, and I will then explore why it still persists. (SS, 0001a, L1)

xii). When the potential shell-noun phrase is completely repeated in one sentence, only the first occurrence of the noun is included into the analysis. This is shown in (3.26):

(3.26) The **need to surrealise such atrocious actions** is common to all massacres as it takes subhuman mental strength to kill another person, but in the cases of Turkey and Nanking this **need** seems more overtly prevalent than in some other cases. (AH, 0012a, L1)

By contrast, partial repetition instances also occurred in this study. Different from complete repetition instances (as in (3.26)), partial repeated instances are accounted as shell-noun uses. Because the noun phrase in a partial repetition instance is not an identical version of the preceding noun phrase but an information-loaded noun phrase of a discourse segment allowing specific interpretation of the partially repeated shell unit. This is shown in (3.27) below:

(3.27) The cashier was hypnotised by a detective and shown a photo line-up, where she picked out Mayes, despite failing to identify him in a prior physical line-up. Hypnosis appears is based on the premise that **suggestion** is readily accepted during intense absorption (i.e.

where a witness concentrates on physical/mental activity, Spiegel, 1995), forensically, this **suggestion** is that memory will improve. (NS, 0020g, L3)

Whilst the syntactic pattern of the assumed shell noun (the second *suggestion*) is *th-SN*, its contextualized meaning seems to draw on the succeeding underlined *SN-be-that*-clause. Based on the preceding stretch, it may be indicated that the appearance of hypnosis is not a random case but a decision based on a premise in which suggestion might be more acceptable during intense absorption. By referring forward in the following discourse, it will be discovered that the meaning of the shell-noun phrase *this suggestion* is gradually unfolded until the *memory will improve* is reached. Therefore, the partially repeated noun phrase in this case is treated as a shell-noun use in the syntactic form of *SN-be-that*-clause, provided that the succeeding clause in the discourse allows specified context-dependent understanding of the partially repeated shell unit.

The separation of shell and non-shell uses is of great importance to this study. Only if such a distinction is clear can the analysis in the following chapters be meaningful. In conclusion, the twelve types of excluded instances discussed in section 3.4 may be summarized as follows:

i). The node noun that does not function as a complementizer in the query of potential shell nouns in the syntactic pattern of *SN-that*.

ii). The node word in the syntactic patterns *SN-be-that/SN-be-to* is a nominal phrase. Only the noun that semantically equivalent to the content coded by the *to*-infinitive clause and *be-that* clause is included as a shell noun instance.

iii). Premodifying nouns are discarded from the analysis inasmuch as they are

non-referential (e.g. *leniency* in the noun phrase *leniency bias*).

iv). Nouns in *it*-cleft sentences are not counted as shell nouns whilst nouns in existential *there* constructions are considered as instances of shell-noun uses.

v). First-order entity: visible and tangible real-world entities (e.g. *tree, earth*).

vi). First-order entity: closed set of items (e.g. *scale, tax*).

vii). First-order entity whose meaning is associated with the measurement of something, especially the length of time. (e.g. *time, period*).

viii). Second-order entity that is not meta-discursive in meaning: meanings of these nouns are in the writer's/reader's world knowledge or background knowledge on the topic (e.g. *overpopulation* in the noun phrase of *this overpopulation*).

ix). Unwanted items: the wrong item (e.g. names such as *Huntley*).

x). The referent of an assumed shell noun is smaller than a clause.

xi). The use of an assumed noun is generic and self-contained: the meaning recovery of these nouns does not require a context-specific semantic gaps. (e.g. *essay* in the noun phrase *this essay*).

xii). Complete repetitions: only the partially repeated nouns are accounted as shell-noun uses, provided that the context-dependent meaning is expressed in the succeeding clause.

3.4.2. Statistical measurements

After cleaning the data, I identified the head nouns in the concordances and calculated their frequencies Excel. As noted by Biber et al. (1999, p.263),

‘normalization is a way to adjust raw frequency counts from texts of different lengths so that they can be compared accurately’. Therefore, the final frequencies of shell nouns were normalized to the base of 1,000,000. The normalized frequencies in this study are reliable for conducting a direct and reasonable comparative analysis between the two corpora. To determine the number of occurrences of a shell noun per million words (represented by x' on the left of the following formula), the researcher divided the raw frequency (represented by x on the right of the following formula) by the total number of words in the corpus (represented by n on the right of the following formula) and then multiplied the result by one million. The calculation was performed according to the following formula: $x' = x / n * 1000000$.

The descriptive statistics presented thus far provided a basic approach to working with shell noun frequencies. However, in order to determine whether the findings are significant and not just due to chance, it is important to introduce a statistical significance test in relation to the size of the corpora under consideration. Two possible statistical tests were investigated for this study: Chi-squared values (Pearson, 1900) and log-likelihood values (Cochran, 1952).

While both Chi-squared and log-likelihood values are part of the power-divergence family of statistics (Cressie and Read, 1989, p.20), the log-likelihood test is generally preferred in corpus linguistics because it does not assume, as the chi-squared test does, that word tokens will be normally distributed in a corpus (e.g. in a bell-curve shape) (Groom, 2007, p.72). Moreover, the chi-squared test becomes unreliable when the expected frequency is less than 5 and may

overestimate results when comparing a small corpus to a much more larger one (Rayson and Garside, 2000, p.2). For these reasons, the present study has chosen to use the Log-likelihood values (henceforth LLV value). Very briefly, the LLV value compares the differences between observed frequencies and expected frequencies, and the smaller the score, the more likely the result is due to chance. Conversely, the larger the score, the more likely the result is influenced by factors other than chance. Dunning (1993) and Rayson (2003; 2008) provide a detailed mathematical explanation of the Log-Likelihood. A simplified version of the calculation is provided by Rayson and Chapelle (2019), and an online wizard for Log-Likelihood and effect size calculation is presented in Rayson (2016). In this study, the log-Likelihood score was measured using the online calculator. The LLV value needs to be above 3.84 for the difference to be significant at the $p < 0.05$ level (Rayson, 2016). The simplified version of Log Likelihood ratio calculation is illustrated in the formula (e.g. Rayson, 2016):

$$LL = 2 * ((a * \ln(a/E1) + (b * \ln(b/E2)))$$

$$E1 = C * (a + b)/(C + D)$$

$$E2 = D * (a + b)/(C + D)$$

Furthermore, the SFL framework adopted in this study is a theoretical framework for analyzing language use. It focuses on how language is used to perform social functions within specific contexts. While inter-rater reliability tests are a statistical measure used to assess the consistency and agreement between two or more raters in their evaluations of the same data, they are not always necessary or appropriate in all

research contexts.

In the case of this research, the focus is on analyzing the shell-noun uses with specific contexts. Therefore, the primary concern is to ensure the accuracy and validity of the analysis of the language structures and their meanings, rather than the agreement between raters. Moreover, SFL is often used in qualitative research to gain an in-depth understanding of language use within a particular context, rather than to measure and compare results across different raters. In this case, the use of inter-rater reliability tests may not be necessary or relevant.

Therefore, while inter-rater reliability test is an important tool for ensuring the accuracy and consistency of data analysis in many research contexts, their use is not necessary or appropriate in the present study.

3.5. Summary

Chapter 3 has explained the research methodological decisions and described the qualitative analytical procedures used in this study. Specifically, it covers the design of the corpus, the determination of shell-noun uses, and the coding of shell-noun semantic classifications. The subsequent chapters, chapter 4 and chapter 5, present quantitative and qualitative findings, respectively, and attempt to answer the research questions.

Chapter 4 Distributions of Shell-noun Use in Disciplinary Student Writing

This chapter addresses the first two research questions of this thesis:

Question 1: What is the characteristic distribution of the shell nouns, the six shell-noun syntactic patterns, and the types of disciplinary-specific transitivity process across the three disciplinary domains?

Question 2: How, if at all, do the three disciplinary domains vary in their construals of disciplinary knowledge according to the distribution of grammatical metaphors manifested by shell-noun constructions across the three corpora?

Question 1, which addresses the shell-noun uses from lexical and syntactic perspectives, is presented in Section 4.1. and 4.2. Question 2, which examines shell nouns as linguistic encodings of construals of human experience from a semantic cognitive perspective, is presented in Section 4.3.

4.1. Distributions of shell nouns across three disciplinary domains

The word counts for each of the three corpora used in the statistical calculations are 745,408 words in Arts and Humanities (AH), 192,061 in Natural Sciences (NS), and 386,882 in Social Science (SS). In the analysis, the raw data has been normalized to a base figure of ‘per million words’, and frequency ratios in the three corpora are interpreted using the log-likelihood test (see Chapter 3). The distributions of cataphoric shell nouns in the three corpora will be investigated first.

4.1.1. Distributions of cataphoric shell nouns

Table 4.1 below shows a total of 3,921 shell-noun tokens retrieved across the three disciplinary domains. By examining the intra-clause cataphoric functions of shell nouns involving the four complement constructions previously mentioned in Section 3.3, the distribution of cataphoric shell-noun tokens across the soft and hard fields demonstrates a striking difference, with 3,453 cases in the soft sciences (AH and SS), averaging 3,070 per million words, and 468 in NS, averaging 2,436 per million words ($LLV=20.8, p<0.001$). In other words, some 70 percent of all shell noun tokens occur in the more discursive soft fields.

Table 4-1

Table 4.1: Distribution of cataphoric shell nouns across three corpora (normalized per million words, *LLV* based on raw frequencies, Signif. Codes: ‘***’ 0.001; ‘**’ 0.01; ‘*’ 0.05)

Disciplinary domains	Essays	Word tokens	SN tokens			SN types	
			Raw	Normalized	<i>LLV</i>	Raw	TTR
AH	300	745,408	2,241	3,006	AH & NS 17.87***	266	0.12
NS	89	192,061	468	2,436	NS & SS -22.07***	139	0.30
SS	147	386,882	1,212	3,133	AH & SS -1.32	216	0.18
Total	536	1,324,351	3,921	2,961		240	

Furthermore, the higher frequencies of the cataphoric patterns examined here in the AH and SS corpora are evenly distributed features of the two corpora. They are reflected not only in the greater number of overall cataphoric shell-noun tokens in each of the corpora, but also in a comparably greater number of shell-noun types. While 139 distinct shell-noun types are represented in the NS corpus, 266 and 216 distinct shell-noun types are represented in the AH and SS corpora respectively. This means that nearly twice as many distinct cataphoric shell nouns are found in the AH as in the NS, and 1.6 times as many distinct cataphoric shell nouns are found in SS as in the NS.

Given that the AH corpus is roughly four times bigger than the NS, it seems as though it is hardly surprising that it would have more distinct types of shell nouns. However, the Type-Token ratios show that there is a somewhat higher number of repetitions of cataphoric shell nouns in the NS corpus (Type-Token Ratio of 0.30) than the AH corpus (TTR 0.12) and SS corpus (TTR 0.18). In overall terms, a greater diversity of shell nouns can be found in the AH and SS corpora. The greater frequency of cataphoric shell nouns suggests a tendency towards overall higher cataphoric

shell-noun uses in these two soft fields. Thus, the higher frequency of cataphoric shell nouns suggests a general tendency in these fields and is not attributed to any single outlier or cause.

To better contextualize these results, it is useful to compare shell-noun frequencies within the sub-corpora with noun frequencies overall. This is supported by a comparison of overall noun frequencies in the sub-corpora. The disciplinary domains with the highest shell-noun frequency in this study, AH (35.0%) and SS (36.5%), consist of 24.1% and 26.1% nouns, respectively. In contrast, the disciplinary domain with the lowest shell-noun frequency, NS, consists of a slightly higher percentage of nouns at 27.1%. Considering this similar, but slightly higher nominal density, it is even more striking that the shell-noun frequency in the soft knowledge disciplinary domain, especially AH, should be so much higher than that of NS disciplinary domain.

The relative lower level of shell-noun uses in NS might suggest a tendency towards a more technical and specialized discourse. In other words, the lower number of shell nouns in NS might result from a greater number of specific technical terms that are ‘fully lexicalized within the community and do not have general-purpose shell noun counterparts’ (Flowerdew and Forest, 2015, p. 94). This explanation is probable since new knowledge constructions in the hard sciences rely more on the assumed validity of certain empirical observations, conventional empirical practices, and laboratory procedures that are distinctive to the hard disciplinary communities (e.g. Hyland, 2004a; 2004b; 2008; 2012).

Additionally, this possible explanation is consistent with those of previous studies (e.g. Omidian and Siyanova-Chanturia, 2021; Durrant, 2014). Omidian and Siyanova-Chanturia's (2021) study reveals that the language of hard knowledge field writing (e.g. empirical research articles) is 'characterized by highly specialized, field-specific discourses that are far from being homogeneous' and may 'not be of equal value to writing in different branches of academia' (Omidian and Siyanova-Chanturia, 2021, p.27). Furthermore, Durrant's (2014) study offers evidence that university students' writing in Science and Technology fields makes greater use of specialized vocabulary to support their knowledge claims compared to Social Sciences and Humanities. By contrast, the higher frequency of shell nouns in AH and SS corpora might indicate a similar mode of knowing. Since 'new perspectives on concepts and entities in soft sciences are often taken from everyday general experiences' (Biber, 2006, p.45), these shell nouns might be useful tools for writers in soft fields to refer to common daily human experiences.

Having established that the two soft fields exhibit a higher frequency than the hard field, this thesis adds nuance to this finding as: 1). on the whole, there is an overall lexical similarity across the three disciplinary domains; 2). the two cognate soft fields have essentially similar preferences in the choice of cataphoric shell noun types. For the moment, turning to the three sets of the cataphoric shell nouns, [Table 4.2](#) and [Table 4.3](#) list the unique cataphoric shell nouns along with their frequencies of occurrences in AH and SS, respectively, while those from NS are listed in [Table 4.4](#) (see Appendix 1). Comparing these three tables, a total number of 240 distinct

cataphoric shell nouns combined from the three corpora are identified. These are listed in [Table 4.5](#) (see Appendix 2). Table 4.6 below shows the most frequent cataphoric shell nouns across the three disciplinary domains.

Firstly, as can be seen from Table 4.6, AH writing tends to prefer *fact* (freq.=406), *attempt* (freq.=239), *ability* (freq.=197) and *idea* (freq.=152). After these four frequently occurring nouns, the normalized frequency rates of nouns in AH quickly drops to *belief* (freq.=99), *need* (freq.=80), and *desire* (freq.=79), followed by a large number of low-frequency shell nouns (freq.<70). Similarly, the frequency rates of the four most frequent nouns in SS *fact* (freq.=364), *ability* (freq.=364), *idea* (freq.=150) and *attempt* (freq.=132) quickly drop to *need* (freq.=121), *argument* (freq.=99), *belief* (freq.=85), *opportunity* (freq.=76), and *view* (freq.=75), followed by a large number of shell nouns whose normalized frequency rates are below 70. The cataphoric shell-noun uses in AH and SS are thus characterized by a wider range of nouns and a frequency drop in a steep Zipfian manner (1936; 1949) from item to item. This distribution approximately follows a mathematical form in which the r^{th} (i.e. frequency rank) most frequent word has a frequency $f(r)$ that scales according to: $f(r) \propto \frac{1}{r^\alpha}$ for $\alpha \approx 1$ (Zipf, 1936; 1949). To put it simply, there are the very few high-frequency words that account for most of the tokens in text and a wide range of low-frequency words.

Turning now to the NS corpus, the frequency data for cataphoric shell nouns in NS exhibits a predominantly high frequency of *fact* and *ability* (freq.=208 for both) as well as a smaller range of low-frequency items.

Table 4-6

Table 4.6. Most frequent cataphoric shell nouns across the three disciplinary domains (*freq.* Rounded figures normalized per million words)

AH			NS			SS		
Rank	SN	<i>Freq.</i>	Rank	SN	<i>Freq.</i>	Rank	SN	<i>Freq.</i>
1 st	fact	406	1 st	ability	208	1 st	fact	364
2 nd	attempt	239	2 nd	fact	208	2 nd	ability	364
3 rd	ability	197	3 rd	need	115	3 rd	idea	150
4 th	idea	152	4 th	attempt	104	4 th	attempt	132
5 th	belief	99	5 th	<i>tendency</i>	73	5 th	need	121
6 th	need	80	6 th	<i>evidence</i>	68	6 th	<i>argument</i>	93
7 th	<i>desire</i>	79	7 th	view	57	7 th	belief	85
8 th	<i>notion</i>	68	8 th	<i>problem</i>	52	8 th	<i>opportunity</i>	76
9 th	view	62	9 th	<i>aim</i>	52	9 th	<i>assumption</i>	54
10 th	<i>failure</i>	56	10 th	<i>advantage</i>	47	10 th	view	75

Furthermore, boldfaced shell nouns in Table 4.6 are found in all three corpora, while shading indicates shell nouns shared by AH and SS but not included in NS. From this, we can see that AH has more in common with SS (seven cataphoric shell nouns in common) than with NS (five cataphoric shell nouns in common). The overall lexical similarity across the three corpora shown in Table 4.6 is complemented by a more significant observation of a common set of 94 cataphoric shell nouns. This common set of 94 shell nouns covers roughly 90% of the total use of cataphoric shell nouns across the three disciplinary domains. [Table 4.7](#) (see Appendix 3) lists the common set of cataphoric shell nouns found across the three corpora. The shell nouns are ranked (*No.*) according to the joint frequencies in descending order, and the four right columns four sets of data: Combined, AH, NS and SS. Each set of data contains three types of statistics: raw frequency (*Raw Freq.*), individual coverage (*Ind Cover*) and cumulative coverage (*Cum. Cover*)

Now consider Table 4.8 below, which is a part of Table 4.7, provides an example using the shell noun *assumption*. This shell noun ranks 11th in frequency and has a combined raw frequency of 64. As an individual item, it covers 1.63% of the total 3,921 uses of shell nouns across the three corpora (2,214 in AH, 468 in NS and 1,212 in SS), as indicated under *Ind Cover*. The *Cum. Cover* of 47.62% suggests that *assumption*, along with the previous top ten shell nouns, jointly cover almost half of the total shell noun uses. In AH, with a raw frequency of 42, along with ten other higher-ranking shell nouns, the *Cum. Cover* of 49.53%, compared with 49.26% in SS. In contrast, the same item in NS corpus has a raw frequency of only 9 and a cumulative coverage of just 35.90%, suggesting that the top 11 shell nouns jointly cover only 35.90% of the total shell noun uses in NS corpus. This example indicates that AH and SS cataphoric shell-noun uses are characterized by a fundamental preference for a similar core set of shell nouns, which jointly cover a large proportion of the total shell-noun uses.

Table 4-8

Table 4.8: A set of common shell nouns in three corpora

No.	SNs	Combined			AH			NS			SS		
		Raw freq.	Ind. cover %	Cum. cover %	Raw freq.	Ind. cover %	Cum. cover %	Raw freq.	Ind. cover %	Cum. cover %	Raw freq.	Ind. cover %	Cum. cover %
1	fact	484	12.34	12.34	303	13.52	13.52	40	8.55	8.55	141	11.63	11.63
2	ability	328	8.37	20.71	147	6.56	20.08	40	8.55	17.09	141	11.63	23.27
3	attempt	249	6.35	27.06	178	7.94	28.02	20	4.27	21.37	51	4.21	27.48
4	idea	175	4.46	31.52	113	5.04	33.07	4	0.85	22.22	58	4.79	32.26
5	need	129	3.29	34.81	60	2.68	35.74	22	4.70	26.92	47	3.88	36.14
6	belief	114	2.91	37.72	74	3.30	39.05	7	1.50	28.42	33	2.72	38.86
7	view	86	2.19	39.91	46	2.05	41.10	11	2.35	30.77	29	2.39	41.25
8	desire	81	2.07	41.98	59	2.63	43.73	4	0.85	31.62	18	1.49	42.74
9	notion	79	2.01	43.99	51	2.28	46.01	6	1.28	32.91	22	1.82	44.55
10	argument	78	1.99	45.98	37	1.65	47.66	5	1.07	33.97	36	2.97	47.52
11	assumption	64	1.63	47.62	42	1.87	49.53	9	1.92	35.90	21	1.73	49.26

12	failure	64	1.63	49.25	26	1.16	50.69	8	1.71	37.61	30	2.48	51.73
13	right	56	1.43	50.68	30	1.34	52.03	4	0.85	38.46	22	1.82	53.55
14	inability	55	1.40	52.08	36	1.61	53.64	6	1.28	39.74	13	1.07	54.62
15	effort	52	1.33	53.40	32	1.43	55.06	7	1.50	41.24	13	1.07	55.69

4.1.2. Distributions of anaphoric shell nouns

We turn our attention now to the distributions of anaphoric shell-noun uses in the three corpora. Table 4.9 below illustrates the distributions of anaphoric shell-noun uses by disciplinary domains as a whole. In the analysis of the immediately preceding section it becomes apparent that the overall cataphoric shell-noun uses in soft are remarkably higher than that of hard field. The picture of this higher frequency of shell-noun uses is not, however, uniform in terms of anaphoric shell noun uses.

Table 4-9

Disciplinary domains	Essays	Word tokens	SN tokens			SN types	
			Raw	Normalized	<i>LLV</i>	Raw	TTR
AH	300	745,408	394	542	AH & NS -140.6****	135	0.33
NS	89	192,061	300	1,432	NS & SS 13.9****	89	0.30
SS	147	386,882	413	1,068	AH & SS -92.3****	132	0.32
Total	536	1,324,351	1,107			204	

Table 4.9 shows that a total of 1,107 instances of anaphoric shell nouns were retrieved from the three corpora, with 204 different types of anaphoric shell nouns identified. It is important to note the distinction between anaphoric and cataphoric shell nouns. Notably, the distribution of anaphoric shell-noun tokens across the soft and hard fields is markedly different. The two soft corpora (AH and SS) yielded 817 cases, averaging 722 per million words, while the NS corpus yielded 275 cases,

averaging 1,432 per million words. In other words, NS accounts for approximately 50 percent of the total number of anaphoric shell-noun tokens in the three corpora, while AH and SS contribute the other half. Interestingly, the figures of anaphoric shell-noun tokens exhibit an opposite trend to that of cataphoric shell-noun tokens in terms of soft versus hard disciplinary divide. Additionally, while the AH and SS corpora have a wider range of cataphoric shell nouns than the NS corpus, the Type-Token ratios reveal some similarities in terms of the diversity of anaphoric shell nouns across the three corpora (TTR: 0.30~0.33).

While the diversity of anaphoric shell-noun types appears roughly to be similar across the three corpora, the rate of occurrence of anaphoric shell nouns in the NS corpus is higher overall, contributing to nearly half of the anaphoric shell-noun tokens in this study. One possible explanation for this finding is that the anaphoric shell-noun uses in the NS corpus are drawn from a small set of anaphoric shell nouns that cover a larger portion of the total uses. This is supported by a detailed investigation of the three sets of anaphoric shell nouns. [Table 4.10](#) and [Table 4.11](#) present the unique anaphoric shell nouns along with their frequency of occurrences in the AH and NS corpora, respectively, while [Table 4.12](#) in Appendix 4 lists those those from the SS corpus.

This thesis identifies a total number of 204 distinct anaphoric shell nouns across the three corpora, as listed in [Table 4.13](#) (see Appendix 5). [Table 4.14](#) below presents the top ten most frequently occurring anaphoric shell nouns across the three disciplinary domains, with shell nouns that are shared by all three domains boldfaced. Shading in [Table 4.14](#) highlights anaphoric shell nouns that are unique to each of the

three corpora.

Table 4-14

Table 4.14. Most frequent anaphoric shell nouns across the three disciplinary domains (*freq.* Rounded figures normalized per million words)

AH			NS			SS		
Rank	SN	<i>Freq.</i>	Rank	SN	<i>Freq.</i>	Rank	SN	<i>Freq.</i>
1 st	way	40	1 st	case	125	1 st	view	47
2 nd	<i>argument</i>	24	2 nd	<i>method</i>	120	2 nd	case	43
3 rd	idea	21	3 rd	way	99	3 rd	way	41
4 th	<i>point</i>	20	4 th	idea	52	4 th	idea	39
5 th	case	19	5 th	view	47	5 th	theory	36
6 th	view	17	6 th	<i>definition</i>	47	6 th	<i>argument</i>	36
7 th	<i>statement</i>	16	7 th	<i>reason</i>	42	7 th	<i>approach</i>	36
8 th	theory	13	8 th	theory	36	8 th	<i>point</i>	34
9 th	<i>question</i>	13	9 th	<i>problem</i>	36	9 th	<i>question</i>	26
10 th	<i>change</i>	12	10 th	<i>evidence</i>	36	10 th	<i>issue</i>	26

This table provides immediate insight into the discourse of the NS discourse, as represented by its distinct shell nouns, such as *method*, *reason*, *problem* and *evidence*. These shell nouns suggest a focus on scientific methods, experimentation evidence, practical issues and so on., as illustrated by examples (4.1) through (4.3).

- (4.1) Before genetic modification, desired traits were achieved through farmers by selective breeding. However this **method** is very time consuming and it is not accurate unlike genetic modification. (NS, 6037e, L1)
- (4.2) The landscape has more cover than the desert regions and it is more humid. For this **reason** tanks of reptiles in captivity must be sprayed with a mist of water at least twice a week. (NS, 6011h, L1)
- (4.3) Should the gene be harmful, to the plant or indeed to humans, the damage this would cause is indeed reason to be cautious. This **problem** can be managed, albeit with difficulty, such as to minimise the risks of contamination. (NS, 0181b, L1)

By comparing the list of shell nouns in the NS corpus of those in the SS and AH

corpora, we can observe that AH and SS share more in common with each other. For instance, they share three shell nouns that are not present in the NS corpus (e.g. *argument, point, question*), as shown in examples (4.4) from SS and (4.5) from AH.

- (4.4) This can be seen in that almost invariably the dominant, high-paid and high-status jobs were given to men (such as Doctor and surgeons), whilst the subordinate, low-paid and low-status jobs were given to women. Whilst this **argument** holds true... (SS, 0075g, L1)
- (4.5) However, since values are intrinsically action-guiding (i.e. they affect our actions) they cannot be part of the fabric of the world. This **argument** follows, presuming you accept Mackie's definition that to be objective something must be characterisable independently of us. (AH, 0113b, L1)

In contrast, if we consider the shaded shell nouns in the NS corpus and compare them to those in the SS and AH corpora, SS shows more similarities with NS than AH. For instance, *approach* and *issue* share semantic affinity with the shell nouns *method* and *problem*, respectively, as demonstrated in examples (4.6) through (4.9). This suggests that, while SS shares some similarities in anaphoric shell nouns with AH, it may also place a larger emphasis on the use of certain anaphoric shell nouns that are closely related to scientific methods and practical issues, similar to the NS corpus.

- (4.6) ...the courts view that it would be unseemly distressing to allow husbands and wives...to use the court as an arbiter for their matrimonial differences. One must question whether this **approach** is incongruous with today's society. (SS, 0148a, L1)
- (4.7) The child is then encouraged to resume sleep and gradually the time between each awakening is lengthened. This **method** appears very harsh to parents and suffers from a low compliance rate. (NS, 0421a, L3).
- (4.8) In midst of their victimhood, women are often pushed to take on greater roles and

responsibilities within the home and beyond. Hence, the crux of the essay is to transcend stereotypes of women's non-agency in wars by adopting a post-modern and social constructivist perspective towards this **issue**. (SS, 0137n, L3)

- (4.9) The impractical clothing places the NHS at risk of legal proceedings, as an employer it has the responsibility of protecting employees from injury at work, back injuries affect a large proportion of nurses and midwives. This **problem** appeared to be tackled... (NS, 3034d, L1))

To further explore this analysis, a set of 45 distinct anaphoric shell nouns commonly used across the three corpora are listed in [Table 4.15](#) (see Appendix 6). As shown in the table, this common set of 45 shell nouns covers roughly 60% of the total use of anaphoric shell nouns in each of the soft disciplinary domains, and 76% in the NS corpus. For instance, *method* is the 5th most frequent anaphoric shell noun, with a combined raw frequency of 36. When examined individually, *method* accounts for 3.26% of the total anaphoric shell noun uses, while in the AH corpus, it covers only 1.98%, and 1.21% in the SS corpus. However, in the NS corpus, *method* covers over 8% of the total anaphoric shell nouns uses. These findings suggest that the anaphoric shell nouns in the NS corpus are characterized by a strong preference for certain nouns, such as *methods*, *reason*, *problem*, and *evidence*, while the frequencies of anaphoric nouns in the AH and SS corpora are more evenly distributed and decrease more gradually.

4.1.3. Changes of the shell-noun uses

Overall, the L1 corpus consists of 330 texts with a total of 678,433 words, while the L3 corpus comprises 206 texts with 645,918 words, indicating a slight decrease in

overall word count. However, there is a significant increase of 52.53% in the average length of text across levels of study. Table 4.16 displays the corpora by levels of study and word length.

Table 4-16

Table 4.16: Corpora by levels of study and word length.

Discipline domain	Level 1	Level 3	Change (%)
AH	413,663	331,745	-19.80
NS	107,705	84,356	-21.68
SS	157,065	229,817	46.32
Average word length of text	2056	3136	52.53
Total words	678,433	645,918	-4.79

This analysis identified 2,828 shell nouns in the L1 corpus, with an average of 41.68 cases per million words, and 2,609 shell nouns in the L3 corpus, with an average of 40.39 cases per million words. This represents a modest decrease of 3% since the first year of study, but the result is not statistically significant (*Log Likelihood*=1.35, $p > 0.05$). The study hesitated to interpret this slight decline as an indication of any dispreference for the use of shell nouns in L3 writers' academic arguments, as it is likely due to the general decrease in overall word length.

4.1.4. Summary: distributions of shell nouns

The distribution of shell nouns across the three disciplinary domains reveals a striking difference between the hard sciences and soft sciences domains in terms of overall shell-noun use. Specifically, the distributions of cataphoric shell nouns showed no statistical significance between the two cognate soft fields. However, the most noteworthy difference relates to the soft versus hard disciplinary divide. Cataphoric shell nouns occur significantly more frequently in the two soft disciplinary

domains (AH and SS) than in the hard disciplinary domain (NS). The soft fields together account for approximately 70% of the total number of shell nouns in the corpus, while the hard field accounts for only 30%. These greater frequencies of cataphoric shell-noun usage appear to be a general trend of the two soft corpora, rather than attributable to a single outlier or cause, as the diversity of shell noun types is also higher than in the NS corpus. Additionally, this thesis adds nuance to this finding by noting that: 1). overall, there is lexical similarity across the three disciplinary domains (e.g. a common set of 94 shell nouns identified in cataphoric shell-noun usage); and 2). the two cognate soft fields have similar preferences in the choice of cataphoric shell-noun types (see section 4.1.1).

Secondly, the distributions of anaphoric shell nouns reveal that they occurred much more frequently in the NS corpus than in the AH and SS corpora. This higher frequency in the NS corpus is mainly due to repetitions of certain shell nouns such as *way*, *method*, and *case*, which denote how an experimental process takes place.

The larger presence of these shell nouns probably suggests a more empirically oriented approach to academic writing in this hard field, with a general concern for finding and testing established methods or procedures whose usefulness may be helpful in addressing the practical concerns in this disciplinary domain.

4.2. Distributions of shell-noun lexico-grammatical patterns

We now move on to the distribution of shell-noun lexico-grammatical patterns, firstly for the shell-noun complement constructions: *SN-that*, *SN-to*, *SN-be-that* and

SN-*be-to*

4.2.1. Distributions of four shell-noun complement constructions

In terms of shell-noun complement constructions across the three corpora, table 4.17 below outlines some discipline-specific tendencies:

Table 4-17

Table 4.17: Distribution of shell noun complement constructions across the three corpora (*Freq.* normalized per million words, *Prop%* based on raw frequencies)

	SN- <i>that-cl</i>		SN- <i>to-cl</i>		SN- <i>be-that-cl</i>		SN- <i>be-to-cl</i>		Total	
	Freq.	Prop(%)	Freq.	Prop(%)	Freq.	Prop(%)	Freq.	Prop(%)	Freq.	Prop(%)
AH	1,468	<u>51.63</u>	1,022	35.96	203	7.13	1,150	5.29	2,843	100
NS	823	34.19	1,015	<u>41.94</u>	385	15.91	193	7.96	2,421	100
SS	1,458	<u>47.12</u>	1,171	37.84	264	8.52	202	6.52	3,094	100
Total	1,372	48.06	1,065	37.29	247	8.65	171	6.00	2,855	100

The most pronounced difference between the corpora concerns the concentration of the most frequently used construction of SN-*that-clause* in AH and SS, which occupies around half of all the shell noun complement constructions in each corpus. However, in NS, SN-*to-cl* is approximately 1.2 times more frequently used than SN-*that-cl*.

Furthermore, while AH and SS reveal a similar use between SN-*be-that-clause* and SN-*be-to-clause* constructions, in NS, SN-*be-that-clause* is over 2 times more frequently used than SN-*be-to-clause*. Overall, SN-*that-cl* is the predominantly key to shell noun syntactic realization in AH and SS, while SN-*to-clause* and SN-*that-clause* are the most frequently used patterns in NS. This observation supports previous research findings (e.g., Flowerdew and Forest, 2015; Jiang and Hyland, 2015) that

SN-*that*-cl is the most frequent syntactic construction in academic prose in terms of shell noun realizations across disciplinary domains.

More importantly, this thesis reveals a visible difference: the NS corpus prefers non-finite *to*- clauses as complement constructions to shell nouns, while AH and SS exhibit a distinctive preference for finite *that*-clause shell noun complement construction. This preference for non-finite *to*-clause in the NS corpus seems surprising, considering the intrinsic nature of hard knowledge fields and the general function of *to*-infinitive complement clauses, which are commonly used to express human intention, future-oriented agency, and human control over action, as suggested by Biber et al. (1998) and Quirk (2010). This preference appears to break from a traditionally stressed scientific paradigm that emphasizes objectivity and neutrality, ‘as far removed from the human touch as possible’ (Mauranen and Bondi, 2003, p. 269), and may indicate a way of constructing authorial voice in meaning-making within this disciplinary domain.

The prevalence of non-finite shell noun complement constructions in the NS corpus could be related to an empirical-oriented and compressed type of writing through *to*-infinitival clauses with a tone of imperative modality, as demonstrated in the examples below. Through these complement constructions, students tend to present or introduce a particular empirical method or procedure into the ‘scene’, thereby raising it into the readers’ consciousness subtly (as in (4.10)). They also establish the novelty and newsworthiness of their studies with an intrusive personal voice (as in (4.11)), and address empirical difficulties or rectify potential problems in

highly standardized, almost shorthand, ways that presuppose a degree of theoretical knowledge and routine practices (as in (4.12)).

- (4.10) There is a quicker **way to do this by simply using Gauss' divergence theorem** (comes about after the time of Green), but I like the originality that Green has shown by introducing the idea of reciprocity into his proof so I have left it as such. (NS, 0323b)
- (4.11) What I will say however, is that it is a path which rejects General Relativity and Quantum Mechanics in an **attempt to provide new answers from fundamental, philosophical principles**. (NS, 6097j)
- (4.12) In an **attempt to overcome this problem software companies like Microsoft have tried to portray their products with a certain image**. (NS, 0228c)

The use of finite *that*-clause construction in the soft knowledge field differentiates itself from the hard knowledge field. The general function of SN-*that*-clauses, as suggested by researchers (e.g., Hyland and Tse, 2005a; 2005b; Jiang and Hyland, 2015), is to help foreground the writer's evaluation of the reliability of information and the assessment of how the readers should understand it. The examples below reveal that finite *that*-clause construction in AH and SS is more likely to be employed as a proactive effort in promoting 'semantic bleaching' or 'grammaticalization', which is a means of changing ideas or possible fact to fact under the disguise of the finite-*that* clause construction (Schmid, 2000, p. 100). Typically, in example (4.13), the writer presents a whole string of claims as one irrefutable objective 'fact', while in example (4.14), the writer reports the critical views from other authorities as 'fact', again not open to discussion. This finding provides empirical support for the impression that AH and SS students favor a

comparatively more discursive and interpretative presentation of meaning-making.

- (4.13) Vincent Cronin has discussed the **fact** that Burckhardt took Renaissance Italy as a whole and made no distinction between what happened in republican Florence and other cities which were ruler over by tyrants. (AH, 0013b)
- (4.14) it could be attributed to the **fact** that as noticed by many critics, since it has adopted a tabloid format in 2002 its drive to become unique, innovative and engaging has resulted in the erosion of its intellectual and critical substance. (SS, 0075h)

4.2.2. Distributions of anaphoric lexico-grammatical patterns

We now examine the distribution of anaphoric lexico-grammatical patterns across the three corpora. Shown below in Table 4.18 are the occurrences of the two anaphoric lexico-grammatical patterns.

Table 4-18

	<i>th-SN</i>		<i>th-be-SN</i>		Total	
	Freq.	Prop.(%)	Freq.	Prop.(%)	Freq.	Prop.(%)
AH	404	86.0	66	14.0	470	100
NS	275	94.8	15	5.2	290	100
SS	413	94.3	25	5.7	438	100
Total	1,092	91.2	106	8.8	1,198	100

Table 4.18 reveals that the most pronounced feature is the highly uniform concentration of the *th-SN* syntactic pattern across the three corpora. Additionally, the *th-be-SN* lexico-grammatical pattern shows marginal frequency rates in the AH, NS and SS corpora.

4.2.3. Changes of shell-noun syntactic patterns

Table 4.19 outlines the distributions of shell-noun lexico-grammatical syntactic patterns across levels of study, and the differences here show no statistical significance. However, further analysis of the variation of shell-noun syntactic patterns by disciplinary domain reveals that this non-significant change is only superficial.

Table 4-19

Table 4.19: Changes of shell-noun syntactic patterns (frequency normalized per million words, <i>LLV</i> based on raw frequency, Signif. Codes: '***' 0.001; '**' 0.01; '*' 0.05)			
	Level 1	Level 3	<i>LLV</i>
N- <i>that</i>	1,462	1,390	1.20
N- <i>to</i>	1,135	1,204	-1.37
N- <i>be-that</i>	233	276	-2.38
N- <i>be-to</i>	195	159	2.31
<i>th-N</i>	1,107	1,013	2.79
<i>th-be-N</i>	66	50	0.22

Table 4.20 presents these diachronic changes of shell-noun syntactic patterns by each of the three disciplinary domains.

Table 4-20

Table 4.20: Disciplinary changes of shell-noun syntactic patterns (frequency normalized per million words, <i>LLV</i> based on raw frequency, Signif. Codes: '***' 0.001; '**' 0.01; '*' 0.05)									
	AH			NS			SS		
	L1	L3	<i>LLV</i>	L1	L3	<i>LLV</i>	L1	L3	<i>LLV</i>
N- <i>that</i>	1,438	1,531	-1.07	743	1,185	-9.79**	2,018	1,262	33.33***
N- <i>to</i>	1,010	1,206	-6.36*	1,003	1,209	-1.83	1,553	1,201	8.51**
N- <i>be-that</i>	220	205	0.19	316	486	-3.48	210	300	-2.95
N- <i>be-to</i>	152	151	0.003	288	119	6.76**	242	187	1.32
<i>th-N</i>	701	742	-0.42	1,838	1,482	3.61	1,674	1,231	12.78***
<i>th-be-N</i>	42	24	0.47	11	14	-1.46	13	12	0.25

There are remarkable differences of shell-noun syntactic patterns across levels of

study within each disciplinary domain, and these differences are not uniform across the three disciplinary domains. For example, the *SN-that* pattern increased significantly in the hard disciplinary domain (e.g. NS) while it showed a remarkable decline in the soft disciplinary domain (e.g. SS).

Before moving further, it is useful to briefly recap the logico-semantic relations grouped under the two umbrella headings of projection and expansion, as outlined in section 2.5. To recapitulate briefly, *Projection* refers to the linguistic representation of thoughts and feelings and it accounts for the representations of ideas (as in (4.15)) and locutions (as in (4.16)):

- (4.15) Initially, it will be shown that the adventure story of late nineteenth century England provided young Englishmen with the **belief** that it was their right to propagate Imperialism because of their natural superiority to other ethnicities. (SS, 0004a, L1)
- (4.16) This rests on the **argument** that New Labour's welfare policies are socially authoritarian. (SS, 0075a, L1)

Expansion involves the account of various kinds of comparisons and contrasts, such as temporal, spatial and casual relationships in discourse, among others. Logico-semantic relations encompass *elaboration*, such as specification, exemplification and clarification; *extension*, which involves addition or variation, replacement, and the presentation of alternatives; and *enhancement*, such as circumstantial and causal expansion. For instance, in example (4.17), the logico-semantic relation is one of *elaboration*, where the shell noun restates or paraphrases the previous text. In example (4.18), the logico-semantic relation is one

of *extension*, where there is an extension of meaning by providing an addition with an adversative feature in the *that*-clause. Finally, in example (4.19), the logico-semantic relation is one of *enhancement*, as the *that*-clause provides the meaning of manner.

- (4.17) The **fact** that these quotes are not from The New Law of Righteousness does not really matter, since...(AH, 0019e, L3)
- (4.18) I feel this example is slightly doctored, as Green feels the need to justify his actions, but provides **evidence** that he tries to link his work to the real world. (NS,0323b)
- (4.19) Humanism is unique in the **way** that it emphasises the importance of subjective experience in understanding an individual. (NS, 3069c)

The concepts of *projection* and *expansion* are useful in understanding how shell nouns function and can shed light on the writers' operation of the decision-making process. For instance, they allow an interpretation of example (4.17) *The fact that these quotes are not from The New Law of Righteousness*, as a case of *expansion*, conveying a truth, and to understand this linguistic decision as the writer's strategic rhetorical choice since the writer could have chosen projection by using a shell noun such as *claim* or *belief*. It may have seemed more effective to the writer to present information as 'a figure of being' rather than 'a figure of sensing' (Halliday and Matthieseen, 1999, pp.144-147) in this case.

The argument can be summarized as follows: the use of the SN-*that* pattern enables writers to create a logical and cohesive flow of information by helping readers understand 'how information can be traced in the text' (Dahl, 2004, p.1820). In other words, this pattern helps establish a cohesive frame of reference by setting up

prospective references for what is to come. As noted by Hunston (2013, p.626), full expositions (*that*-clause) are often chosen when overt assessment of the proposition is required. Therefore, the **SN-*that*** pattern is useful not only for establishing a cohesive frame of reference but also foregrounding the writer's interpretation and evaluation of further claims with a coherent logic chain. By using this pattern, the writer can guide the reader through a well-structured and persuasive argument.

The soft domain is characterized as having relatively 'loosely knit academic communities' (Becher and Trowler, 2001, p.33), with research conducted within less precisely defined boundaries of knowledge and less clearly agreed disciplinary problems than the hard sciences. As a result, 'the criteria the audience will apply are not clear-cut and universal, nor is it certain what intellectual framework they will bring to the reading' (Bazerman, 1988, p. 34). Given the discursive nature of the soft disciplinary domains, it seems that the **SN-*that*** pattern might be a useful rhetorical choice to align the writer's personal 'interpretative persuasion' and 'sympathetic understanding' (Hyland, 2008, p.16) with the readers.

In contrast, writers in hard knowledge disciplinary domains conduct readers along pathways with more precisely defined ways of seeing the world, and new knowledge is often recognized by broadly accepted paradigms and methods. Therefore, the cohesion of texts is often implicitly established by virtue of their 'craft skills in specialized discourse' and 'tacit knowledge from their daily work' rather than explicitly marked lexical realization (Myers, 1991, p. 6).

However, the observed increase of the **SN-*that*** pattern in NS's L3 writing and its

decrease in SS's L3 writing may reflect some changes in disciplinary academic writing that are not as straightforwardly explained by the typical soft vs. hard disciplinary divide. This suggests the complexity of shell noun use and its role in academic writing. This interesting and counter-intuitive change will be addressed in chapter 5 through a detailed analysis of lexicalizations of shell nouns. By examining the specific ways in which writers use shell nouns in their writing, we can gain a deeper understanding of the factors that influence their rhetorical choices and how these choices contribute to the effectiveness of their arguments.

4.2.4. Summary: distribution of shell-noun syntactic patterns

In terms of shell-noun syntactic patterns, this thesis reveals that the NS corpus exhibits a preference for non-finite *to*-clauses as complement constructions to shell nouns, while AH and SS exhibit a distinctive preference for finite *that*-clauses. This finding is novel and interesting because finite *that*-clauses tend to construe neutral facts as arguments, while *to*-infinitival embody modality of varying degrees of human intentions and efforts (see examples (4.10) through (4.12) in section 4.2.1). All three corpora show a concerted effort to exploit the textual coherence functions of the *th-SN* syntactic pattern. In contrast, the *th-be-SN* syntactic pattern remains at a low level of use in AH and is completely absent from the NS and SS corpora.

It is worth noting that the *SN-that* syntactic pattern has increasingly become a powerful persuasive device in NS writing, rising by 60% from a low base in the first

year of study. This pattern enables NS writers to foreground their preferred assessment of the reliability of what follows and build an inclusive relationship with readers. The remarkable decreases in the use of the *SN-to* syntactic pattern in the two soft disciplinary domains might indicate a move away from expressing the experience of human intention, future oriented agency, and human control over action. Examples (4.20) through (4.21) emphasize the volitional force of personal intention, which is increasingly replaced by examples that refer to the manner in which actions are conducted (examples (4.22) and (4.23)):

(4.20) His whole demise is derived though a **desire to stay young, mostly courted by the compliments he receives regarding his beauty and comments** such as ... (AH, 0012b, L1)

(4.21) I feel the **need to explore this term further, in order to form a central understanding of what 'naturalism' refers to.** (AH, 0024a, L1)

Examples (4.22) and (4.23), which were retrieved from level 3 sub-corpora differ from examples (4.20) and (4.21) in that the information focus of the *to*-infinitival clause is not so much on the volitional force of the personal experiences, but rather on scientific methods.

(4.22) As few people took part in the Revolution itself, the best **way to analyse the amount of support the Bolsheviks actually had,** is to examine events leading up to the Revolution. (AH, 0029b, L3)

(4.23) with the suggestion that fewer object request statements parallels the pre-linguistic period where language is used as a **means to engage people socially rather than to engage with the surrounding environment.** (AH, 6206d, L3)

The following analysis focuses on the divergent uses of shell nouns from a semantic perspective. To facilitate a fine-grained analysis of these uses, this thesis adopts the semantic classification system proposed by Schmid (2000) for the sub-categorization of shell nouns.

4.3. Distribution of types of construal instantiated by shell nouns across the three corpora

To unveil disciplinary variations in the use of shell nouns, the following sections provide a detailed investigation through an analysis of the specific distributions of transitivity processes construed by distinct shell nouns across the three disciplinary domains. The analysis will first focus on cataphoric shell nouns (see section 4.3.2 for anaphoric shell nouns).

4.3.1. Distribution of types of construal instantiated by cataphoric shell nouns

The semantic and cognitive patterns of preferences for different transitivity processes further support the typical soft vs hard divide. Table 4.21 displays the frequency distributions of different types and sub-types of transitivity processes across the three corpora:

Table 4-21

Table 4.21: Distributions of different types of transitivity process construed by cataphoric shell nouns across three disciplinary domains. (*freq.* Rounded figures normalized per million words, *prop%* based on raw frequency)

Types of transitivity process	Sub-types of transitivity processes	AH			NS			SS		
		Freq.	Prop. %	Type	Freq.	Prop. %	Type	Freq.	Prop. %	Type
Attributive relational process	Comparative	11	0.36	5	2	0.42	2	18	0.56	1
	Attitudinal	38	1.25	52	32	6.65	10	57	1.75	10
	Epistemic	47	1.56	21	9	1.87	4	49	1.51	4
	Deontic	219	7.27	109	46	9.56	13	331	10.19	16
	Dynamic	334	11.11	78	104	21.62	13	411	12.66	13
	Sub-total	648	21.55	266	193	41.26	42	866	27.64	44
Identifying relational process	Neutral	431	14.19	36	41	8.52	2	395	12.18	5
	Partitive	66	2.17	104	15	3.12	10	96	2.95	11
	Sub total	496	16.51	141	56	11.96	12	491	15.67	16
Mental process	Evidential	72	2.39	26	15	3.12	3	31	0.96	3
	Conceptual	135	4.47	73	23	4.78	8	145	4.46	14
	Creditive	581	19.14	292	51	10.90	27	633	19.51	40
	Dubitative	13	0.44	16	5	1.06	2	36	1.11	1
	Emotive	24	0.80	47	1	0.21	1	5	0.16	2
	Volitional	251	8.27	198	41	8.76	20	385	11.86	37
	Sub total	1,077	35.83	651	136	29.05	61	1,236	39.45	62
Verbal process	Propositional	4	0.13	16	0	0.00	0	8	0.25	3
	Assertive	369	12.30	167	35	7.52	15	302	9.68	27
	Rogative	3	0.09	5	2	0.43	1	3	0.08	1
	Directive	13	0.44	16	0	0.00	0	10	0.33	3
	Commissive	7	0.22	21	2	0.43	1	41	1.32	4
	Expressive	0	0	0	0	0	0	0	0	0
	Sub total	396	13.19	224	39	8.39	17	364	11.66	38
Material process	General	4	0.13	16	1	0.21	1	13	0.40	5
	Specific	294	9.77	115	37	7.69	12	85	2.72	14
	Attitudinal	83	2.74	42	3	0.62	3	70	2.22	3
	Sub total	381	12.67	172	41	8.76	16	168	5.36	22
Total	2,998	100	1,453	465	100	148	3,125	100	182	

Table 4.21 shows that the mental process construed by cataphoric shell nouns occurs most frequently in both AH (35.83%) and SS (39.45%) disciplinary domains, while attributive relational process accounts for the largest proportion of shell-noun use in NS (41.26%). These findings are in line with the findings about complement

types discussed in section 4.2.1. Finite *that*-clauses tend to be adopted by the AH and SS writers to convey their internal, mentally-oriented experiences associated with ideas and beliefs, in order to develop a new body of knowledge. In contrast, *to*-infinitival clauses, which embody varying degrees of modality and are associated with shell nouns that serve to construe the desirability of specific, established and definite empirical procedures, are frequent in the NS disciplinary domain. To further extend this analysis, consider [Table 4.22](#) (see Appendix 7), which summarizes the different types and sub-types of transitivity processes construed by distinct shell nouns across the three disciplinary domains. Starting with *attributive relational* process in NS first, we can see from the most frequent attributive relational process shell nouns in this disciplinary domain (Table 4.23 below) that half of these shell nouns are closely related to dynamic relational processes, which denote the experience of making judgments about the degree to which it is preferable that specific activities will occur under certain empirical conditions, such as *ability*, *tendency* and *capacity* (as in examples (4.24) through (4.26)).

- (4.24) It also has the **ability** to vaporise large involatile molecules that would decompose through vaporisation by heat. (NS, 0388E, L3)
- (4.25) ...there is often a **tendency** to have lots of activity on the interesting bits of code, while the more mundane areas are not done. (NS, 6101c, L3)
- (4.26) this evidence suggests that if monkeys can do the same (albeit with a third photopsin) and develop the **capacity** to use their extra photopsin, humans may be able to. (NS, 0014d, L3)

While it may not be surprising that shell nouns closely related to the dynamic

modality are the most frequent in a hard science discourse (as already noted in section 4.2.1), readers (e.g. EAP researcher and practitioners) familiar with academic writing may be surprised to find attitudinal shell nouns such as *problem*, *disadvantage* and *advantage* as the most frequent within the category (as in examples (4.27) through (4.29)). this finding challenges the notion of academic writing as a purely objective and impersonal discourse, and highlights the persuasive nature of scientific writing, which is often hidden behind a veneer of objectivity (Hunston and Thompson, 2000, p.177; Hyland, 2012, p.127).

(4.27) A **problem** with this method is that unless the patient is willing to let themselves be in their feared situation, it is very difficult to implement this method. (NS, 0017b, L1)

(4.28) An equally important **disadvantage** is that stage models are often inaccurate or simply wrong. (NS, 0011a, L1)

(4.29) The **advantage** of Scheme C is that it would reduce traffic congestion between the airport and the city, whereas the others would add to it. (NS, 3095e, L1)

Examples (4.27) through (4.29) illustrate how the hard knowledge field tends to use these nouns to denote quality, comment on the relative desirability or the downside of particular aspects of scientific methods. These evaluations offer a clear signal to the reader what a writer considers important or worth attending to, and can help to refute disagreements and gain support for their claims.

Table 4-23

Table 4.23: Most frequent cataphoric shell nouns in <i>attributive relational</i> process in NS			
Rank	Sub-types	Shell nouns	Norm. Freq.
1	Dynamic	ability	208
2	Deontic	need	115
3	Dynamic	tendency	73
4	Attitudinal	problem	52
5	Attitudinal	advantage	47
6	Dynamic	capacity	36
7	Dynamic	inability	31
8	Attitudinal	disadvantage	26
9	Dynamic	way	26
10	Epistemic	chance	21

If we go back to Table 4.23, it is further observed that the distribution of the most frequent attributive relational process shell nouns in NS follows a Zipfian's law (see section 4.1.1) (Zipf, 1936; 1949). For example, the most frequent shell noun, *ability* (208 per million words), has a frequency nearly double that of the next most frequent shell noun, *need* (115 per million words), and three times of the third-rank shell noun, *tendency* (73 per million words), and so on. The reason for the high frequency of *ability* is likely because it represents an extremely important module topic in this disciplinary domain. This is evident from the fact that *ability* most often occurs in the SN-*to*-infinitival clause syntactic pattern, highlighting topics that are closely related to the industrial apparatus (as in (4.30)), experimental skills/measurements (as in (4.31)), scientific observations (as in 4.32), implications of experimental results (as in 4.33) and requirements of practicality (as in 4.34). This is in consistent with the NS's preference for attitudinal shell nouns, such as *problem*, *advantage* (see examples (4.27) through (4.29)) and the distribution feature of its grammatical patterns (see section 4.2.1).

- (4.30) Work by John Fenn and colleagues developed ES for commercial use 16 years ago, and it has increased in popularity ever since due to its **ability to ionise a wide range of compounds**. (NS, 0388e, L3)
- (4.31) Their real measure is in the **ability to function in a manner acceptable to himself and the group of which he is part**. (NS, 3064e, L1)
- (4.32) Instead of the universe consisting of tiny, point-like particles, it consists of strings. A point-like particle has only the **ability to move in different directions**. (NS, 6097d, L1)
- (4.33) The relevance of this shows how females in a predominantly dichromatic species can develop the **ability to enhance their colour vision by developing, and critically using, an extra photopsin**. (NS, 0014d, L3)
- (4.34) The **ability to disable personal password changes at the workstation** had to be represented in the design of the new system. (NS, 6108g, L3)

In short, the higher relative frequency of the *attributive relational* process in NS can mainly be attributed to the predominance of dynamic shell nouns, such as *ability*. Additionally,, another feature associated with the preference for the *attributive relational* process is the prevalence of attitudinal shell nouns, such as *problem*, *advantage*.

Having dealt with the *attributive relational* process that stands out in the hard field, let us take a look at the *mental* process in soft field. Table 4.24 displays the most frequently used cataphoric shell nouns in the *mental* process for two soft knowledge disciplinary domains.

Table 4-24

Table 4.24: Most frequent cataphoric *mental* shell nouns in AH and SS

AH				SS			
Rank	Sub-types	Shell nouns	Norm. Freq.	Rank	Sub-types	Shell nouns	Norm. Freq.
1	Creditive	idea	152	1	Creditive	idea	150
2	Creditive	belief	99	2	Creditive	belief	85
3	Volitional	desire	79	3	Creditive	view	75
4	Conceptual	notion	68	4	Conceptual	notion	57
5	Creditive	view	62	5	Creditive	assumption	54
6	Evidential	<u>evidence</u>	39	6	Volitional	will	47
7	Creditive	assumption	35	7	Volitional	desire	47
8	Creditive	sense	23	8	Volitional	intention	44
9	Creditive	impression	20	9	Volitional	aim	44
10	Volitional	purpose	20	10	Volitional	purpose	28

We need to examine each of the mental processes across two disciplinary domains to identify any considerable divergences from the sub-types of *mental* process. One sub-type (evidential) that stands out is evidential, with 39 occurrences per million words for the shell noun instance *evidence* in AH, compared to no occurrences at all in the SS corpus. Apart from this, we can say that the general picture of the *mental* process is fairly consistent across these two disciplinary domains. How, now, do we account for the higher relative frequency of the evidence in AH? One way to find an answer to this question is to compare it with the other disciplinary domain discussed in this research, NS. Table 4.25 below illustrates this comparison:

Table 4-25

Table 4.25: Most frequent cataphoric shell nouns in mental process in AH and NS

AH				NS			
Rank	Sub-types	Shell nouns	Norm. Freq.	Rank	Sub-types	Shell nouns	Norm. Freq.
1	Creditive	idea	152	1	Evidential	<u>evidence</u>	68
2	Creditive	belief	99	2	Creditive	view	57
3	Volitional	desire	79	3	Volitional	aim	52
4	Conceptual	notion	68	4	Creditive	assumption	42
5	Creditive	view	62	5	Creditive	belief	36
6	Evidential	<u>evidence</u>	39	6	Creditive	theory	36
7	Creditive	assumption	35	7	Creditive	notion	31
8	Creditive	sense	23	8	Volitional	purpose	26
9	Creditive	impression	20	9	Creditive	idea	21
10	Volitional	purpose	20	10	Volitional	desire	21

Interestingly, AH seems to have more in common with NS than with SS regarding the sub-types of mental process. This is because the shell noun *evidence* is the most frequent within the *mental* process in the NS corpus (68 per million words) and appears in sixth place in AH as well (39 per million words), but doesn't appear in the top ten most frequent *mental* shell nouns in SS. In short, AH shares more with NS (8 shell nouns in common) than with SS (7 shell nouns in common). If we compare the rank sequence of sub-types for the *mental* process in AH with those in the NS and the SS, however, AH nevertheless has more in common with SS than NS. From this perspective, it seems that while AH shares with NS a stronger emphasis on the evidential sub-type of *mental* process, the actual shell-noun uses (modes of knowing) in AH are actually closer to those in SS.

4.3.2. Distribution of types of construal instantiated by anaphoric shell nouns

We will now consider the frequency data for anaphoric shell nouns. Since the *th-be-SN* syntactic pattern is completely absent from the NS and SS corpora, this analysis mainly focuses on the *th-SN* syntactic pattern. Table 4.26 shows a breakdown of the distributions of the different types and sub-types of transitivity process across the three corpora. Furthermore, [Table 4.27](#) (see Appendix 8) shows the distributions of anaphoric shell nouns according to the major divisions of the three disciplinary domains and different types of transitivity process, but together this time along with distinct anaphoric shell-noun items and their frequency.

Table 4-26

Table 4.26: Distributions of different types of transitivity process construed by anaphoric shell nouns across three disciplinary domains. (*freq.* Rounded figures normalized per million words, *prop%* based on raw frequency)

Types of transitivity process	Sub-types of transitivity processes	AH			NS			SS		
		Freq.	Prop. %	Type	Freq.	Prop. %	Type	Freq.	Prop. %	Type
Attributive relational process	Comparative	13	3.34	5	10	1.01	2	21	2.41	2
	Deontic	12	3.01	6	21	2.02	2	34	3.92	7
	Attitudinal	11	2.68	4	73	7.07	3	31	3.61	3
	Epistemic	5	1.34	4	5	0.51	1	10	1.20	3
	Dynamic	3	0.67	2	31	3.03	3	26	3.01	5
	Sub-total	44	11.04	141	27	13.64	11	121	14.16	20
Identifying relational process	Neutral	4	1.00	5	4	2.02	2	26	3.01	2
	Partitive	12	3.01	21	8	4.04	4	47	5.42	6
	Sub total	16	4.01	26	12	6.06	6	72	8.43	8
Mental process	Conceptual	105	26.09	73	65	32.83	15	199	23.19	18
	Creditive	70	20.74	130	23	13.64	11	189	25.00	25
	Volitional	28	7.02	73	9	4.55	6	47	5.42	12
	Emotive	4	1.00	16	2	1.01	1	13	1.51	3
	Evidential	1	0.33	5	13	6.57	3	10	1.20	2
	Dubitative	0	0.00	0	0	0.00	0	0	0.00	0
	Sub total	208	55.18	297	112	58.59	36	458	56.33	60

Verbal process	Assertive	64	16.05	78	16	8.08	10	90	10.54	7
	Rogative	13	3.34	5	4	2.02	1	26	3.01	1
	Directive	11	2.68	31	4	2.02	3	13	1.51	3
	Propositional	3	0.67	10	12	6.06	2	21	2.41	2
	Commissive	3	0.67	5	0	0.00	0	3	0.30	1
	Expressive	0	0	0	0	0.00	0	0	0.00	0
	Sub total	94	23.41	130	36	18.18	16	153	17.77	14
Material process	General	28	7.02	26	11	5.56	5	47	5.42	10
	Specific	7	1.67	26	0	0.00	0	3	0.30	1
	Attitudinal	4	1.00	16	0	0.00	0	5	0.60	1
	Sub total	39	9.70	68	11	5.56	5	54	6.33	12
Total	401	100	630	198	100	74	858	100	114	

As table 4.26 shows, the distribution of anaphoric shell nouns in different types and sub-types of the transitivity process across the three corpora is quite different from that of cataphoric shell nouns, as it shows a general similarity in the preference for *mental* process. Table 4.26 demonstrates that the types of transitivity process are, for the most part, consistent across the major disciplinary divisions, but with one notable exception. If you read the sub-types of *mental* process in Table 4.28 from top to bottom, starting with *conceptual* and ending with *dubitative*, you can see that the frequency for each sub-type is on a similar declining trend, with *conceptual* being the most frequent and *dubitative* being the least frequent, and with the intervening subtypes in the same rank order of frequency. There is, however, one exception, *evidential*, for the NS corpus, which has a higher frequency than the norm and would be in a different position if the mental process in Table 4.26 was strictly ordered by rank; it is the third most frequent subtype, not the penultimate one as suggested by its position in the table. Let us now turn to a more quantitative reading. Table 4.28 below shows the frequency of the top ten *mental* anaphoric shell nouns for the overall corpus, split across the AH, NS and SS.

Table 4-28

Table 4.28: Most frequent shell nouns in <i>mental</i> process across the three corpora						
Rank	AH		NS		SS	
	Shell nouns	Norm. Freq.	Shell nouns	Norm. Freq.	Shell nouns	Norm. Freq.
1	idea	21	<u>evidence</u>	57	view	46
2	point	20	idea	52	idea	39
3	view	17	view	47	point	36
4	theory	13	theory	36	theory	36
5	concept	11	Knowledge	31	issue	26
6	issue	11	concept	21	notion	21
7	interpretation	11	issue	21	perspective	18
8	notion	7	<u>purpose</u>	16	reasoning	13
9	image	5	hypothesis	16	assumption	10
10	hypothesis	5	point	10	concept	10

Here, we need again to compare the relative frequency of the subtypes of *mental* process for each disciplinary domain. There is not much variation around the overall distribution trend of subtypes, with a large proportion of conceptual and creditive shell nouns, except for NS, where the evidential shell noun *evidence* (underlined) is substantially frequent (as already noted in section 4.3.1). Additionally, while the volitional shell noun *purpose* (underlined) does not appear at all in this list in AH or SS corpora, it stands out as the most frequent shell noun for NS with a relatively high frequency of 16 per million words, ranking as the eighth most frequent. How do we explain these two discrepancies? Let us take the two shell nouns in question for NS one by one. Starting with the evidential shell noun *evidence*, my qualitative reading of the corpus suggests that, as researchers (e.g. Hyland, 2008; Flowerdew and Forest, 2015) have previously argued, texts in natural science are more concerned with presenting evidence and identifying the basis for an assertion in an argument than exploring relationships in arguments or discussing abstract ideas. This can be observed in examples (4.35) through (4.37) below:

- (4.35) Furthermore, over expression of superoxide dismutase in Drosophila also extends life span. This **evidence** supports the theory that free radical damage is a significant determinant of ageing. (NS, 6012e, L3)
- (4.36) It is estimated that most of the deuterium in the atmosphere was produced within the first two minutes after the big bang. This is new **evidence** shown in an article by Lee Siegel in the year 2000 and this shows that more evidence for the big bang is still being found. (NS, 6094b, L1)
- (4.37) Although Cohn et al (1989) found that this can limit the female's colour vision in some respects, it has been shown that the woman can compensate for this by scanning. This **evidence** can be linked... (NS, 0014d, L3)

Now, let's consider the higher frequency of volitional shell noun *purpose* in NS. This may be because, as we mentioned earlier in our discussion of Table 4.22, in NS, the experimental measurement the writer employed plays an important role (as in (4.38)). Besides, the usefulness of experimental apparatus is also a matter of concern to the writers in the NS domain (as in (4.39)):

- (4.38) I have not been able to find any measures of the severity of a person's alcoholism whilst on my current placement however, Moos et al (1990) cite the addiction severity index and Brief drinker profile as measurement tools which can be used for this **purpose**. (NS, 3032h, L1)
- (4.39) If mufti were to be washed to the same extent could it with stand the numerous washes at sixty degrees combined with tumble drying (Oxtoby, 2003) when it is not designed for this **purpose**. (NS, 3034d, L1)

Moving on to the two soft knowledge fields, although the greatest affinity is between broadly cognate fields, AH and SS share seven out of ten items, however, the frequency rate of each item in SS is double that of the corresponding in AH. For example, the most frequent shell noun *view* (46 per million words) in SS occurs at a

frequency of around double that of the most frequent shell noun *idea* (21 per million words) in AH. Similarly, the second most frequent shell noun *idea* in SS occurs twice as often (39 per million words) as the corresponding item *point* (20 per million words) in AH. Therefore, while AH shares with SS a similar emphasis on the subtypes of *mental* process, SS is nevertheless more concerned with discursively exploring theories, ideas and so on.

4.3.3. Changes of types of construal instantiated by shell-noun uses

Table 4.29 below outlines the change of types of transitivity process by levels of study. Furthermore, [Table 4.30](#) (see Appendix 9), [Table4.31](#) (see Appendix 10) and [Table 4.32](#) (see Appendix 11) summarize the different types of transitivity process construed by the distinct shell nouns across levels of study in each of the three disciplinary domains, respectively. As shown in Table 4.28 , for three of the five patterns of transitivity process, the normalized frequencies are increasing. The frequency numbers only decrease for *identifying relational* and *mental* processes. However, these decreases suggest the two most significant changes of all the types of transitivity process. *Identifying relational* process decreased significantly by approximately 24% due to a significant decline in the use of neutral factual nouns (30%), while the 15% decrease in *mental* process is rooted in a remarkable decline in the use of creditive shell nouns (41%).

Table 4-29

Table 4.29: Changes of types of transitivity process across level of study (rounded figures normalized per million words, *LLV* based on raw frequencies, Signif. Codes: '****' 0.001; '**' 0.01; '*' 0.05)

		L1	L3	LLV
Attributive relational process	Comparative factual	18	36	-4.08*
	Attitudinal factual	63	102	-6.09*
	Epistemic modal	52	54	-0.04
	Deontic modal	217	200	0.46
	Dynamic modal	408	483	-4.16*
Subtotal		758	875	-5.57*
Identifying relational process	Neutral factual	425	297	14.90***
	Partitive factual	46	59	-1.10
Subtotal		470	356	10.45**
Mental process	Evidential	124	113	0.33
	Conceptual	349	207	24.14***
	Creditive	601	619	-0.17
	Dubitative	24	12	2.34
	Emotive	19	37	-3.88*
	Volitional	355	302	2.87
Subtotal		1,473	1,291	7.87**
Verbal process	Propositional	22	20	0.06
	Assertive	308	374	-4.25**
	Rogative	38	12	9.17**
	Directive	25	30	1.20
	Commissive	28	19	1.27
	Expressive	0	0	-
Subtotal		422	455	-0.85
Material process	General	50	33	2.50
	Specific	223	271	-3.14
	Attitudinal	53	76	-2.69
Subtotal		326	378	-2.55
Total		3,474	3,421	0.27

Let's first consider the two types of transitivity process where the most significant changes occurred one by one. Table 4.33 shows the frequency of the top ten *identifying relational* shell nouns across the levels of study.

Table 4-33

Table 4.33: Most frequent *identifying relational* shell nouns across the levels of study (rounded figures based on normalized frequencies)

Shell nouns	L1	L3
fact	1,062	891
phenomenon	46	3
feature	36	6
<u>example</u>	28	67
<u>factor</u>	28	40
<u>aspect</u>	20	34
<u>basis</u>	13	39

exception	13	0
thing	11	9
function	7	6

From Table 4.33, we can immediately see that the shell noun *fact* decreased significantly in the L3 text. Thus, the decrease in *identifying relational* process can primarily be attributed to the L3 texts' lesser emphasis on the use of the shell noun *fact*. The underlining in Table 4.33 indicates shell nouns that stand out, as they show an opposite increasing trend in compared to the overall decreasing trend across the levels of study. While all these four shell nouns are members of the *partitive relational* process, there is a slight difference in the relational nature of these nouns, or more precisely, the linguistic effects of their relational nature. The nouns *factor*, *aspect*, *basis* bring the semantic component that the highlighted relation is general 'part-whole' (Schmid, 2000, p.118). In contrast, *example* conveys the additional meaning that the 'whole' is a class or set of entities, and the 'part' is a member of the class. In this case, the 'class-member' relation. If we compare these two categories of partitive shell nouns, it is quite interesting that the increase of *example* (increased by 140%) exceeds that of the overall increase of *factor*, *aspect* and *basis* (increased by 85%). This may be because higher-level students deal with issues that are more subject to contextual caprice than those studied in their primary level, and thus writers frequently add a 'descriptive gloss' to their accounts (Halliday and Matthiessen 2014, p.464). As the extracts below show, these propositional elaborations may either exemplify the argumentative grounds of the discussion (as in (4.40)) or specify an extending interpretation (as in (4.41)) (see section 5.1.2):

- (4.38) Modern Arctic explorers have developed similar abnormalities by consuming the livers of large carnivores such as polar bears. In addition to strengthening the case for the increasing importance of animal protein within the diet, this **example** is also significant for two further reasons. (NS, 3016b, L3)
- (4.39) It was only with the imposition of power constraints from 1988, and the exogenous variable of the need for reform regarding Eastward enlargement, that the isolation of the CAP became restricted. The inference drawn from this **example** highlights the potential danger that policies which are implemented may remain functioning outside the arena of political scrutiny,...(SS, 0399e, L3)

Turning now to the mental process, Table 4.34 shows the frequency of the top ten *mental* shell nouns across levels of study.

Table 4-34

Table 4.34: Most frequent *mental* shell nouns across the levels of study (rounded figures based on normalized frequencies)

Shell nouns	L1	L3
idea	606	396
view	274	232
theory	227	102
belief	217	91
desire	189	113
<u>assumption</u>	154	202
<u>notion</u>	143	138
<u>evidence</u>	131	344
purpose	113	47
decision	107	90

Table 4.34 shows that the general picture, a trend of decrease, is fairly consistent across the levels of study, with decreases in nearly each of the ten most frequent shell nouns in mental process. However, there are two outliers that show a different changing trend: 131 per million words for *evidence* in L1 texts as opposed to 344 per million words for the L3 texts (an increase of a massive 163%); and 154 per million words for *assumption* in L1 texts as opposed to 202 per million words in L3 texts (an

increase of 32%). First, taking the most striking feature, the significantly greater use of *evidence* in L3 texts, my qualitative reading of the corpus suggests to us that academic writing in higher levels, especially in soft fields, is increasingly concerned with expressing something of a scientific ideology that emphasizes empirical evidence over interpretative ideas. These extracts (4.40) through (4.42) give some flavour of this:

- (4.40) Non-verbal stimuli have therefore also been studied with **evidence** that children with SLI show poorer performance on tactile perception tasks. (AH, 6206e, L3)
- (4.41) There are at least five terms in Thucydides' description that only occur in the Hippocratic corpus. Although this **evidence** is useful in showing how Thucydides was aware of the Hippocratic writings it does not show whether he related his theories to them. (AH, 6109a, L3)
- (4.42) The Monticchio core of southern Italy produced a variable record of tree, shrub and vine pollen to herb pollen. This herb pollen identifies a colder open landscape with little tree cover in northern Europe (ibid). This **evidence** is invaluable ...(AH, 6171e, L3)

How do we account for the higher relative frequency of the creditive shell noun *assumption* in this general picture of decrease? One way to find an answer is to compare it with other related creditive shell nouns, such as *belief*. Both of these shell nouns belong to the ‘creditve-belief’ family, as they have the potential to evoke a cognitive representation of psychological states of knowing and believing. However, *belief* carries the semantic component that the highlighted feature is ‘the feeling that something is definitely true or definitely exists’ (LDCE), while *assumption* conveys the meaning that ‘something that you think is true although you have no definite proof’ (LDCE). Therefore, the relative higher frequency of *assumption* is likely due to the

fact that the legitimate routes to knowledge-making and understanding in higher levels of study largely lie in ‘interpretation, speculation, complexity and theoretical reasoning’, rather than expressing ‘codified beliefs’ (Bazerman, 1988, p.126), which value the representation of personal beliefs and perceptual reasoning. The examples below illustrate this difference, with examples such as (4.43) and (4.44) referring to ‘codified beliefs’, which are increasingly replaced by those that refer to the assumptions in which they are scientifically proposed (4.45), utilized/supported (4.46) and refuted (4.47).

**L1
texts**

(4.43) The articles in the Declaration of the Rights of Man, which expounds on the freedom of Man, may have led to the **belief** that it was the Enlightenment that was the ideological inspiration behind the revolution. (AH, 0019c, L1).

(4.44) Initially, it will be shown that the adventure story of late nineteenth century England provided young Englishmen with the **belief** that it was their right to propagate Imperialism because of their natural superiority to other ethnicities. (SS, 0004a, L1)

(4.45) The only **assumption** made on U and V was that they had no singular values. (NS, 0323b, L3)

**L3
texts**

(4.46) This step was guided by our **assumption** that the values are related and by a simple relation as well. (AH, 0311k, L3)

(4.47) Secondly, the **assumption** that there are no barriers to trade such as tariffs and that there are no transport costs are unrealistic. (SS, 0187b, L3)

Regarding the types of transitivity process that show an increased normalized value, of the three increased types, only the *attributive relational* process exhibits a statistical significance. Both the *attributive relational* and *material* processes increased by a similar extent of approximately 15% respectively, while the verbal process shows a marginal increase of nearly 8%.

The *attributive relational* process, as defined in this thesis, comprises various sub-categories. It is chiefly associated with the modal class, which attributes different types of epistemic, deontic, and dynamic modality to the writer's propositions. In addition, it is partly linked to the sub-classes of factual shell nouns, which attribute comparative features and different attitudinal evaluations to general facts and states of affairs. Table 4.35 shows the frequency of the top ten *attributive relational* shell nouns across levels of study.

Table 4-35

Table 4.35: Most frequent *attributive relational* shell nouns across levels of study (rounded figures based on normalized frequencies)

Shell nouns	L1	L3
ability	595	625
need	332	279
right	150	103
tendency	135	133
opportunity	123	159
<u>problem</u>	118	293
potential	103	66
possibility	68	95
<u>advantage</u>	37	99
<u>difference</u>	27	74

The shell nouns *problem*, *advantage* and *difference* are underlined because the most significant changes lies in the use of these three items. It can be seen that the attitudinal shell noun *problem* (293 per million words) is the most frequent in L3 texts, at a frequency of over double that of the same item in L1 texts (118 per million words). Similarly, the ninth shell noun *advantage* (attitudinal relational item) is used substantially more by writers in the third level (99 per million words), with a frequency of nearly three times that of 37 per million words in L1 texts. In addition to using more attitudinal shell nouns, writers in the third level also make nearly three

times more use of comparative relational shell noun *difference*, amounting to 74 per million words compared with 27 per million words. Once again, these three shell nouns can probably explain how *attributive relational* process has increased significantly in the third level of study. This quantitative observation, along with my qualitative reading of the corpus, indicates that writers in the higher level of study (Level 3) are perhaps more concerned with evaluating investigated issues and exploring the relations between them, rather than merely using abstract nominals to define what they are. This may be observed in the following four examples: the first two examples (4.48) and (4.49) refer to the writers' evaluations of the issue, and the latter two examples (4.50) and (4.51) refer to the comparative relations between two issues/methodologies.

- (4.48) The **problem** is that these benefits probably do not economically outweigh the hefty investment that would be required in order to bring the new members into line with existing EU law. (SS, 0244m, L3)
- (4.49) Contradictory to the idea of equality which founded the Bretton Woods system, this development highlighted an emergent asymmetry in the international economy. As an economically dominant power, the US utilised this **advantage** and took on a new role as an international hegemonic power. (SS, 0399c, L3)
- (4.50) The only **difference** is that the upper jaw of the bear aligns closer to the bottom jaw, giving more of a chewing mechanism than the dog... (NS,6181d, L3)
- (4.51) Thus we can see how the Satyricon is an example of how cinema can reconstruct antiquity in different ways depending on the filmmaker's use of sources, just as literary history can argue different viewpoints by interpreting sources in different manners. This **similarity** in methodology can therefore be used to argue that.... (AH, 6053h, L3)

Next, the increase in *material* process may reflect a trend where students in the

third level are increasingly taking the opportunity to construct knowledge based on their experiences of the neutral existence of physical observable facts and empirical results, rather than cognitive activities or psychological-states. Turning now to a more quantitative reading, Table 4.36 shows the frequency of the top ten *material* shell nouns across the levels of study.

Table 4-36

Table 4.36: Most frequent *material* shell nouns across levels of study (rounded figures based on normalized frequencies)

Shell nouns	L1	L3
attempt	431	495
failure	101	165
effort	72	118
change	44	15
<u>struggle</u>	35	3
act	34	24
action	20	15
choice	16	10
option	11	20
move	6	15

Based on the overall increasing trend, we would anticipate increases for some typical material shell nouns such as *attempt*. Comparing these two figures, we can see that, consistent with what we saw for the overall trend (see Table 4.28), the most significant increases occur in the frequencies of the top three most frequent shell nouns, especially *failure* and *effort*, which are the second (101 per million words) and third (72 per million words) most frequent shell nouns in L1 texts respectively, and are significantly increased by more than 60% in L3 texts. Examples such as (4.52) through (4.54) suggest that the increase in these *material* nouns may reflect a broad tendency for students in the third level to take an experimental angle and make greater emphasis on a research-based approach to the construction of new knowledge.

- (4.52) In an **attempt** to reduce percentage of crop raiding caused by baboons living in the forest, the non-native 'Zairois,' from neighbouring Zaire, are given land closest to the forest. (SS, 3001h, L3)
- (4.53) The second component concentrates on implementing a new approach towards their child in an **effort** to change the child's learned responses. (NS, 0421a, L3)
- (4.54) Indeed if the discount rate is increased further to 8%, the gains are in fact reduced to zero. As we can see, this **failure** occurred in an industry where competition did not operate at its most efficient level. (SS, 0202a, L3)

However, against the overall trend, the fifth most frequent shell noun *struggle* (underlined) is markedly decreased by 91% (35 per million words for L1 texts vs. 3 per million words for the L3 texts). The reason why this shell noun is almost absent from the third-year's writing may be that students in the third level are less concerned with discussing subjective attitudinal nature of research topics or event descriptions, as compared to other related attitudinal eventive shell nouns such as *failure* (as in example (4.54)). This assumption is supported by the observation that in first-year texts, the attitudinal shell noun *struggle* often occurs in a shell-noun pattern (SN-*to*-clause) pre-modified by possessive determiners, which functions somewhat as common collocations, as in examples (4.55) and (4.56) below:

- (4.55) And discussing Dowell's **struggle** to narrate properly she offers the point,.. (AH, 3158a, L1)
- (4.56) In their **struggle** to preserve the dynamics of nature, 'the ecocentric very much restates the ideas of the transcendentalists. (AH, 0129e, L1)

Moving now to the last of the three increased transitivity processes, *verbal*

process, which shows a slight rise with no statistical level of importance. In fact, if we go back to Table 4.28, we could already have identified that this slight rise is mainly due to the significant increase in assertive shell-noun uses. In contrast, rogative shell nouns decreased markedly. Table 4.37 shows the frequency of the top ten *verbal* shell nouns across the levels of study.

Table 4-37

Table 4.37: Most frequent *verbal* shell nouns across the levels of study (rounded figures based on normalized frequencies)

Shell nouns	L1	L3
argument	240	115
conclusion	112	78
question	112	54
claim	109	130
statement	96	72
criticism	73	25
<u>suggestion</u>	27	135
assertion	10	25
comment	7	12
proposition	7	12

The shell noun *suggestion* is underlined because it shows the largest change across levels of study, increasing by a striking 400% and seemingly playing an important role in shaping academic arguments in the third level. Interestingly, with regard to the shell-noun use of *suggestion*, it seems that students in the first-year are more concerned with the directive aspect of advising (semantic feature: linguistic-directive) while the third-year students attribute more importance to the assertive component (semantic feature: linguistic-assertive/tentative) based on empirical research. Examples (4.57) and (4.58) refer to directive speech acts, which are increasingly being replaced by those that refer to assertions based on evidence, as in examples (4.59) and (4.60).

- L1 texts**
- (4.57) 'Mr Brown is looking into ways of freeing up the planning system to boost house building, as well as encouraging people to make more use of fixed rate mortgages.' (John Hawksworth, 2001) I think this **suggestion** is just one of the things that needs to be carried out in order to prepare the UK for membership. (NS, 3109c, L1)
- (4.58) As a means of emancipation the ideal speech situation does appear to have some merit such as the idea of using culture to force through change. And if considered as a (very) long term solution, this **suggestion** does seem more practical than I originally thought. (SS, 0010e, L1)
- (4.59) This has led to the **suggestion** that SLI is not in fact a specific disorder of language but a more general disorder affecting cognition.... (AH, 6206e,L3)
- L3 texts**
- (4.60) This gives **suggestion** that bears too, might be omnivorous, but only when they revert back to their carnivorous instincts that the animal shows the maximum size of the species. (NS, 6181d, l3)

An analysis of the overall changes is just an entry point, and it turns out that these changes related to levels of study are not evenly distributed across the three disciplinary domains. Table 4.38 below shows the changes in types of construal across levels of study by disciplinary domain.

Table 4-38

Table 4.38: Changes of types of transitivity process across levels of study by disciplinary domains (normalized per million words, *LLV* based on raw frequencies, Signif. Codes: '****' 0.001; '***' 0.01; '**' 0.05)

Types of transitivity Process	Sub-types	AH			NS			SS		
		L1	L3	LLV	L1	L3	LLV	L1	L3	LLV
Attributive relational	Comparative factual	10	30	-4.15*	19	24	-0.12	38	48	-0.20
	Attitudinal factual	29	45	0.25	158	368	-8.27**	51	113	-4.40*
	Epistemic modal	41	27	0.30	56	71	-1.91	96	87	0.07
	Deontic modal	193	154	1.67	241	178	0.91	261	274	-0.06
	Dynamic modal	331	398	-2.26	501	794	-1.32	548	492	0.56
Subtotal		604	654	-0.73	1,003	1,363	-5.25*	993	1,014	-0.04
Identifying relational	Neutral factual	474	286	17.04***	186	308	-2.94	458	309	5.54*
	Partitive factual	29	45	-1.32	37	59	-0.50	96	78	0.32

subtotal		503	332	12.94 ***	223	367	-3.42	554	387	5.60*
Mental	Evidential	152	99	4.08*	111	356	-12.99 ***	13	61	-6.13 **
	Conceptual	276	199	4.55*	362	296	0.62	478	191	12.25
	Creditive	515	648	-5.65*	548	533	0.18	923	609	24.25 ***
	Dubitative	17	6	1.95	28	0	-	38	26	0.5
	Emotive	19	21	-0.03	28	36	-0.08	57	44	0.36
	Volitional	285	317	-0.60	334	202	3.12	554	318	12.32 ***
subtotal		1,264	1,290	-0.10	1,411	1,423	-0.01	2,063	1,244	38.79 ***
Verbal	Propositional	12	3	2.10	65	59	0.02	19	30	-0.48
	Assertive	297	313	-0.16	204	225	-0.10	407	374	0.26
	Rogative	36	9	6.30*	19	36	-5.20	57	9	7.87 **
	Directive	34	99	-12.67 ***	19	59	-2.17	6	61	-9.04 **
	Commissive	12	9	0.16	19	0	-	76	39	2.32
	Expressive	0	0	-	0	0	-	0	0	-
subtotal		392	434	-0.81	325	379	-0.40	567	513	0.49
Material	General	36	33	0.05	37	83	-1.72	96	13	13.95 ***
	Specific	225	344	-8.26 **	186	178	0.02	242	200	0.74
	Attitudinal	68	84	-0.68	9	24	-0.63	45	78	-1.72
Subtotal		329	461	-8.26 **	232	285	-0.50	382	292	2.29
Total		3,135	3,171	-0.07	3,194	3,817	-5.27*	4,558	3,451	28.95 ***

The normalized figures in Table 4.37 show that there is a moderate increase in the use of shell nouns in AH, though not statistically significant. However, a scrutiny of each type and sub-type of transitivity process reveals that writers in the AH disciplinary domain substantially reduced their shell-noun uses in the construal of *identifying relational* process ($LLV=12.94, p<0.001$). Furthermore, the shell-noun uses in the construals of *material* process have undergone a remarkable increase ($LLV=-8.26, p<0.01$). Evidential and conceptual shell-noun uses for *Mental* process displayed substantial decreases, while creditive use has recorded a significant increase. In addition, although the directive shell-noun uses under *verbal* process showed a

substantial increase, with normalized figures increasing by a massive 190 percent, it occurred with very low frequencies, which provide no analytical value.

In terms of NS, normalized figures of shell-noun uses show that there is a significant increase of 20% ($LLV=-5.27, p<0.05$). This thesis may confidently say that NS L3 writers have embraced shell nouns in their linguistic repertoires and are increasingly investing in them to enhance the accessibility of technicality of disciplinary knowledge and the persuasiveness of their academic arguments. It is further observed that evidential shell-noun uses under *mental* process records the most significant rise of 220% ($LLV= -12.99 p<0.001$), and the dynamic modal relational shell-noun use under *attributive relational* process shows a 58% rise ($LLV= -6.38, p<0.05$). The large increase in evidential shell-noun uses reflects the growing importance of empirical research ‘based on the evidence from the real world’ (Adams et al., 2007, p.28). The increase in dynamic modal shell-noun uses indicates that L3 writing in NS tends to use shell nouns to attribute various types of modality to potential occurrences under certain empirical conditions. This corresponds to the ways this disciplinary domain typically constructs knowledge. As Hyland (2008, p.15) points out, new knowledge in hard sciences is accepted based on empirical demonstrations and empirical outcomes.

Overall, the normalized figures in Table 4.37 show that there is a significant decrease of 24% in the uses of shell noun in SS overall ($LLV=28.95, p<0.001$). It is easily observed that the two cognate soft disciplinary domains tend to share the same changing pattern. L3 writing in SS substantially reduced its shell-noun uses in the

construals of *mental* process ($LLV=38.79$, $p<0.001$) and *identifying relational* process ($LLV=5.60$, $p<0.05$). This finding ties with the intrinsic nature of soft sciences on the study of these nouns, but a closer look at Table 4.37 adds nuance to this finding as SS diverge from AH.

It is to be noted that there is a remarkable decrease in all but one of the six sub-types of *mental* process construed by shell nouns in SS, while AH shows a significant increase of 26% in creditive shell-noun uses. A possible explanation of this difference is that L3 writing in AH has shifted its focus from projecting the abstract relations into mentally perceivable ideas to the construals of psychological status of believing and knowing. By contrast, SS L3 writing might have witnessed a relatively significant move away from conceptually-oriented knowledge construction.

In addition, under the *material* process, whilst AH L3 writing exhibits a 53% increase in the uses of specific eventive shell nouns, SS L3 writing shows a significant decrease of general eventive shell-noun uses, decreasing by 86%. This difference is a possibly due to AH L3 writing's increasing preference for characterizing specific event. By contrast, the need to generalize and simplify various social activities might be less urgent and necessary, as the complexity of social activities and diversity of research entities have perhaps increased in SS.

4.3.4. Summary: distributions of types of construal

This section analyzed the distributions of different types and sub-types of transitivity process construed by shell nouns across disciplinary domains and levels of

study. The results are summarized below.

The distribution of types of transitivity process construed by cataphoric shell nouns was strongly different between soft and hard fields. Hard science showed a strong preference for the *attributive relational* process, while the two soft fields favored *mental* process. This finding offers evidence that roughly substantiates the typical divide between the more empirically-oriented style of hard sciences writing and the more interpretive, discursive and cognitive style of soft science writing (e.g. Hyland and Jiang, 2019ab; Benitez-Castro, 2021; Omidian and Siyanova-Chanturia, 2021).

However, the distribution of types of construal instantiated by anaphoric shell nouns exhibits a similarity across the three corpora. The uniform popularity of *mental* process across the three corpora was mainly accounted for by conceptual and creditive shell-nouns, such as *idea* and *view*. *Mental* shell nouns colligating with *th-SN* syntactic pattern provide writers with a useful means of encapsulating information that is conceptually dependent on the writers' interpretations and understandings and incorporating it into the ongoing discourse. Although this finding is less straightforward to explain, it may be related to the more argumentative nature of the academic essay included in the three corpora, which expects student writers to demonstrate their critical thinking skills and their ability to present a well-reasoned academic arguments.

In terms of changes across the levels of study, conceptual shell-noun use subsumed in *mental* process and neutral factual shell-noun use subsumed in

identifying relational process show the largest decrease over the years, decreasing by 40% and 30%, respectively. By contrast, there is a rise in the normalized frequency in all but one of the five sub-categories in the *attributive relational* shell-noun use. Comparative and attitudinal factual shell-noun uses exhibit the largest rises with increases of 100% and 60%, respectively, and coming to play an important role in shaping academic arguments.

It is also important to note that these changes are not evenly distributed across the three disciplinary domains. The most prominent difference occurs between soft and hard sciences, as the two soft disciplinary domains share a similar changing pattern, while the hard disciplinary domain is increasingly featured by its pronounced technicality and positivism.

Although AH and SS show some similarities, the picture is not uniform. The first major difference is the change occurring in the creditive shell-noun uses. The significant increase of creditive shell-noun uses in AH indicates a possible shift from one sub-type of *mental* process to another, while the substantial decrease in all sub-types of *mental* process in SS suggests a more complete move away from construals of mental entities (see section 5.2.2.2).

The second major difference is the change occurring in the *material* process. The remarkable increase of one sub-type of material shell-noun uses in AH L3 writing might indicate a growing preference for characterizing specific events (see section 5.1.3). By contrast, the increasing reluctance to simplify and generalize complex social activities and events is reflected by a significant decrease in the general

eventive shell-noun uses in SS L3 writing (see section 5.2.1.3).

4.4. Distribution of the five types of transitivity process across the six SN lexico-grammatical patterns in three disciplinary domains.

As we have seen, the analysis of lexical, semantic, and grammatical features of shell nouns has offered evidence of differences across the three corpora, indicating a number of disciplinary variations, as expected. These differences deserve closer attention, and the section below attempts to probe further into the semantic and grammatical differences observed above. Figures 4.1, 4.2 and 4.3 present the distributions of types of transitivity process across six shell noun lexico-grammatical patterns in the three disciplinary domains.

Beginning with AH, Figure 4.1 shows that the predominant preference for the *mental* process (see section 4.3.1) is largely associated with the remarkably frequent use of finite **SN-that** complement construction (see section 4.2.1). This salient inclination for *mental* process could similarly explain the use of the *th-SN* and *th-be-SN* patterns. Furthermore, the distinctive inclination towards *identifying relational* and *verbal* processes could be linked to the use of **SN-that** and **SN-be-that** patterns.

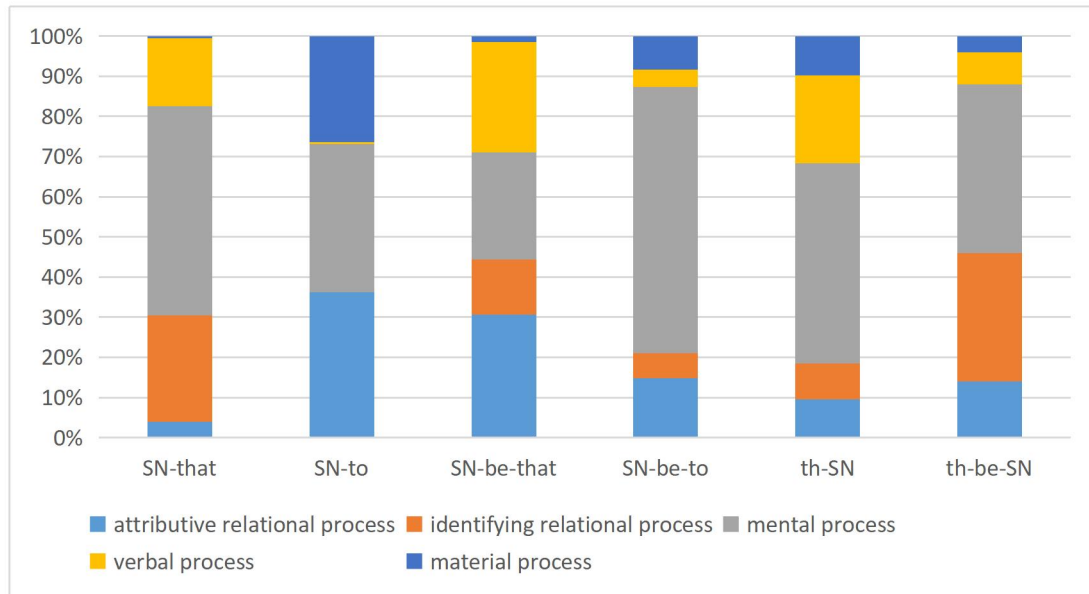


Figure 4.1: Distributions of five types of transitivity process across six SN lexico-grammatical patterns in AH.

Figure 4.2 below shows that, in NS, the primary preference for the *attributive relational* process (as described in section 4.3.1) is the main factor that explains the chief choice of the **SN-to-inf** (as discussed in section 4.2.1). Additionally, the secondary preference for the *mental* process is mostly associated with the comparatively smaller use of the **SN-that-cl** clause construction. Furthermore, the preference for the material process could also be linked to the conspicuous use of **SN-to-inf** construction. Finally, the primary preference for the *mental* process (as described in section 4.3.2) could be linked to the relatively greater use of the **th-SN** pattern (see section 4.2.2).

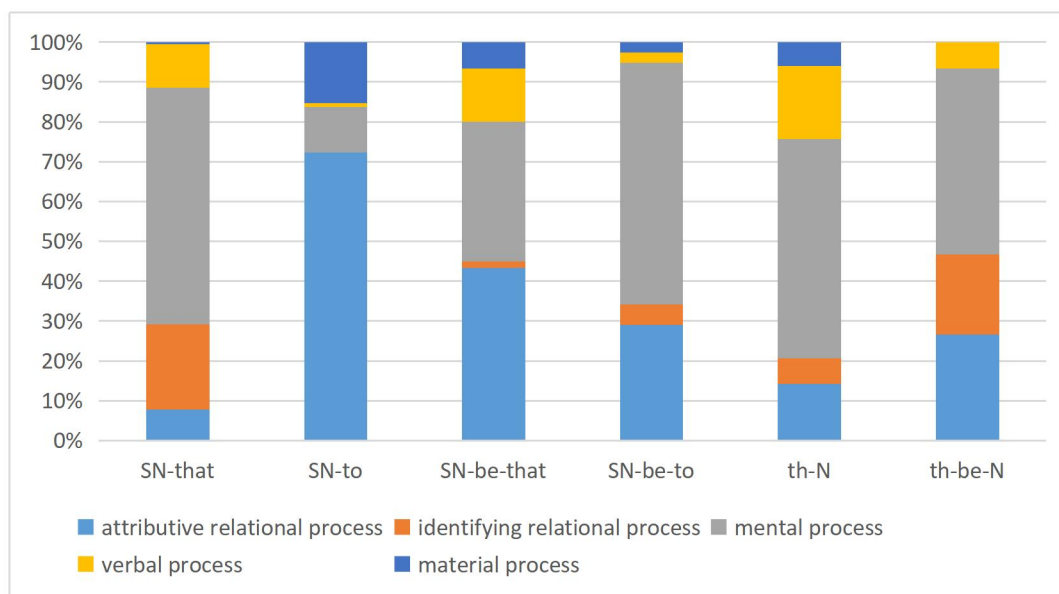


Figure 4.2: Distributions of five types of transitivity process across six SN lexico-grammatical patterns in NS.

Moving on to the SS corpus, it can be observed that, similar to the AH corpus, the primary preference for the *mental* process (as described in section 4.3.1) can largely be attributed to the greatest use of the *SN-that-cl* pattern (as discussed in section 4.2.1). The second primary inclination for the *attributive relational* process is mainly related to the use of the *SN-to-inf* construction, which, compared to the stronger association between the inclination for the *mental* process and the use of *SN-to-inf* in AH, appears to indicate a closer association to the use in NS. However, the additional preference for the *verbal* and *identifying relational* process is mainly related to the use of the *SN-that-cl* and *SN-be-that-cl* constructions, which, compared to the more conspicuous association between the proclivity to the *attributive relational* process and the use of *SN-be-that-cl* in NS, seems to imply a relationship that is closer to the use in AH. Once again, these mixed characteristics found in SS seem to reflect a writing style that is argumentative in nature, similar to AH, and certain empirical

dimensions of knowledge constructions in NS, making it a middle ground between AH and NS (as discussed in section 5.2.2.2).

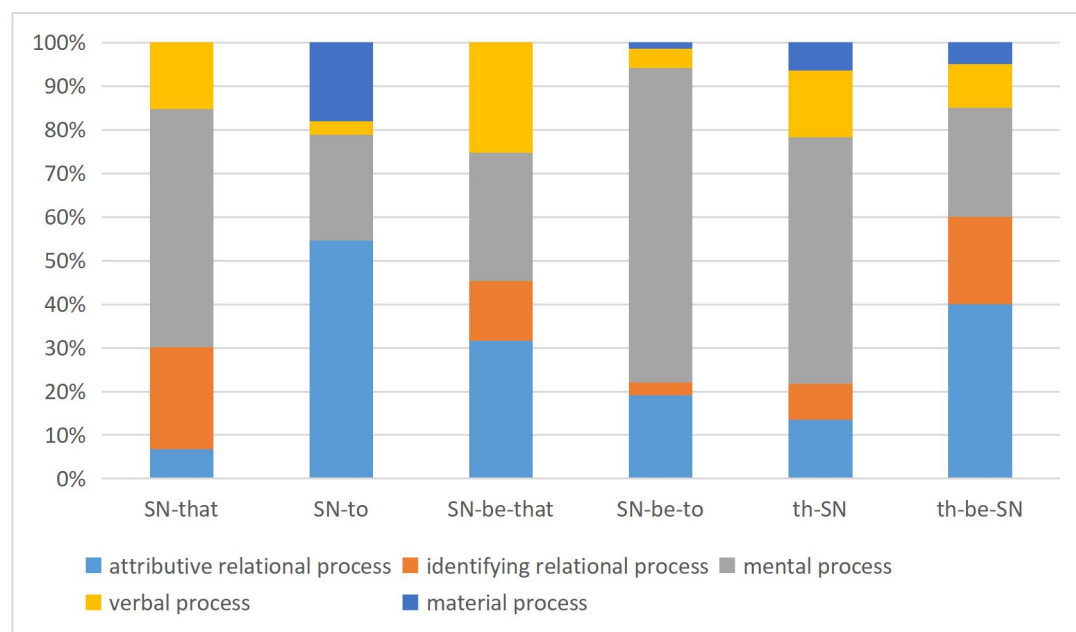


Figure 4.3: Distributions of five types of transitivity process across six lexico-grammatical patterns in SS.

4.5. Summary: chapter 4

Chapter 4 reveals that the quantitative analysis of shell-noun usage suggests disciplinary associations straddling the levels of lexis, syntax and semantics (question 1 and 2). Three disciplinary domains show a certain degree of similarity in the choice of shell nouns, indicating that while they differ from each other, they share broadly similar discourse characteristics in the use of shell nouns as meta-discursive devices. However, chapter 4 also identifies major differences in lexico-grammatical patterns between the three disciplinary domains. Specifically, AH and SS tend to favor the use of finite SN-*that*-clause complements with shell nouns, while NS primarily uses non-infinitival clause SN-*to*-clause complements. At a semantic cognitive level,

chapter 4 finds that soft and hard fields prefer different types of transitivity processes with shell nouns. *Attributive relational* process of dynamic modality are saliently favored in NS, while *mental* processes instantiated in creditive and volitional shell-noun uses are favored in AH and SS. This finding is in line with the differences in complement types previously discussed. Based on the epistemological orientations maintained by the three disciplinary domains, it can be argued that NS is more inclined to construe the experience of exploring the relations between entities, emphasizing practical activities in the real world, while AH and SS, or perhaps soft sciences broadly, are more inclined to construe the experience of characterizing the psychological status of believing and thinking.

However, a complete understanding of the profound influence of disciplinary concern cannot be achieved without conducting exhaustive analyses of the changes of shell-noun lexicalizations within each disciplinary domain. Therefore, the following section presents further analyses to address this issue.

Chapter 5 Changing Patterns of Shell-noun Use in Disciplinary Student Writing

This chapter addresses the third research question of this thesis, which is:

Question 3. What differences, if any, are evident in the uses of shell noun across levels of study within each disciplinary domain? Have these differences been consistent across the three disciplinary domains?

The distributions of shell-noun use reported in chapter 4 suggest a number of disciplinary associations that span the levels of lexis, syntax, and semantics (question 1 and 2). However, to fully understand the disciplinary associations of shell-noun use, it is necessary to conduct exhaustive analyses of the changing patterns of such use across levels of study within each disciplinary domain. To provide a detailed account of shell noun uses, this chapter will delve into the analysis of certain shell nouns' lexicalization and discuss their discourse features, as well as their functional workings in first-year writing in comparison to third-year writing within each of the three disciplinary domains. The focus of this chapter is on the qualitative results obtained by the analysis of examples.

The analysis is conducted by examining the surrounding context of a shell noun using the key word in context (KWIC) function of Sketch Engine, and an examination individual texts are examined when necessary. The functional work is analyzed by the

shell-noun referential and syntactic functions, the shell-noun phrase structure, and phraseology. The lexicalization is analyzed incorporating factors related to the meaning of the referent, as well as the English rhetorical patterns (e.g. clause relations and text patterns).

To make the analysis more manageable in size, this chapter's data analysis is limited to the shell nouns that occurred in the types of construal that show statistically significant differences in either of the three corpora ($p < 0.05$, see section 4.3.3). However, shell nouns without statistical significance will be included when they are closely relevant to the specific shell-noun use under discussion.

5.1. Changes in AH

According to Table 4.37 (see section 4.3.3), two types of construal instantiated by shell nouns show significant changes: *identifying relational* ($LLV=12.94$, $p < 0.001$) and *material* processes ($LLV=-8.26$, $p < 0.01$). Therefore, the following analysis will focus on the shell-noun uses for the construals of these two patterns of transitivity process.

5.1.1. *fact* in the construal of *identifying relational* process

The decline in the number of neutral factual shell-noun uses is chiefly accounted for by the overall decrease of identifying relational process. Neutral factual shell-noun uses construe human experiences related to general facts, states of affairs or aspects, features, such as *fact*, *thing*, *feature example* and so on. Specifically, shell noun *fact*

records the most significant decrease of nearly 40% ($LLV=19.8$, $p<0.001$) in AH L3 sub-corpus. Thus, the analysis of *identifying relational* process mainly focuses on *fact* and partly on *example*. Table 5.1 below outlines the changes in the syntactic patterns of the shell noun *fact* across levels of study in AH.

Table 5-1

Table 5.1: Changes of shell noun *fact* syntactic patterns (normalized per million words, LLV based on raw frequency)

	AH			
	L1	L3	LLV	p
SN- <i>that</i> post nominal	488	295	18.9***	<0.001
SN- <i>be</i> clause	0	9	-	-
<i>th</i> -SN	17	6	1.9	>0.05
<i>th-be</i> -SN	0	0	-	-

It can be seen from Table 5.1 that the SN-*that* post nominal pattern occurs most frequently, comprising around 95% of all shell-noun syntactic patterns in the two corpora, respectively. This finding corresponds to the results of previous studies that have found the combination *the fact-that*-clause to be the most frequent single colligation in written English, particularly in academic writing (e.g. Schmid, 2000, p.98; Jiang and Hyland, 2015, p.15). Since the use of *fact* predominantly occurs in the SN-*that* post nominal pattern, the analysis of its use primarily focuses on this type.

5.1.1.1. AH L1 writing

Out of 216 instances of *fact* in the AH L1 sub-corpus, over 90% (198 instances) occur in the phrase *the fact that*. The Phraseologies that include the three-word phrase are shown in Table 5.2 below:

Table 5-2

Table 5.2: Phraseologies including *the fact that* in AH L1 writing (raw frequency) (Search Term Position: On Left)

			AH L1		
Rank	Phraseology	Freq.	Rank	Phraseology	Freq.
1	<i>∅ The fact that</i>	54	23	<i>notwithstanding the fact that</i>	1
2	<i>to the fact that</i>	27	24	<i>lit up the fact that</i>	1
3	<i>by the fact that</i>	18	25	<i>know the fact that</i>	1
4	<i>despite the fact that</i>	16	26	<i>highlight the fact that</i>	1
5	<i>is the fact that</i>	11	27	<i>explain the fact that</i>	1
6	<i>and the fact that</i>	10	28	<i>emphasis the fact that</i>	1
7	<i>in the fact that</i>	5	29	<i>disguise the fact that</i>	1
8	<i>was the fact that</i>	4	30	<i>discuss the fact that</i>	1
9	<i>on the fact that</i>	4	31	<i>describe the fact that</i>	1
10	<i>from the fact that</i>	4	32	<i>deny the fact that</i>	1
11	<i>for the fact that</i>	3	33	<i>plus the fact that</i>	1
12	<i>of the fact that</i>	3	34	<i>disregards the fact that</i>	1
13	<i>ignore the fact that</i>	3	35	<i>denote the fact that</i>	1
14	<i>underline the fact that</i>	2	36	<i>could be the fact that</i>	1
15	<i>resent the fact that</i>	2	37	<i>cite the fact that</i>	1
16	<i>reflect the fact that</i>	2	38	<i>bemoan the fact that</i>	1
17	<i>given the fact that</i>	2	39	<i>belittle the fact that</i>	1
18	<i>but the fact that</i>	2	40	<i>being the fact that</i>	1
19	<i>with the fact that</i>	1	41	<i>appreciate the fact that</i>	1
20	<i>support the fact that</i>	1	42	<i>affect the fact that</i>	1
21	<i>stress the fact that</i>	1	43	<i>push the fact that</i>	1
22	<i>but the fact that</i>	1	44	<i>then the fact that</i>	1

The frequency information shows that three out of the four most frequent collocations include a preposition. Moving on to the semantic sequence of these collocations, the subsequent categorizations of the semantic sequence in Table 5.3 with *the fact that* are drawn from these collocations.

Table 5-3

Table 5.3: Semantic groups of *the fact that* identified in AH L1 writing (Search Term Position: On Left)

1. Fact is the basis for a practical outcome or reasoning
[entity] <i>due to the fact that</i>
[entity] <i>be strengthened by the fact that</i>
[entity] <i>be illustrated in the fact that</i>
[entity] <i>supports the fact that</i>
[entity] <i>explains the fact that</i>
[entity] <i>reflects the fact that</i>
[entity] <i>points to the fact that</i>
[entity] <i>denotes the fact that</i>
[entity] <i>could be the fact that</i>
[entity] <i>was the fact that</i>

be exemplified by the fact that
be proven by the fact that
be supported by the fact that
be seen in the fact that
be based on the fact that
be added by the fact that
be linked to the fact that
be down to the fact that
derives from the fact that
rely on the fact that

2. Fact explains something

[result/effect] for the fact that
[result/effect] is the fact that
by the fact that
be explained by the fact that
be implied by the fact that
be characterized by the fact that
be furthered by the fact that

3. Fact is the cause of a problem or its solution

[problem/difficulty] is raised by the fact that
[problem/answer] lies in the fact that
[problem] is shown in the fact that
[problem] stems from the fact that
the fact that
What constitutes [problem] is the fact that (all sentence initial)
be caused by the fact that

4. Something shows/uses/assumes a fact

[entity] makes apparent the fact that
[entity] lit up the fact that
[entity] this is not to deny the fact that (all sentence initial)
[entity] underline the fact that
[entity] be contradiction of the fact that
notwithstanding the fact that
being the fact that
and the fact that
but the fact that
disguise the fact that
despite the fact that
in addition to the fact that
in regard to the fact that
plus the fact that
given the fact that

5. Be aware or unaware of a fact

[people] highlights the fact that
[people] emphasis the fact that
[people] underlines the fact that
[people] knows the fact that

6. People discuss a fact

[people] stress the fact that
[people] discuss the fact that
[people] comments on the fact that
[people] describe the fact that
[people] cite the fact that

7. Affective reaction to a fact

[people] *appreciate the fact that*
 [people] *are oblivious of the fact that*
 [people] *resent the fact that*
 [people] *ignore the fact that*
 [people] *push the fact that*
 [people] *bemoan the fact that*
 [people] *be contented with the fact that*
 [people] *disregard the fact that*

To make these sequences more intelligible and manageable, this thesis draws on the umbrella term ‘semantic motifs’ from Groom (2007) to divide them into three broad areas. In brief, a semantic motif is ‘a group of semantic sequences or other phraseological items which can be grouped together insofar as they share a similar broad meaning’ (Groom, 2007, p.102). The three semantic motifs are presented in Table 5.4 below:

Table 5-4

Table 5.4: Three semantic motifs of <i>the fact that</i> in AH L1 writing		
Area	Description	Phraseology group
1. The ‘cause/effect’ motif	Facts are the cause/effect of things-outcomes, reasons, problems and solutions.	Group 1, 2, 3
2. The ‘orientation’ motif	Facts are either ignored or taken into consideration, Things are oriented around facts	Group 4
3. The ‘human response’ motif	Facts are discussed, responded by human	Group 5, 6,7

Specifically, the ‘cause’ motif of *the fact that* indicates that the fact is a cause; the ‘orientation motif’ of *the fact that* implies that the fact builds up a setting or that the lack of the fact is part of the orientation of the happening and consideration of other things; the ‘human response’ motif of *the fact that* is often used to indicate the human response and reaction to a fact. Thus, most instances of this ‘human response’ motif indicate the writer’s judgement of a fact in terms of its significance or validity, etc.

Additionally, with regard to the instances where these three-word shell noun phrases occur in the clause initial position in the form of *the fact that* (with no preposition preceding the shell-noun phrase), this thesis identifies their relations with the three motifs outlined above by examining the verb group that takes *The fact that* as the subject. Examples of these occurrences are illustrated in Table 5.5 below:

Table 5-5

Table 5.5: Three motifs of clause-initial *The fact that* in AH L1 writing

The cause motif	(5.1)	The fact that <u>people especially the employee at the factory see her as naïve</u> leads to her destruction. (AH, 0061c, L1)
	(5.2)	The fact that <u>he has inherited a mortal illness - syphilis - from his father</u> makes him comparable to Nora whose social determinism is the fault of her own father.(AH, 0024a, L1)
The orientation motif	(5.3)	The fact that <u>her well-intentioned action in forging her father's signature in order to save her husband's life</u> is illegal reflects the woman's subordinate position in society.
	(5.4)	However the fact that <u>it clearly does highlight these issues</u> does suggest that cinema needs to rethink its way of communicating sound space to the audience.(AH, 0224e, L1)
The human response motif	(5.5)	The fact that <u>we can now assign different mental faculties to different parts of the brain</u> casts some doubt on Descartes assertion that the soul is indivisible and also goes a long way to helping us understand the relationship between the body and the mind. (AH, 0292c, L1)
	(5.6)	The fact that <u>the practical application of eugenic science had already been undertaken elsewhere</u> was critical for the Nazi government in justifying their actions. (AH, 0252n, L1)

Based on these semantic categorizations of all three-word shell-noun phrases (e.g. *the (The) fact that*), the distribution of the three motifs in AH L1 sub-corpus is shown in Table 5.6 below:

Table 5-6

Table 5.6: Distribution of the three motifs of the fact that in AH L1 sub-corpus

AH L1		
Rank	Motif	Freq. (%)

1	The ‘cause’ motif	47%
2	The ‘human response’ motif	36%
3	The ‘orientation’ motif	17%

As shown in Table 5.6 above, the most frequently occurring motif is the ‘cause’ motif, followed by the ‘human response’ motif, while the ‘orientation’ motif has the least portion of the shell-noun uses of *the fact that*. It is worth mentioning that the relatively high frequency of the ‘human response’ motif might indicate AH L1 writing’s preference for taking evaluative stances toward the topic of discussion and the averred fact.

The typical situation is that when the shell-noun phrase *the fact that* is primed to occur in the subject-complement and object positions, it is often associated with strong evaluative uses, as shown in the following examples:

- (5.7) What is striking about this topic is the **fact** that Salem seemed to be an average, peaceful settlement and yet the outbreak of frenzy surpassed all precedents in its span and magnitude. (AH, 0261d, L1)
- (5.8) This can primarily draw from the unavoidable **fact** that before the Peninsular War and the crises of legitimacy that it provoked, there was little serious prospect for an independent Spanish America. (AH, 0267a, L3)

The example (5.7) belongs to ‘human response motif’ and the shell-noun phrase *the fact that* occupies the subject complement position, presenting focused information that is explicitly evaluated through ‘evaluative *what*-cleft’ construction (*What*+*[V-link]*+*[ADJ]*+*[be]*+*that* pattern) with the evaluative adjective *striking*. Example (5.8) belongs to the ‘cause’ motif, indicated by the phrasal verb *drawn from*. Here, the shell-noun phrase *the fact that* lends epistemic credence and certainty to the

succeeding new information by referring to it as an *unavoidable fact*. The shell-noun phrase's interpretation is not open to denial but should be taken for granted; the reader is presented with no options but is instead positioned to accept the way the writer evaluates the fact, implicitly establishing an evaluative construction.

Furthermore, a close examination of the concordance lines shows that when AH L1 writers use the shell-noun phrase *the fact that*, they are not necessarily assessing the status of a fact objectively. That is, calling something a fact does not necessarily make it so. The hypothetical propositions following *the fact that* can often have a factual status without being labelled. Further corroboration of this hypothetical aspect or negotiability of *fact* comes from examining the types of *that*-clauses succeeding this node.

A detailed analysis of the syntactic constructions of *that*-clauses motivates this thesis to outline a categorization of four types of appositive *that*-clauses in which the hypothetical aspect (negotiability) tends to play a visible and important role. The grammatical environments of these *that*-clauses are mainly prominent in the variants of verb tenses and forms, as shown in Table 5.7 below:

Table 5-7

Table 5.7: Categorizations of subordinate *that*-cause in AH L1 sub-corpus (Search Term Position: On Right):

	Types of <i>that</i>-clause	Examples
a.	The noun <i>fact</i> is accompanied by a <i>that</i> -clause with hypothetical verb forms such as modal auxiliaries. (5.9)	...this could link to the fact that <u>women may have felt they had no other way of acquiring status, their language use was the only way.</u> (AH, 6174a, L1)
b.	The noun <i>fact</i> is accompanied by a <i>that</i> -clause with epistemic modal auxiliaries conveying future. (5.10)	..yet Böll's characters are still completely oblivious of the fact that <u>their defeat will come in a matter of hours.</u> (AH, 0061a, L1)

- c. The noun *fact* is accompanied by a *that*-clause with subjunctive mood triggered by conjunction word *if*. (5.11) Further to this, the **fact** that we could only exist if the world was ordered takes nothing away from the fact that the level of orderliness is astonishing. (AH, 6179i, L1)

In about three-quarters of the 198 instances of the *that*-clause, however, the grammatical items (e.g. verb tense) show a less conspicuous signal of negotiability. Thus, what is significant is the type of knowledge that the clause construes. Here, this thesis recognizes two broad types of knowledge conveyed by the *that*-clause, as presented in Table 5.8 below:

Table 5-8

Table 5.8: Semantic categorization of the shell-content in AH L1 sub-corpus (Search Term Position: On Right):

Category (knowledge type)	Related human experience	Example
a. The 'claim/assertion' from others	The shell noun <i>fact</i> construes the experience of conveying a string of claims or ideas from others. (linguistic expression)	(5.12) Vincent Cronin has discussed the fact that <u>Burckhardt took Renaissance Italy as a whole and made no distinction between what happened in republican Florence and other cities which were ruled over by tyrants.</u> (AH, 0013b, L1)
b. The 'beliefs/ideas' on the writer's part	The shell noun <i>fact</i> construes the experience of asserting writer's own beliefs and ideas. (conceptual belief)	(5.13) However this query doesn't do anything to belittle the fact that <u>we can still doubt our senses at any given time, we can never be certain that they are being truthful.</u> (AH, 6134b, L1)

As expected, an in-depth qualitative analysis shows that the shadow of negotiability constantly hangs over more than half of the occurrences of *the fact that* in AH L1 writing. The shell-noun uses of *fact*, or more accurately, the use of the shell-noun phrase *the fact that* in AH L1 writing, makes the objective relation between human experience and the real-world state of affairs to which it corresponds irrelevant. It seems that what AH L1 writing is concerned with is not so much the factivity

(Schmid, 2000, p.99) involved but rather the rhetorical persuasiveness realized by the shell-noun use of *fact*.

5.1.1.2. AH L3 writing

Turning now to the AH L3 writing, what is similar to AH L1 writing is that of the 103 instances of the shell-noun uses of *fact*, about 93% occur in the phrase *the fact that*. However, the difference is that AH L3 writing reveals a tendency for the objective relation between experience construed by the shell noun *fact* and the real world state of affairs to which it corresponds to be always relevant. To provide further descriptive and illustrative detail, phraseologies including that three-word shell-noun phrase are presented in Table 5.9:

Table 5-9

Table 5.9: Phraseologies of *the fact that* in AH L3 sub-corpus (raw frequency) (Search Term Position: On Left)

		AH L3			
Rank	phraseology	Freq.	Rank	Phraseology	Freq.
1	<i>∅ the fact that</i>	35	14	<i>was the fact that</i>	1
2	<i>to the fact that</i>	16	15	<i>resented the fact that</i>	1
3	<i>by the fact that</i>	7	16	<i>miss the fact that</i>	1
4	<i>with the fact that</i>	4	17	<i>is the fact that</i>	1
5	<i>on the fact that</i>	4	18	<i>in the fact that</i>	1
6	<i>for the fact that</i>	3	19	<i>identify the fact that</i>	1
7	<i>of the fact that</i>	2	20	<i>given the fact that</i>	1
8	<i>despite the fact that</i>	2	21	<i>enhance the fact that</i>	1
9	<i>from the fact that</i>	2	22	<i>belie the fact that</i>	1
10	<i>as the fact that</i>	2	23	<i>at the fact that</i>	1
11	<i>that the fact that</i>	2	24	<i>altered the fact that</i>	1
12	<i>and the fact that</i>	2	25	<i>affect the fact that</i>	1
13	<i>recognize the fact that</i>	2			

At first glance, the underlying difference might not be clear, but collocations including prepositions reveal an interesting difference between levels of study, which

is prominently featured in AH L3 writing. Specifically, the most frequent of these (nine of the top ten) include a preposition. Using the categorization approach described earlier, Table 5.10 below shows the recurring semantic groups based on the shell-noun phrase *the fact that* preceded by a preposition.

Table 5-10

Table 5.10: Semantic groups of the shell-noun phrase *the fact that* in AH L3 sub-corpus (Search Term Position: On Left):

1. Fact is the basis for a practical outcome or reasoning
<i>[reason] was the fact that</i> <i>due to the fact that</i> <i>be supported by the fact that</i> <i>stems from the fact that</i> <i>be emphasized by the fact that</i> <i>based on the fact that</i>
2. Fact explains something
<i>given the fact that</i> <i>derives from the fact that</i> <i>be criticized by the fact that</i>
3. Fact is the cause of a problem or its solution
<i>[problem] be summarized by the fact that</i> <i>be helped by the fact that</i> <i>[problem] arises in the fact that</i> <i>[problem] be attributed to the fact that</i> <i>be reinforced by the fact that</i> <i>[problem] is the fact that</i>
4. Something shows/uses/assumes a fact
<i>[entity] believes the fact that</i> <i>and the fact that</i> <i>[entity] affect the fact that</i> <i>[entity] centered on the fact that</i> <i>[entity] accounts for the fact that</i> <i>despite the fact that</i> <i>as well as the fact that</i> <i>together with the fact that</i> <i>[entity] point to the fact that</i> <i>coupled with the fact that</i> <i>[entity] emphasizes the fact that</i> <i>[entity] alter the fact that</i> <i>there is emphasis on the fact that</i> <i>[entity] link up to the fact that</i> <i>[entity] enhance the fact that</i> <i>and the fact that</i> <i>combined with the fact that</i> <i>Representative of the fact that</i> <i>In recognition of the fact that</i>
5. Be aware or unaware of a fact
<i>[people] aware of the fact that</i>

[People] <i>emphasize the fact that</i>
[people] <i>miss the fact that</i>
[people] <i>recognize the fact that</i>
[people] <i>identify the fact that</i>
6. People discuss a fact
[people] <i>refer to the fact that</i>
[people] <i>argue that the fact that</i>
7. Affective or reaction to a fact
[people] <i>expressed concern at the fact that</i>
[people] <i>be concerned about the fact that</i>
[people] <i>draw attention to the fact that</i>
[people] <i>appeal to the fact that</i>
[people] <i>altered [people] to the fact that</i>

Table 5.11 below shows the distribution of the three areas of motif of *the fact that* in AH L3 writing.

Table 5-11

Table 5.11: Distribution of three motifs in AH L3 sub-corpus

	AH L3 Freq. (%)
The 'orientation' motif	43%
The 'cause' motif	42%
The 'human response' motif	15%

As Table 5.12 below shows, the changes pertaining to the three motifs for the shell-noun phrase *the fact that* between AH L1 and L3 writings:

Table 5-12

Table 5.12: Changes of distribution of motifs across levels of study in AH (frequency normalized per million words, *LLV* based on raw frequencies)

	AH			
	L1 Freq.	L3 Freq.	<i>LLV</i>	<i>p</i>
The 'cause' motif	227	121	12.1***	<0.001
The 'orientation' motif	97	124	-1.21	>0.05
The 'human response' motif	155	42	24.63***	<0.001

As shown in Table 5.12, there are decreases in two of the three motifs as we move from AH L1 writing to AH L3 writing. The 'human response' motif features the greatest drop of 73%, while the 'cause' motif shows the least drop of 47%. These

changes are perhaps surprising given the typical impressions of ‘more discursive and argument-based uses of shell noun’ on the soft end of a disciplinary continuum uncovered in previous research on shell nouns in disciplinary writing. (e.g. Flowerdew and Forest, 2015, p.169; Jiang and Hyland, 2015, p.538; Benitez-Castro, 2021, p.140). Most unexpected is the massive decline in the ‘human response motif’, which conveys the discussions, reactions and evaluations of facts. While there is no ‘faceless’ writing, especially in soft knowledge fields, and all language choices represent a certain degree of rhetorical efforts in strengthening the persuasiveness and affecting the reader’s interpretation of the knowledge claims, it is interesting to see fewer human-related motifs, which might be an indication of a less interpretative and personal ‘take’ on AH L3 writing.

Furthermore, in terms of the ‘cause’ and ‘orientation’ motifs, there is a shift in the type of knowledge that AH L3 writers may feel confident that they can infer as a fact, or at least epistemologically acceptable as ‘factivity’ (e.g. Steinberg and Osgood, 1971; Francis, 1986). Specifically, the majority of types (95%) of knowledge construed by shell-noun phrases of *the fact that* in these two motifs can be categorized as following shown in Table 5.13:

Table 5-13

Table 5.13: Types of knowledge construed by *the fact that* in AH L3 sub-corpus (Search Term Position: On Right)

Category (knowledge type)	Related human experience	Example
Existences in the real world	status of affairs that has actual existence in the physical observable world.	The ‘cause’ motif: (5.14) Roanoke was not suitable for cultivation probably due to the fact that <u>it was an island and thus the earth was saline</u> . (AH, 0029m. L3)

		<p>The ‘orientation’ motif: (5.15) The imagery and camera work, as well as the fact that <u>Draba was a black slave</u>, would have had a great impact on the film audience. (AH, 6073g, L3)</p>
Quantitative data	Objective evidence supported by quantitative data	<p>The ‘cause’ motif: (5.16) The fact that <u>their numbers dwindled</u> is indicated by the increasing numbers of other food sources such as birds in the assemblage. (AH, 6033e, L3)</p> <p>The ‘orientation’ motif: (5.17) The fact that <u>only approximately 12.6% of the stone artifacts are formal re-touched tools</u> does not take away from the possibility that the animals may have been butchered. (AH, 3019h, L3)</p>
Empirical evidence	Findings resulted from an empirically oriented activity or event.	<p>The ‘cause’ motif: (5.18) It is suggested that the reason these two disturbances occur together is due to the fact that <u>the neurological areas responsiblepartially overlap, or are in close proximity in the brain.</u> (AH, 6174d, L3)</p> <p>The ‘orientation’ motif: (5.19) The fact that <u>the students pronounced them differently</u> suggested that they were using analogy to read the words. (AH, 6174e, L3)</p>

To conclude, in AH L1 writing, the use of shell noun *fact* in SN-*that* post nominal pattern appears to be a useful rhetorical device for marking the epistemic certainty of knowledge claims or propositions presented as true in academic arguments. However, in AH L3 writing, there has been a significant decrease in the use of the shell noun *fact* in SN-*that* post nominal pattern, indicating a shift towards using it to convey its nature of facts and reality meaning. In other words, AH L1 writers tend to accept or even intend the change of the conceptual status of ‘abstract relations or meta-discursive meanings’ (Jiang and Hyland, 2018) from ideas, claims, or possible

facts to accredited facts. In contrast, L3 writers tend to construe the experience of objective conditions in the real world.

Interpreting the reasons for these differences of knowledge construction behind AH L1 and L3 writings here is not entirely straightforward, but one possible interpretation is that such a shift may be partly because of AH L3 writers' greater awareness of building arguments through reference to the accredited truth and a greater involvement in empirically-oriented events, or at least a greater carefulness about the rhetorical impact of language choices in the crafting of academic arguments and the construction of knowledge claims.

5.1.2. *example* in the construal of *identifying relational process*

There has been a remarkable increase (increased by 60%, $LLV = -4.4$, $p < 0.05$) in the use of partitive factual shell noun *example*. It is worth drawing attention to the distinction made between the different ways of using this shell noun *example* in realizing the discourse act of exemplification. The rationale for selecting the shell noun *example* for an in-depth analysis is twofold: a). its generally increasing use in comparison to the massive decrease of the shell noun *fact* in AH L3 sub-corpus, and b). exemplification has been found to play an important role in ensuring persuasiveness and accessibility of abstract expository academic writings (e.g. Hyland, 2007, p.270).

Before moving on to uncover some disciplinary-specific features observed in this

section, it is necessary to discuss the notion of ‘pattern grammar’ (Francis 1993; Hunston and Francis 2000; Hunston 2015) and ‘local grammar’ (Hunston and Sinclair, 2000). In brief, a pattern grammar is an approach to the grammar of English that seeks to account for, not the whole of a language, but one ‘prioritizes the behaviour of individual lexical items’ (Hunston, 2002b, p.169). The most prominent difference between pattern grammar and traditional/general grammar is that pattern grammar sees patterns and meanings as inseparable. That is to say, words with different sense tend to occur in different patterns, and words in a given pattern can be divided into a limited set of meaning groups, such that words in the given pattern share certain aspects of meaning. This approach is rooted in the observations from Sinclair (1991) and Hudson (1984) that meaning and patterning are connected, and that grammar and lexis should be treated as a unity rather than distinct phenomena in a description of English. In this respect, it is thus believed that investigating diachronically *example*’s uses in AH writing by drawing on insights from the concept of pattern grammar is useful because the underlying basis of the pattern grammar approach is consistent with the present study’s theoretical framework, with the view that shell nouns and their syntactic constructions together act as a lexico-grammatical unity to construe disciplinary knowledge.

A local grammar is always ‘a grammar of a discourse function’ (Hunston and Su, 2019, p.571). Some examples of local grammars are grammars of exemplification (Su et al., 2021), evaluation (Hunston and Su, 2019), requests (Su, 2017), apologies (Su and Wei, 2018), disclaimers in company reports (Cheng and Ching, 2018), and affect

(Bednarek, 2008). The most notable difference between local grammar and traditional/general grammar is that the former attempts to assign functional-pragmatically transparent labels to the corresponding pattern elements, whereas the latter prefers traditional grammatical terminologies, such as object, subject (Su et al., 2021, p.122). Speaking of which, the fine-grained categorization of semantic relationship between syntactic units in this study's investigation of exemplification realized by shell noun *example* is drawn from the notion of local grammar (see below).

In conclusion, a pattern-based approach to local grammar appears to offer the researcher a useful means of comparing the discourse act of exemplification instantiated by the shell noun *example* across levels of study in AH. It was found that instances containing the shell noun *example* as an exemplificatory marker may realize different local grammar patterns across levels of study. Furthermore, some nuanced but notable changes are also revealed.

5.1.2.1. AH L1 writing

First, in AH L1 sub-corpus, the resultant local grammar types of syntactic pattern identified for the exemplification that are realized by shell-noun uses of *example* are presented in the following Table 5.14:

Table 5-14

Table 5.14: Overview of local grammar patterns of <i>example</i> in AH L1 sub-corpus (with examples)				
Pattern 1	Exemplified	Indicator	Hinge	Exemplification (subcategory)
	It is unlikely therefore that she would speak with total pessimism.	Another example	is	that one could interpret the class divisions and indeed the social conditions as

				being the reason for people - such as the narrator - to become educated and want to better themselves.(AH, 0061d, L1)
Pattern 2	Exemplification (subcategory)	Hinge	Indicator	Exemplified
	So this	is	example	of how plays pushed at cultural norms in Athens. (AH, 0098c, L1)
Pattern 3	Exemplification (subcategory)	Indicator	Exemplified	
	<u>The designation of the seduction/abduction of Helen by Paris as the 'Rape of Helen' was established....</u>	In this example		from a social point of view, we can observe the severe male bias against the female.(AH, 6192c, L1)

5.1.2.2. AH L3 writing

Second, in AH L3 sub-corpus, the resultant local grammar types of syntactic pattern show that there are not just occurrences of the first three patterns but another three new local grammar types of syntactic pattern (patterns 4, 5 and 6) emerge. Syntactic patterns occurring in AH L3 sub-corpus are outlined below in Table 5.15 below:

Table 5-15

Table 5.15: Overview of local grammar patterns of example in L3 sub-corpus (with examples)				
Pattern 1	Exemplified	indicator	Hinge	Exemplification (subcategory)
	Brecht weaves a number of themes into the opera which are duplicated in his later work.	An example	is	that <u>of business transcending love in an amoral, capitalist world.</u> (AH, 0024g, L3)
Pattern 2	Exemplification (subcategory)	Hinge	indicator	Exemplified
	This	was	example	of Enlightenment thought in relation to the imprisonment of convicts (AH, 0029o, L3)
Pattern 3	Exemplification (subcategory)	Indicator	Exemplified	

	<u>An agricultural worker executed in Kent for arson in 1830 refused to profess his guilt and 'refused to pull down over his eyes, saying he wished to see the people as' as he died.'</u>	Although example	this	spans past the seventeenth century phenomena of the 'set piece' speech, it displays a victim's efforts to maintain his dignity and die well not only through penitence, but through this supportive lifebelt surrounding him, the crowd. (AH, 0044b, L3)	
Pattern 4	Exemplified	indicator	Exemplification (subcategory)		
	It is possible to show exactly how much the people were aggrieved by their situation if the two strikes at the Erikson phone and the Atlas Engineering plants in June to August of 1906 are viewed.	In another example	that <u>workers wanted pay for the day of Nicholas II's coronation in May 1896 at the Rossiiskaia cotton mills in St. Petersburg.</u> (AH, 0244g, L3)		
Pattern 5	Exemplified	Hinge	Exemplification (citation)	indicator	Supporting information
	That H2O and "Water" are not interchangeable <i>salva veritatae</i>	is shown	by Burge (2002)	with the example	that <u>it is possible for a person to think/assert that they do not believe that "water is not fit to drink" but not that "H2O is not fit to drink".</u> (AH, 3019f, L3)
Pattern 6	Exemplification (citation)	Initiator	Hinge	Indicator	Exemplified
	<u>"everything we hold sacred is profane, and conversely they permit what for us is taboo"</u> (Tacitus, Histories, p4).	This	is	example	of how.... (AH, 6053g, L3)

Finally, the distribution of local grammar syntactic patterns across levels of study in AH is illustrated in Table 5.16 below:

Table 5-16

Table 5.16: overview of local grammar across levels of study in AH (normalized per million words, *LLV* based on raw frequency)

No.	Pattern	AH		
		L1	L3	<i>LLV</i>
Pattern 1	Exemplified-Indicator-Hinge -Exemplification (subcategory)	12	24	-0.96
Pattern 2	Exemplification (subcategory)-Hinge	24	6	4.2*

	-Indicator-Exemplified			
Pattern 3	Exemplification (subcategory)-Indicator-Exemplified	21	24	-2.9
Pattern 4	Exemplified-indicator-Exemplification (subcategory)	0	12	-
Pattern 5	Exemplified-Hinge -Exemplification (citation)-indicator-Supporting information	0	21	-
Pattern 6	Exemplification (citation)-Initiator-Hinge -Indicator-Exemplified	0	21	-

As can be seen from the above set of tables, AH L3 writing evinced the most pronounced changes in the occurrences of the last three types of local grammar the in AH L3 sub-corpus, compared to their complete absence from the AH L1 sub-corpus. Clearly, the ways in which AH L3 writers choose to exemplify by incorporating other's work into their own, ranging from casual acknowledgement as in pattern 5 to extended discussion as in pattern 6, can both influence the rhetorical presentation of their cases and effectiveness of knowledge construction. Exemplifying in this way, AH L3 writers mainly situate their arguments in the scaffolding of references to the collaborative construction of a body of new knowledge.

The rhetorical practice of using citation to report the views of authorities and integrate claims into current knowledge reflects what Palmer (1986, p.51) refers to as the 'quotative mode of knowing'. The reasons for the strengthened function of *example* in exemplification (citation) are unclear, but they may be related to the AH L3 writers' increasing proficiency in integrating claims into current knowledge and their confidence in positioning themselves in relation to their topic and disciplinary community (Hyland, 2012). It is also likely, however, that this change is perhaps influenced by the AH L3 writers' heightened awareness of creating a rhetorical context in which exemplification (e.g. citation) is ever more important in contributing

to clarity and guiding readers (e.g. subject lecturers, peer students) to preferred interpretations.

Furthermore, it has also been observed that another difference concerns pattern 2 of the local grammar which is used significantly less in AH L3 writing. Local grammar pattern 2 identified in this thesis is predominately realized by the instances where exemplificatory marker *example* colligates with *th-be-SN* syntactic pattern. *example* in these instances have a great reliance on the prepositional post-modification, as shown in the following Table 5.17:

Table 5-17

Table 5.17: Pattern 2 of local grammar (with examples) in AH L1 writing			
Exemplification (subcategory)	Hinge	Indicator	Exemplified
This	was	an example	of oligarchic capitalism's repression coming into contact with a burgeoning working class identity that demonstrates the dialectical process in bringing about militant revolutionary activity. (AH, 0267c, L1)
This	is	an example	of how the knowledge of pragmatics and communication therapy can help speech therapists. (AH, 6067f, L1)

Previous studies have suggested that, in phrasal terms, elaboration is apparent in the shell-noun uses favoring appositive *of*-phrases instead of condensed nominal premodifiers (Benitez-Castro, 2021, p.147). Presumably, the prevalence of local grammar pattern 2 in L1 writing is related to a more involved, elaborated style of academic writing. The decrease in the uses of pattern 2 in the AH L3 sub-corpus might indicate that AH L3 writing has moved away from a focus on elaboration and amplification towards a rhetorical style that emphasizes succinct explanation and illustration, at least in the dimension of the discourse act (exemplification) performed

by shell-noun uses of *example*. This finding is interesting as it is inconsistent with the results reported in Gray's study (2015b, p.64) which identified that prepositional phrases exhibit 'less variation..and there is no systemic trend of increase nor decrease' from soft disciplines to hard disciplines.

This discrepancy between the findings of my study and previous ones appears to have stemmed from the overlook of diachronic change within a disciplinary domain in the previous study or its focus on academic articles, which are relatively more solidified in the way of meaning-making and knowledge construction. By contrast, perhaps due to the accumulation of disciplinary knowledge and the AH L3 writers' familiarization with the operations of empirical research, AH L3 writing tends to increasingly favor presenting clarification and exposition with simplicity and clarity, exhibiting a compressed style of writing which is less typical in the soft field (e.g. Hyland, 2008; Durrant, 2017; Hyland and Jiang, 2019a; Benitez-Castro, 2021).

Based on these observations, the following conclusions can be tentatively drawn, although it should be cautioned that these conclusions may only apply to the one specific shell noun *example* in limited syntactic patterns (e.g. *SN-that*; *SN-be-that*; *th-SN*; *th-be-SN*). In summary, the partitive factual shell noun *example* showed an opposite changing direction compared to the neutral factual shell noun *fact*: the shell noun *example* has undergone an increase in AH L3 writing. Perhaps more importantly, this increase generally coincides with citations, which underlines the fact that AH L3 writing shows a growing interest in construing the experience of illustrating arguments, viewpoints, observations, and so on., by citing relevant studies rather than

showing subcategories.

Firstly, this trend aligns with the observations drawn in Nesi's (2021, p.1) corpus study, which found that students' use of citation 'was found to increase as they progressed through the years of undergraduate study'. Besides, among the new patterns of exemplification, the advent of patterns 5 and 6 is particularly notable. These two patterns roughly correspond to 'parenthetical' (Swales, 2014b, p.122) or 'non-integral' citations (Hyland, 1999, p.344; Thompson and Tribble, 2001, p.92). Briefly, such citations are usually placed within parentheses and play no explicit grammatical role in a sentence. The choice of non-integral citation is driven by disciplinary convention in scientific writing, where the writers tend to 'de-emphasize the role of the cited authors' (Thompson and Tribble, 2001, p.99). In this respect, it is interesting to find that AH L3 writing shows a tendency to suppress human agency in knowledge-making and place greater emphasis on the relevant studies. This finding offers additional evidence to support Hyland and Jiang's (2019b, p.82) observation that soft fields have shown a greater use of non-integral citations and a movement towards greater technicality or scientism with an associated trend towards scientific reporting styles over the past 50 years. Thirdly, the significant decrease of prepositional bundles: noun+appositive *of* phrase (pattern 2) shows an increasing trend of 'simpler and clearer' ways of exemplifying (Su et al., 2021, p.130) or a 'concise and compressed' style of writing, which is prevalent in hard science writing (e.g. Biber and Gray, 2016, p.170; Gardner et al., 2019, p.657; Benitez-Castro, 2021, p.140). In conclusion, in terms of the use of *example*, it thus seems to be plausible to

say that AH L3 writing reflects some aspects that are closer to not just a disciplinary professionalism but also to hard field writing in comparison to AH L1 writing.

5.1.2.3. Summary: shell-noun uses in the construal of *identifying relational* process.

To summarize, the most significant change observed in AH L3 writers is the marked decline in the use of *fact*. It seems that AH L3 writers are more interested in presenting their ideas as objective reality, and they make less explicit effort to describe the process of shifting conceptual status of the shell content. This is not to say that AH L3 writers are no longer crafting texts that take rhetorical needs into account, or that they completely lose the discursive character of modes of knowing, but that this is done with less obvious ‘semantic bleaching’ or ‘grammaticalization’ (Schmid, 2000, p.99). This point can also be justified by the change of shell-noun use of *example*.

Although the increase in the use of *example* is not as prominent as the greater decrease in the use of *fact*, it is an important piece of evidence that shows that shell nouns seem nevertheless continue to be an important part of AH writers’ rhetorical arsenal. The growing number of shell-noun uses of *example* for the purpose of exemplification corresponds with a general upward trend towards greater exemplification in abstract expository academic articles (Hyland, 2007, p.270). This may be related to the writers’ enhanced awareness of the readers’ possible processing

needs for clarification or at least to their greater care in crafting arguments to support their claims, so that readers can better make sense of their knowledge claims and hopefully accept their propositions. What is interesting about the increase in exemplification is the growing frequency of citations, especially non-integral citations, occurring alongside *example* across levels. This observation is consistent with those of Swales's (2014b, p.136) study of the MICUSP corpus and Nesi's (2021, p.12) study of the BAWE corpus, both of which found that higher relative frequencies of references and citations are observed in more advanced levels or years. This trend may be due to the intrinsic nature of the soft knowledge field, where part of the epistemological and social framework for the acceptance of knowledge is often based on the elaboration, reinterpretation and dialogue of the topic under discussion. (e.g. Hyland, 2008; 2018). It may also be unsurprisingly related to a wider scope of reading in module assignments of a more advanced level.

5.1.3. *attempt* in the construal of *material* process

In AH L3 writing, there has been a substantial increase of 40% in the use of *material* transitivity processes ($LLV=-8.26$, $p<0.01$). Specifically, the only sub-category of material processes that increased across the levels of study was that used to construe specific events and activities. Figures show that the rise in material processes is chiefly due to the increase in the use of *attempt* (increased by 67%, $LLV=-6.7$, $p<0.01$). Since the shell noun *attempt* is mainly colligated with adjacent *to*-clauses in the syntactic form of [*attempt to*-clause], the analysis of shell-noun use

of *attempt* is focused on this pattern.

Although there is no observable difference in the main syntactic patterns hosting the shell noun *attempt* across the two sub-corpora, it is helpful to specify the position of the shell-noun clause in a sentence in order to reveal some possible corpus-specific syntactic features. Thus, this study has formulated more fine-grained sub-types of syntactic patterns based on the core patterns in Schmid (2000). Before moving on to the classification of syntactic patterns that occurred in this study, a discussion pertaining to the simplification of some components of the sentence (e.g. subject, verb) is in order. This is illustrated in Table 5.18 below:

Table 5-18

Table 5.18: Simplifications of terminology in the formulation of sub-types of N-CL

Components of sentence	description
SN	stands for the shell noun.
Np	stands for a noun phrase which occurs at the subject position.
be	stands for <i>be</i> -verbs or linking verbs (e.g. <i>appear</i> , <i>seem</i> , <i>remain</i>).
V	stands for a verb that takes shell noun as an object.
Vp	stands for a verb functioning as predicate and which takes a shell noun as subject.
Non-clausal	stands for adverbial phrase or prepositional complement.
SN-CL	stands for N- <i>to</i> infinitive clause.

The resultant sub-types of the syntactic pattern identified for the core syntactic pattern SN-CL hosting the shell noun *attempt* were the following four types.

Examples of each of the syntactic patterns are shown in Table 5.19 below:

Table 5-19

Table 5.19: sub-types of syntactic patterns formulated for the shell noun *attempt*

syntactic functions	Types of pattern	examples
Subject-based	subject	SN-CL-Vp (5.20) His attempt to use logical, scientific method to prove what is, at its most basic level, an intuition, is faintly ludicrous.(AH, 0215d, L1)
	subject complement	Np-be-SN-CL (5.21) The play is an attempt to understand the Persian situation by translating it into Greek

			<u>terms.</u> (AH, 0098c, L1)
Object-based (direct/indirect object)	Np-V-SN-CL	(5.22)	but the community of practitioners makes little attempt to <u>develop the theory towards solutions of the problems.</u> (AH,03111, L3)
Adverbial phrase or prepositional complement	Non-clausal SN-CL	(5.23)	and finally it stimulated European industries in their attempt to <u>surpass the East through emulation and innovation.</u> (AH, 0019g, L3)

5.1.3.1. AH L1 writing

Based on these four sub-types of SN-CL, Table 5.20 below presents the distribution of these four types of pattern hosting *attempt* in AH L1 sub-corpus.

Table 5-20

Table 5.20: frequent sub-types of SN-CL hosting *attempt* and their distributions in AH L1 sub-corpus.

Syntactic functions	AH L1	
	Types of sub-pattern	Freq.(%)
Adverbial phrase/prepositional complement	Non-clausal N-CL	50%
Object (direct/ indirect)	Np-V-N-CL	30%
Subject-based	Np- <i>be</i> -N-CL	10%
	N-CL-Vp	10%

In AH L1 sub-corpus, non-clausal SN-CL uses are proportionally higher than other sub-types, while object uses are relatively less frequent. Subject-based uses show the lowest proportion of all. An examination of all the premodifying structures of *attempt* reveals that the indefinite article of *attempt* stands out in L1 writing, as shown in Table 5.21 below:

Table 5-21

Table 5.21: Premodifying structures of *attempt* in AH L1 sub-corpus

AH L1		
Rank	Premodifying structures	Freq.(%)
1	Indefinite article	46.7%
2	Possessive determiner	27.0%
3	Adjective	22.2%
4	Determiner	4.1%

It is further observed that the prevalence of possessive determiner and the non-clausal SN-CL syntactic pattern seems to be correlated. In other words, the three-word shell-noun phrase *an attempt to*, preceded by a preposition, almost always occurred in a non-clausal SN-CL pattern. Table 5.22 outlines these frequently occurring phraseologies, including the three word *an attempt to* in L1 the AH sub-corpus:

Table 5-22

Table 5.22: phraseologies of *attempt to* in non-clausal SN-CL syntactic pattern in AH L1 sub-corpus (raw frequency) (Search Term Position: On Left)

AH L1		
Rank	phraseology	Freq.(%)
1	<i>In an attempt to</i>	57.9%
2	<i>As an attempt to</i>	15.8%
3	<i>Of an attempt to</i>	10.5%
4	<i>Through an attempt to</i>	5.3%
5	<i>By an attempt to</i>	5.3%
6	<i>On an attempt to</i>	5.3%
7	<i>After an attempt to</i>	5.3%

As can be seen from Table 5.22, the phrase *in an attempt to* is particularly frequent. A close reading of the context of *in an attempt to* shows that the shell-noun uses of *attempt* in these non-clausal SN-CL syntactic patterns are almost always construed as being a personal volitional status, more accurately, an agent-oriented intention or goal of such events, as shown in the following examples (5.24) and (5.25):

(5.24))Rene Descartes was by most accounts a committed Catholic and wrote his philosophy in an **attempt to** dissuade the Church from hostility to modern ideas.

- (5.25) There is a theory that Desdemona yearns for authority because her father influences her, she models herself on him in an **attempt** to acquire power on a level with Brabantio. (AH, 3005b, L1)

Further corroboration of the focus on attributing a goal-oriented feature to the specific individual agent of such an event in AH L1 writing comes from the relatively sizeable proportion of possessive determiners across the other two sub-types of syntactic patterns (subject-based and object syntactic function), as exhibited in the following examples:

- Subject-based (5.26) John Burnet's **attempt** to reconcile the two halves of Parmenides' work also appeals to the context in which Parmenides was writing. (AH, 0215a, L1)
- (5.27) The first reminder we see in both 'Mateo Falcone' and 'Tamango' of Mérimée's **attempt** to control the reader is the use of a narrator. (AH, 0061b, L1)
- Object (direct/ indirect) (5.28) The fluidity of the sentences and the lack of grammatical and syntactical rules show Woolf's **attempt** to get away from the 'formal railway line of the sentence'. (AH, 3080f, L1)

Another observation concerns the verbs in the *to*-infinitive clause. A scrutiny of all the transitive verb forms in the infinitive *to*-clause in AH L1 sub-corpus reveals some corpus-specific features. Table 5.23 below presents all the verbs in the *to*-infinitive clause succeeding *attempt* and their distributions in AH L1 writing:

Table 5-23

Table 5.23: verbs in infinitive *to*-clause and their distributions in AH L1 sub-corpus (Search Term Position: On Right).

AH L1					
Ranks	Verbs	Freq. (%)	Ranks	Verbs	Freq. (%)
1	create	8.1	21	reconcile	2.7
2	show	8.1	22	recapture	2.7
3	make	5.4	23	place	2.7
4	avoid	5.4	24	justify	2.7

5	keep	5.4	25	increase	2.7
6	control	5.4	26	get away	2.7
7	understand	5.4	27	find	2.7
8	give	2.7	28	explain	2.7
9	use	2.7	29	escape	2.7
10	turn away	2.7	30	end	2.7
11	trick	2.7	31	dissuade	2.7
12	shift from	2.7	32	detect	2.7
13	seduce	2.7	33	compensate	2.7
14	rival	2.7	34	change	2.7
15	re-write	2.7	35	attack	2.7
16	reproduce	2.7	36	answer	2.7
17	remove	2.7	37	acquire	2.7
18	rely	2.7			
19	relieve	2.7			
20	recover	2.7			

As can be seen from Table 5.23, the top three most frequent verbs are related to the human experience of ‘doing’: *make*, *creat*, and *show*. According to the definition of *attempt* in the *Longman Dictionary of Contemporary English* (LDOCE) and the examination of the most frequent transitive verbs, such a combination would normally have an intrinsic meaning of taking actions or performing a piece of work in the physical world. However, the close analysis of the contexts shows that this is not the case in AH L1 writing. Quite the opposite, the shell noun *attempt* seems to be mainly construed as the experience of ‘projecting interior content, ideas into existence’ (Halliday and Matthiessen, 1999, p. 129). Examples (5.29) through (5.31) below give a flavour of this:

(5.29) This plus use of the flare, seems to be an **attempt to make clear that this scenery, unlike the previous landscapes, cannot be romanticised in the same way**, it is too different, too bare and ugly. (AH,0224b,L3)

(5.30) These may be to emphasise how our view of the world is completely relative, but may also be an **attempt to show how our ideas on identity, and hence contradiction, may be flawed**. (AH, 0294b,L3)

- (5.31) In order to support this assertion, the reader would have to view the poem's framework of mythical references and allusions to Vegetation rites as an **attempt to create a sense of continuity and ingrained pattern within its structure**. (AH, 3008h, L3)

In example (5.29), *make* itself is a part of phrasal verb *make clear*. The meaning of the referent (underlined) is to eliminate any cognitive doubts, equivocality or ambiguities. Similarly, the meanings of the *to*-infinitive clauses in example (5.30) and (5.31) emphasize on projecting cognitively perceivable ideas.

5.1.3.2. AH L3 writing

Turning now to the L3 AH sub-corpus, for the purpose of comparison, Table 5.24 below shows the changes in the syntactic functions of *attempt* across levels of study:

Table 5-24

Table 5.24: Changes of types of patterns hosting *attempt* across levels of study (normalized per million words, *LLV* based on raw frequency)

Syntactic functions	Types of pattern	AH		
		L1 freq	L3 freq	<i>LLV</i>
Subject-based	N-CL-Vp	12	42	-6.7**
	Np- <i>be</i> -N-CL	22	42	-1.9
Object (direct/ indirect)	Np-V-N-CL	22	30	-0.2
Adverbial phrase/prepositional complement	Non-clausal N-CL	51	72	-2.2

Whilst AH L1 and L3 writings show preference for the last syntactic function patterns, AH L3 writing is by no means the same as AH L1 writing. Subject-based patterns are common in AH L3's writing, compared to being practically non-existent in AH L1 writing.

Furthermore, these subject-based syntactic functions suggest a strong correlation between the writer's ideological stance (evaluation) and the referent of *attempt* under

discussion. That is, when construing the experience of undertaking activities in the pursuit of aims or intended accomplishments, AH L3 writers use more evaluative markers, such as adjectives conveying positive or negative evaluations.

Partial corroboration of this correlation comes from the quantitative data of the distribution of pre-modifying structures in AH L3 writing. Table 5.25 below shows the changes in premodifying structures for the shell-noun phrase *attempt to* across levels of study:

Table 5-25

Table 5.25: Changes of premodifying structure for the *attempt to* across levels of study in AH (normalized per million words, *LLV* based n raw frequency)

Premodifying structures	AH		
	L1 Freq.	L3 Freq.	<i>LLV</i>
Indefinite article	51	90	-4.2*
Adjective	24	66	-7.64**
Determiner	5	21	-4.2*
Possessive determiner	31	9	3.9*

Whilst the overall preference for the indefinite article has been uniform across the levels of study, it is to be noted that the most notable difference ($LLV=-7.64$, $p<0.01$) concerns the adjective, which ranks second in L3 writing's pre-modifying structures, compared to their less frequent occurrences in L1 writing. During the process of coding and examining concordance lines for *attempt to* in AH L3, the researcher began to notice an increasing occurrence of evaluative adjectives (as in (5.32) through (5.35)). This observation led to the hypothesis that the tendency of evaluation might be one important distinguishing feature between AH L1 writing and AH L3 writing.

(5.32) Godbole's **unsuccessful** attempt to make Mrs Moore, the wasp and a stone equal shows that it is an ability only capable of by God. (AH, 3005a, L3)

(5.33) The music highlights the desperate **attempt** to search for work and it seems that the

faster they pedal the more in vain their quest becomes. (AH, 0061f, L3)

(5.34) Jean Delumeau argues that the period marked the first serious **attempt** to christianise rural Europe by attacking the pagan superstitions that were seemingly rife. (AH, 0144c)

(5.35) Whereas the actions of the mother shown in example (1) and (2) is clearly a significant **attempt** to increase the child's 'communicative competence... (AH, 6048a, L3)

Further corroboration of this evaluative tendency for *attempt* comes from a qualitative analysis of the contexts in which the shell-noun phrase *attempt to* occurs. This close and intensive analysis of *attempt* is mainly due to the fact that evaluative meaning is notoriously difficult to pin down, as it is often ‘implicitly’ expressed and ‘cumulative’ (Hunston, 2011, p.3). The results show that almost all (but three cases) of the subjective-based use of *attempt* are related to the writer’s evaluation, as shown in the following examples drawn from L3 writing:

	(5.36)	The attempt <u>to divide and conquer peasant enclaves</u> was decidedly unsuccessful. (AH, 0255g, L3)
Subject	(5.37)	Thus any attempt <u>to discover a single principle that guided the Levellers on the issue of the franchise</u> would seem fruitless... (AH, 0040c, L3)
Subject-based syntactic function of <i>attempt to</i>	(5.38)	An attempt <u>to show that wealth acquired during his reign, distanced him from his peasant concerns and made him 'a tool of conservative forces,'</u> is unfounded. (AH, 0280h, L3)
Subject-complement	(5.39)	There was a great deal of investigation into where <u>the break between man and ape, and between person and non-person existed</u> . This was a rational scientific attempt <u>to order the universe</u> . (AH, 0029o, L3)
<i>It</i> -cleft notional subject	(5.40)	It seems that any attempt <u>to describe or justify a particular inductive principle</u> is plagued with problems... (AH, 0177b, L3)
Existential <i>there</i> notional subject	(5.41)	Only with the third and final voyage in 1587, was there a genuine attempt <u>to colonize the island and</u>

attempt to construct a permanent settlement. (AH, 0029m, L3)

The reason for this increasing preference for expressing an attitude towards an event encapsulated by *attempt* is probably due to the evaluative and interpretative rationale inherent in soft fields. This explanation is probable as its reasonableness is accentuated by intensive observation of non-clausal SN-CL patterns. Since this sub-syntactic pattern is predominantly frequent in AH L1 writing (see section 5.1.3.1), it may provide some more direct comparisons. It shows that such a preference for evaluative use is not confined to subjective-based patterns, but is also prevalent in non-clausal SN-CL pattern in AH L3 writing, as shown in the example below:

- (5.42) Sacrificed by their allies Britain and France in a futile **attempt** to avoid war by ceding substantial Sudetenland territory in the Munich Settlement of 29 September 1938, full occupation was really only a matter of time. (AH, 0318f, L3)
- (5.43) ...her drama is in possession of an overreaching feel in it's grand **attempt** to make sense of the world rather than a specific event or social agenda. (AH, 0126f, L3)
- (5.44) For if she fails in her magical **attempt** to make Anselmus marry Veronika, then this is not diminished by her obvious gift of prophecy... (AH, 0229c, L3)

Now let's turn to the transitive verbs in the *to*-infinitive clause on the right side of the node. Table 5.26 outlines all the transitive verbs that occurred in the infinitive *to*-clause for *attempt* in L3 writing.

Table 5-26

Table 5.26: Transitive verbs in infinitive *to*-clause for the shell-noun phrase *attempt to* in L3 texts (Search Term Position: On Right).

AH L3					
Ranks	Verbs	Freq. (%)	Ranks	Verbs	Freq. (%)

1	demonstrate	6.7	23	marry	1.7
2	show	6.7	24	launch	1.7
3	capture	5.0	25	use	1.7
4	explain	3.3	26	integrate	1.7
5	prove	3.3	27	ascertain	1.7
6	solve	3.3	28	harness	1.7
7	regain	3.3	29	describe	1.7
8	make	3.3	30	equate	1.7
9	find	3.3	31	establish	1.7
10	create	3.3	32	escape	1.7
11	base	3.3	33	divide	1.7
12	avoid	1.7	34	characterize	1.7
13	reduce	1.7	35	develop	1.7
14	discover	1.7	36	validate	1.7
15	overcome	1.7	37	convey	1.7
16	improve	1.7	38	colonize	1.7
17	change	1.7	39	surpass	1.7
18	reach out	1.7	40	authorize	1.7
19	project	1.7	41	suggest	1.7
20	produce	1.7	42	label	1.7
21	give	1.7	43	reflect	1.7
22	order	1.7	44	arrive	1.7

Table 5.26 casts light on the uses of *attempt* in AH L3 writing, as most transitive verbs occur only in AH L3 writing compared to their complete absence in AH L1 writing. Three observations might be made about this result: L3 writing appears to be driven by 1). its concern with identifying difficulties/situations faced by experiencers (as in (5.45)); 2). offering detailed accounts of identifying and ascertaining the methods/ways/reactions through which the experiencers either solve or avert problems (as in (5.46)); 3). exploring correlations/causalities between states of affairs based on empirically-oriented evidence (as in (5.47)):

(5.45) The text therefore may be an **attempt** to demonstrate the difficulty faced by women in trying articulate their voices within the constraints of masculine language. (AH, 3008f, L3)

(5.46) Changing farming methods may have been adopted in the **attempt** to overcome these imbalances as demonstrated by a changing pollen record between stages A and B of

Ballynagilly, Co. (AH, 6171c, L3)

- (5.47) A new understanding of evolutionary processes prompted F. Clark Howell to begin work in 1951 in an **attempt** to explain Neanderthal morphology as a result of genetic isolation in addition to glacial adaptation. (AH, 6171e, L3)

5.1.3.3. Summary: shell-noun uses in the construal of *material* process

Since *material* process encoded by eventive shell-noun uses construes the human experience of actions and processes, as well as real-world evidential cases, it is expected that the significant increase of *attempt* might indicate a greater tendency towards a more research-based procedural approach to the shell-noun use of *attempt*, revealing some tendencies towards scientific modes of knowing in AH L3 writing. The context-sensitive analysis of *attempt* has provided several meaningful findings in support of this initial hypothesis. For example, the analysis of the verbs in the infinitive *to*-clause suggests a prominent discourse and cognitive-act perspective in AH L1 writing, while the approach to shell-noun use of *attempt* in AH L3 writing tends to be comparatively more pragmatically- and empirically-based.

However, this thesis adds nuance to these findings. The analysis of data in this section suggests that AH L3 writing reveals a trend towards making both overt and implicit evaluations towards a specific event. The shell-noun use of *attempt* in L3 writing seems to serve as an important part of the writers' rhetorical armoury to strengthen the persuasiveness of their arguments. This finding casts light on how writers in different levels within the same disciplinary domain perceive the necessity of conveying evaluations towards a specific event under discussion in order to garner

ratification of their knowledge claims and arguments. It seems that as AH L3 writers' understanding of disciplinary knowledge develops and matures, not just does their consciousness of authorial evaluative positions increase, but also their confidence in indicating them.

5.2. Changes in SS

Overall, according to the normalized figures in Table 4.37 (see section 4.3.3), there has been a significant decrease in the use of shell nouns in the SS corpus ($LLV=7.52$, $p<0.01$). Further scrutiny of the specific patterns of transitivity processes reveals that writers in this disciplinary domain have substantially reduced their use of shell nouns in the construals of *identifying relational* ($LLV=16.04$ $p<0.001$) processes and *mental* processes ($LLV=7.02$, $p<0.001$). Whilst there has been a significant increase in the use of directive shell-noun uses under *verbal* process, the starting level was extremely low, which limits its analytical value for this systemic analysis. Therefore, this section mainly focuses on the two patterns of transitivity process instantiated by shell-noun uses, which are *identifying relational* and *mental* processes. Specifically, this analysis emphasizes neutral factual shell-noun uses in the construal of *identifying relational* processes and evidential, conceptual, as well as volitional shell-noun uses in the construal of *mental* processes.

5.2.1. *fact* in the construal of *identifying relational* process

The remarkable decrease of neutral factual shell-noun uses in SS is due to a considerable decrease of one shell noun item: *fact* (decreased by 33%, $LLV=5.7$ $p<0.05$). Therefore, the analysis in this section mainly focuses on the shell-noun use of *fact*.

5.2.1.1. SS L1 writing

As can be seen from table 5.27 below, while the frequency of the SN-*that* post nominal pattern has decreased significantly by 40%, it still remains the most frequent pattern across the levels of study. In addition, the *th*-SN pattern records the most significant increase, with a massive 130%.

Table 5-27

Table 5.27: Changes of shell noun *fact* syntactic patterns (normalized per million words, LLV based on raw frequency)

	L1 freq	SS L3 freq	LLV
SN- <i>that</i> post nominal	446	270	8.3**
SN- <i>be</i> clause	13	13	-0.4
<i>th</i> -SN	13	30	-3.8*
<i>th-be</i> -SN	0	0	-

The decrease in the use the post nominal SN-*that* has offered evidence that shell-noun use of *fact* is mainly driven by this syntactic pattern. Furthermore, the increase in the use of the *th*-SN pattern in SS echoes Benitez-Castro's (2021) finding that anaphoric shell-noun use is more frequent in soft disciplines such as sociology and business. This finding is also consistent with that of Jiang and Hyland (2017),

who found that research articles increasingly invest efforts in strengthening anaphoric cohesive ties with prior moves by using shell nouns.

This trend in SS L3's writing may point to a higher preference for professionalism in crafting a smooth flow of discourse and establishing textual coherence, in this pattern, *fact* functions as an anaphoric discourse marker, gathering up the previous discourse and in so doing, moving the argument forward while introducing a new topic. As Schmid (2000, p.343) observes, shell-noun phrases in the *th-SN* pattern not only endow a fairly high accessibility of antecedents but also signal a discourse change. The following analysis starts with an investigation of anaphoric shell-noun uses of *fact* in SS and ends with a comparison of cataphoric shell-noun uses of *fact* in two broadly cognate soft disciplinary domains, namely AH and SS.

A close analysis of the concordance lines of the shell noun phrase *this fact* reveals a nuanced difference in how SS L1 and L3 writings view the world and develop a body of new knowledge. Starting with SS L1 writing, there are two observations that might be made about shell-noun uses of *this fact*. First, SS L1 writers tend to devote their efforts mainly to pinpointing and discussing a cause-result causal relation by directing the reader's attention to the cause component in the preceding discourse, whilst explicitly referring to the result component in the succeeding discourse (Schmid, 2000, p.102), as shown in example (5.48) below. In this example, *fact* easily allows a congruent replacement by shell noun *reason*. Perhaps more importantly, such a causal relation introduced by the use of *fact* is often accompanied by the writer's evaluation and interpretation.

- (5.48) Most living primates are arboreal and it is generally agreed that it is this **fact** which led them to successfully develop many of their physical traits. Their stereoscopic vision, enabled by the positioning of the eyes in the head, gives primates great depth perception which allows them to carry out complex tasks in the tricky environment of fine terminal branches. (SS, 3088a, L1)

In this extract, the writer explains why primates can develop their useful physical traits. The meaning of *this fact* is expressed in the preceding discourse, which is a cause segment, stating the habitation of primates. The causality construed by *this fact* is indicated by the phrase *led them to successfully...* The anaphoric shell-noun phrase *this fact* functions as a link connecting the situation to the succeeding clause that expresses the writer's evaluation, as *successfully*. In fact, the whole following sentence is a manifestation of the writer's evaluation. Semantically, this sentence can be simplified and represented as follows:

Cause (antecedent): Most living primates are arboreal

←*this fact*

Result (evaluation): they successfully develop many physical traits which are useful and advanced

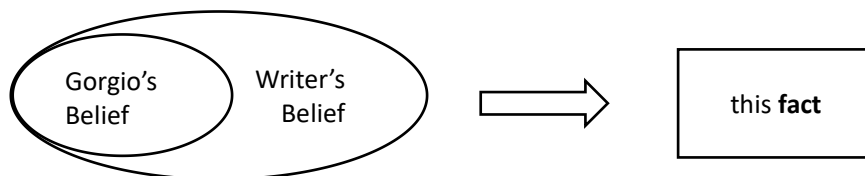
Secondly, in SS L1 writing, the shell-noun use of *fact* in the *th-SN* syntactic pattern strongly reflects the process of 'semantic bleaching' or 'grammaticalization' (Schmid, 2000, p.99). As discussed earlier in section 5.1.1, this feature is saliently prevalent in the *fact-that* post nominal clause, especially in the AH L1 sub-corpus. Something of this similarity can be seen from the following example (5.49):

- (5.49) On the contrary, the Gorgio population regard this animal as probably the cleanest of all,

because of this **fact**, and that they also bury their faeces, hence their popularity as common, domestic pets. (SS, 3055a, L1)

The paragraph from which the example (5.49) is taken is a discussion about beliefs of cleanliness held by two different groups of people within what the writer refers to as ‘gypsy society’. The antecedent of *this fact* (underlined) provides a belief about the cleanliness that is reflected in the viewpoints toward cats. What follows after the shell-noun phrase *this fact* is an ethological description of cats that supports the previous claim. Cognitively, the shell-noun phrase *this fact* is employed to transform a conceptual combination in which the writer’s belief overlays Gorgio’s belief into a true fact. This sequence can be illustrated as following:

Claim1 (Antecedent): this animal as probably the cleanest of all
 ←*this fact*
 Result (claim2): their popularity as common, domestic pets.



As some philosophers (e.g. Zagzebski, 1996; Sosa, 1997) have argued, knowing a fact involves the manifestation of epistemic virtue in the truth of one’s belief. It seems that SS L1 writers tend to build knowledge by making the effort to be trustworthy informants.

5.2.1.2. SS L3 writing

Turning now to the shell-noun uses of *this fact* in SS L3 writing, it is

characterized by two features. Firstly, when the shell noun phrase *this fact* occupies the theme position of a clause, a main feature is that *this fact* functions to create an evidential relation between the two segments, as show in example (5.50) below:

(5.50) The establishment of genocide and crimes against humanity require fulfilment of the definitions in Articles 6 and 7 of the Rome Statute respectively. The crime of genocide requires specific intent that is targeted at particular groups of individuals as listed in Article 6 of the Rome Statute. This requirement sets a rather high level of intent and thus not every harmful act perpetrated will fall into the category of genocide. The essential elements are the numbers and specificity. This **fact** is highlighted in the Chief Prosecutor's communication concerning the situation in Iraq. Although "genocide remains a central conceptual component of international criminal law..." certain factual and procedural aspects have to be satisfied before a thorough investigation procedure. This may be considered a hindrance to a further development of genocide if what is essentially required are statistics and pedantry. (SS, 0312i, L3)

The essay, from which this extract is taken, discusses how adherence to rules affects the intervention of international criminal law in cases of human rights atrocities committed by individuals. *This fact* refers to the immediately preceding segment (underlined), which states the criteria for establishing crimes against humanity. At first glance, the shell noun phrase *this fact* seems to resemble that of SS L1 writing, as the shell content (antecedent) seems to shed some argumentatively and disputatiously relevant insights on shell noun phrase *this fact*. It is so, if only look at the immediate context of a single sentence. However, the aspect of ‘truth’ and ‘factivity’ of shell noun content is clearly justified by the Rome law in the extended segment that stretches into several sentences preceding the shell noun content (wavy-lined). The shell-noun phrase *this fact* in the *th-SN* construction seems to indicate an evidential relation, as it allows a replacement by shell noun *evidence*.

Second, when the shell-noun phrase *this fact* occupies the rheme position, it functions to highlight a factual contrast, as shown in the example (5.51) below:

- (5.51) Although this statement depicted the function of trusts in the seventeenth to the eighteenth century it can equally be applied to the nineteenth century period. However it is important to highlight that the use of trusts was also utilized by farmers, tradesmen, professionals and even charities; trusts were no longer limited to the landed classes. Dickens was also aware of this **fact**. (SS, 0312h, L3)

The essay, from which this extract is taken, is about the assessment of the usefulness of Dickens' Bleak House in interpreting the law and practice in the work of the old court of Chancery. The referent of *this fact* (underlined) is a description of the broadened spectrum of the use of trusts during the nineteenth century period. This description shows a difference in the use of trusts between the eighteenth and nineteenth centuries, and is introduced by the adversative conjunction *however*, which signals the contrast between the previous statement and the current discussion. In other words, the information encapsulated by the shell noun phrase *this fact* represents a contradiction or inconsistency in the light of the issue about trusts. The shell-noun phrase *this fact* thus functions to highlight a factual comparative aspect of the issues or affairs under discussion, as it allows a replacement by the shell noun *contradiction*. The tenability of this point is supported by the definitions of the noun *contradiction* offered by Oxford Advanced Learner's Dictionary (OALD) as 'a lack of agreement between facts, opinions, actions, etc.' and by the Oxford English Dictionary (OED) as 'a state or condition of opposition in things compared; variance; inconsistency, contrariety'.

It is worth mentioning that the overall decrease in the cataphoric shell-noun use of *fact* has been relatively uniform across the two soft disciplinary domains (AH and SS) examined in this thesis. On this account, it seems helpful to round off the analysis in this section by comparing the changes of cataphoric shell-noun use of *fact* (e.g. *fact that*) across levels of study within the two cognate disciplinary domains.

5.2.1.3. Comparison of *fact that* between the two cognate disciplinary domains

The shell-noun use of *fact* over the levels of study in the two soft disciplinary domains has decreased by a similar extent, although the frequency of *fact* in AH greatly exceeded that of SS. Statistically, both corpora have witnessed a drop of approximately 30%. Interestingly, however, there have been considerable differences in the types of knowledge shelled by *fact*. Here this thesis recognizes three main types of knowledge, as shown in Table 5.28 below:

Table 5-28

Table 5.28: Three types of knowledge construed by *fact that*

Types of knowledge	Related human experience	Examples
A. Logical reasoning	Conceives of facts as coherence (e.g. rationality, logical consistency) (5.52)	A possible objection to this is the fact that <u>we scream due to pain, or jump back due to the pain caused by a fire.</u> (AH, 0294c, L1)
B. Objective condition	Conceives of facts as correspondence (e.g. real-world objects, empirical accuracy) (5.53)	This is indicated by the fact that <u>the number of people under 35 who read a daily newspaper has dropped from two thirds in 1965 to one third in 1990.</u> (SS,0135c, L1)

C. Disciplinary Familiarity	Conceives of facts as community rule-of-thumb practices and usual beliefs (e.g. disciplinary-specific appreciation tradition and typicality)	(5.54)	The fact is that <u>judicial punishments are always left in the hands of "uncontrolled discretion of judges or juries"</u> . (SS,01371, L1)
-----------------------------	--	--------	--

Based on this categorization, Table 5.29 below demonstrates the distribution of the types of knowledge across levels of study by disciplinary domains. It is worth noting that knowledge related to two of the three types: logical reasoning and disciplinary familiarity, has fallen across the two soft disciplinary domains, while knowledge relating to objective condition has remarkably risen in AH but decreased significantly in SS.

Table 5-29

Types of knowledge	AH			SS		
	L1 freq	L3 freq	LLV	L1 freq	L3 freq	LLV
Logical reasoning	41	33	0.31	38	22	0.87
Objective condition	34	81	-7.56**	32	9	4.65*
Disciplinary Familiarity	447	226	24.13***	382	305	1.58

Arguably, these changes may represent a shift in what writers feel confident that they can recognize and present as accredited knowledge or, at least, epistemologically accepted in their communities. The most dramatic shifts have been observed in AH, where objective condition has increased nearly 240% while disciplinary familiarity has fallen by nearly 50%. These changes probably reveal a shift in AH writing's preference, from averring claims as if they are true realities, to a greater focus on presenting objective aspects of conditions in practices. The increase of objective condition in AH L3 writing may possibly be related to what has been perceived by

some as increasing scientism (e.g. Glynos and Howarth, 2007) in soft science fields, resulting from a more empirical and hard science orientation in primary research methods and approaches. In the present study, this feature is especially apparent in certain disciplines, such as applied linguistics, where the L3 students' skilled exploitation of technologies that permit quantitative support for language enquiry and more precise data measurement significantly contributes to the establishment of legitimate disciplinary knowledge. This finding lends further support to the results of previous investigations (e.g. Rezaei et al., 2019; 2021) that identified increasing scientism in applied linguistics practices.

Regarding the decrease in disciplinary familiarity in AH L3 writing, one possible explanation is that as the AH L3 writer's understandings of the field matures and develops, their urgent need to label their interpretations as '*fact*' declines. In other words, it may simply be safer for them to be less assertive about their understanding and interpretation of theories and beliefs.

Interestingly, the figure for one particular type of knowledge in another soft field exhibits an opposite tendency compared to AH. In SS, the figure that shows the greatest fall is in those relating to objective condition, which has declined by 72% when normalized for text. This suggests that in SS L3 writing, writers show less confidence in arousing interest in their studies by highlighting the generalization and homogeneity of real-world conditions. This apparent contradiction is possibly a result of the interwoven relationship between the inherent reiterative and contextual nature of soft knowledge construction and the capricious nature of asymmetry, non-linearity,

and unpredictability characteristic of human societies. With a deeper understanding of disciplinary knowledge and greater involvement in intractable of specific social problems, there seems to be more obligation and need for SS students to infer rather than determine factivity.

5.2.1.4. Summary: shell-noun uses in the construal of identifying relational process

To recapitulate, SS L1 writing tends to use the shell noun *fact* in *th-SN* construction to construe the experience of conveying their evaluations, demonstrating logical reasoning, and well-reasoned academic arguments. Although never as conspicuous in ‘grammaticalization’ (Schmid, 2000, p.99) as the collocation *fact-that* post nominal clause, the anaphoric shell-noun use of *fact* in the *th-SN* syntactic pattern has nevertheless been an important element of SS L1 writers’ rhetorical armoury. This finding mirrors the evaluative and reiterative style of AH L1 writing (see section 5.1.1.1). The reason for this writing style might be related to their levels of disciplinary-knowledge understanding and types of readership. Speculatively, SS L1 writers’ understanding of the field has not yet fully matured, nor do they have the sufficient data-driven factors to comprehensively support their knowledge claims. Therefore, the *fact* put forward for truth demonstration throughout their studies is often presented in an equivocal light. Moreover, first year writing tends to ease its novice readers/audiences into the specialized field of study by using rhetorical

presentations of reasoning to introduce the fundamental concepts of disciplinary knowledge in relatively plain and categorical factual claims (e.g. *this fact; the fact that...*).

By contrast, the anaphoric shell-noun use of *fact* in SS L3 writing mainly contributes to the illustration of a factual aspect pertaining to the evidential relations (as in (5.50)) or contradictory state of affairs (as in (5.51)). Although the shell noun *fact* itself has little, if anything, to do with the integration, change or development of ideologies and concepts, the anaphoric use of *fact* in SS L3 writing seems to reflect L3 writers' subtle changes in thinking about the question 'what is it to know a fact'. In other words, what is suggested in SS L3 writing is a tendency toward a more detailed marking of certain specific factual aspects of social affairs that contrast or correlate. This conveys the impression that studies in L3 are being reported more objectively in this soft disciplinary domain.

A common cataphoric shell-noun use of *fact* that takes the form of SN-*that* and suggests some notable differences between two cognate, soft disciplinary domains. In terms of the types of knowledge shelled by *fact*, AH L3 writing reflects a trend toward a more scientifically-oriented knowledge structure, as the construals of knowledge based on objective condition have increased substantially. In fact, this feature is apparent in texts written by writers from applied linguistics (Hyland and Jiang, 2019b), a discipline that 'takes on the rhetorical style of more established, and more 'scientific' academic communities' (ibid, p. 74). As discussed in section 5.2.1.3, AH L3 writing tends to identify specific factual relations among capricious social affairs

and to prioritize credibility and objectivity over personal judgments and evaluations. However, SS L3 writing also shows a greater circumspection in the construals of knowledge based on objective condition. The reason for this seemingly contradictory finding is unclear, however, we cannot rule out the pressure that the SS L3 writers are under to get access to the massive amounts of transactional social data continuously generated by ‘corporations who are dominating data-driven social science’ (Ruppert, 2013, p.269). This pressure may be especially acute given the academic backgrounds of L3 writers, who are, after all, still undergraduate students rather than experts.

5.2.2. *idea, evidence, solution* in the construal of *mental* process

Turning now to the *mental* process construal, the findings in this section offer further linguistic evidence to support the often-invoked arguments in recent studies of academic writing that some social science disciplines (e.g. business, economics), share certain empirical and procedural orientations with hard knowledge fields, despite being classified as soft knowledge fields (e.g. Hyland and Jiang, 2018a; Benitez-Castro, 2021).

As Table 4.37 in section 4.3.3 illustrates, there has been a substantial overall change in the use of *mental* processes, with various changes in sub-categories. The most prominent difference pertains to two sub-categories: conceptual (e.g. *idea*) and volitional shell nouns (e.g. *solution*), which have seen a 35% and 43% decrease, respectively. Interestingly, it appears that the change in *mental* processes is polarized

based the different tendencies between evidential shell nouns and other sub-categories. The sub-category of evidential shell nouns (e.g. *evidence*) has undergone a significant rise (increased by 280%) across the levels of study. Therefore, the analysis in the following section is mainly focused on the uses of creditive, evidential and volitional shell-noun uses. Since the occurrences of these shell nouns in *mental* process are not as frequent as those of *identifying relational* process, such as *fact*, the in-depth analysis in this section primarily focuses on the meanings of referents and rhetorical patterns (e.g. clause relations, text patterns).

5.2.2.1 *idea* in creditive shell-noun use

The decline of the shell noun *idea* is primarily due to the overall decline of creditive shell nouns, which have decreased by nearly 70%. The most frequent pattern in which the shell noun *idea* occurs is **SN-that** post nominal clause. Although the core patterns examined in the two sub-corpora fail to reach a significant difference, it is illuminating to draw conclusions concerning the information and attention status from the clause position. Specifically, this thesis identifies 7 fine-grained sub-types of syntactic patterns that house the shell noun *idea* across levels of study, comprising 3 core types of general syntactic pattern (e.g. **SN-CL**, *th-SN*, **SN-be-CL**), as shown in Table 5.30 below:

Table 5-30

Table 5.30: Sub-types of shell noun <i>idea</i> syntactic patterns (with examples)			
Types of syntactic pattern	Description		Example
Np-V-SN-CL	Shell noun phrases in the SN-CL pattern occur	(5.55)	this supports the idea that <u>individuals can cure themselves, or play an active part in</u>

	as focus (rheme/objects)		<u>preventing disease.</u> (SS, 0320g, L3)
SN-CL-Vp	Shell noun phrases in the SN-CL pattern occur as topics (theme/subjects)	(5.56)	However, the idea that <u>institutional control is significant</u> provides a platform from which wide-ranging changes in the relationship of these two forms of schooling can be identified. (SS, 0252j, L3)
Np-V-th-SN	Shell noun phrases in the <i>th</i> -SN pattern occur as focus (rheme/objects)	(5.57)	Cicourel also supports this idea ; however...(SS, 0001b, L1)
<i>th</i> -SN-Vp	Shell noun phrases in the <i>th</i> -SN pattern occur as topics (theme/subjects)	(5.58)	This idea is re-enforced by figure two which compares the incidence of crises with and without capital controls. (SS, L3)
Non-clausal SN-CL	Shell noun phrases in the SN-CL pattern occur as adverbial phrase	(5.59)	Again, this has its foundation in the socialist idea that <u>the welfare of people should supersede that of the state.</u> (SS, 0135e, L1)
Non-clausal <i>th</i> -SN	Shell noun phrases in the pattern <i>th</i> -SN occur as adverbial phrase	(5.60)	Being completely socialised to this idea makes people less likely to react against the system as it is totally installed within them. (SS, 3135c, L1)
SN-be-CL	SN-be-CL	(5.61)	The key idea is that <u>theory is applicable to the real world and does reflect the real world, though with varying perspectives to varying degrees.</u> (SS, 0137g, L1)

Before moving on to the observations made in this thesis, it is necessary to briefly discuss the information status of theme and rheme. Theme and rheme are terms that represent the way in which the information is distributed in a clause, and they form the main system of the textual metafunction. Theme is the initial element of a clause that serves as the ‘point of departure’ (Halliday and Matthiessen, 2014, p.93) of a message. In other words, the theme is ‘given’ or ‘old’ information (Halliday and Matthiessen, 2014, p.94) that provides the context for the remainder of the clause-rheme. Rheme is the ensuing part of the clause that constitutes the wholeness of the sentence, and it typically contains ‘unfamiliar’ or ‘new’ information. (ibid.) that expands upon the topic.

5.2.2.1.1. SS L1 writing (*idea*)

As shown in Table 5.31, the most notable difference in this case emerges with the first subtype of the pattern Np-V-SN-CL, which relates to the rheme element (focus positions) and stands out in SS L1 writing. Syntactic patterns relating to the theme elements (topic positions) remained steady across the levels of study. A possible reason for the decline in the rheme element (focus positions) Np-V-SN-CL is that SS L3 writers are less likely to forward knowledge related to the ‘idea’ family in a manner that emphasizes the distribution of attention. This change may be related to SS L3 writers’ increasingly mature understanding of disciplinary knowledge and their specialization in empirical methodologies.

Table 5-31

Table 5.31: Changes of types of pattern hosting shell noun *idea* (normalized per million words, LLV based on raw frequency)

Types of pattern	SS		
	L1	L3	LLV
Np-V-SN-CL	140	56	7.03**
Np-V- <i>th</i> -SN	38	9	3.90*
SN-CL-Vp	25	13	0.78
<i>th</i> -SN-Vp	19	17	0.01
Non-clausal SN-CL	25	17	0.29
Non-clausal <i>th</i> -SN	13	4	0.83
SN- <i>be</i> -CL	6	4	0.83

Further support for this assumption is found in a close analysis of the concordance lines. This reveals a notable divergence in the information and attention status of the shell contents in appositive *that*-clause. Specifically, the information status, which is either new or given information, differs in SS L1 and L3 writings. This difference may have an impact on the metadiscursive functions of the shell noun

idea, and, more importantly, it may reflect the different ways of new knowledge construction.

Although the subtype of the pattern Np-V-SN-CL, in which the shell-noun phrase *idea that* occurs as the rheme element in a focus position serving to provide new information predominates in the SS L1 writing, in the majority of cases, the information status of the referent of *idea* is actually ‘given’, as in example (5.62) below:

(5.62) However by constructing ethnic minorities as a reserve army of labour racist beliefs are reinforced, as they are viewed as both inferior and only suitable for inferior jobs, but also as a threat to the white workers, creating a sense of racism in white workers.Since the Second World War British politics reinforced have the **idea that those from other races are inferior.**Throughout the 1950s the debates on immigration focused on the **idea that the number of black immigrants entering Britain needed to be restricted as they were posing a threat to British society.** (SS, 0001a, L1)

The essay from which this example is taken is about racism and its persistence in modern society. There are two shell-noun phrases using *idea that* (boldfaced) in the text. At first glance, it seems that these two shell nouns are lexicalized in an ‘unmarked’ (Halliday and Matthiessen, 2014, p.91) way in which ‘new information is at the end of the information unit’ (ibid.), being expressed in an appositive clause. However, viewing the shell contents in a larger context, it can be seen that the information in these two shell contents is actually a repetition of an earlier extended segment (wavy-lined). Something of this repetitive or circular texture structure can be seen from Table 5.32 below:

Table 5-32

Table 5.32: Texture structure of example 5.62	
Clause 1 T1 However...of labour	→ R1 racist beliefs....workers.
	↓↑
Clause 2 T2 Since... war	→ R2 (=R1) politics reinforced...inferior.
↓↑	→ ↓↑
Clause 3 T3(=T2) Throughout the 1950s	R3(=R1) the debates focused on.. society.

In his study of professional written English, Schmid (2000, p. 330) states that ‘when shell noun phrases in the pattern SN-CL occur as topics, one might expect that the shell nouns or the shell contents or both elements represent given or at least accessible information, but cases where the information is actually given are very rare’. Such frequent occurrences of given information suggest a divergence between SS L1 writing and professionally written academic articles in terms of rhetorical context in which they work. One possible explanation for this divergence is that SS L1 writers are inexperienced writers in their primary level of study with a lack both of control of texture coherence and of the ability to form a natural and unmarked flow of information and arguments. Conversely, another explanation for the trend of such a circular discourse realized by the use of the Np-V-SN-CL syntactic pattern is that SS L1 writers may purposefully exploit the rhetorical potential of shell nouns as fully as possible as a persuasive device, thereby strengthening their arguments. In short, within a given scenario or set of contextual conditions, SS L1 writers tend to ‘play with’ (Halliday and Matthiessen, 2014, p.93) the system of theme and the system of information to produce an ‘astonishing variety of rhetorical effects’ (ibid.).

5.5.2.1.2. SS L3 writing (*idea*)

In contrast, the information status of the shell contents of *idea* in the Np-V-SN-CL construction is almost always presented as new in SS L3 writing, as shown in example (5.63) below:

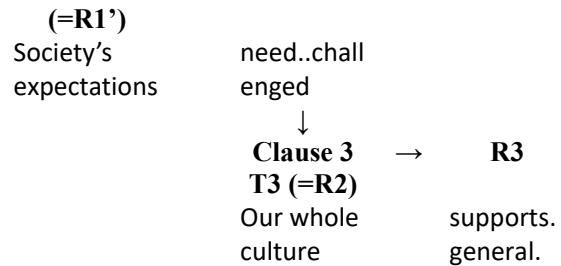
- (5.63) Welldon (1992) however illustrates that the problem of motherhood as a source of subordination lies at the heart of society; it is society's expectations that need to be challenged. Our whole culture supports the **idea** that mothers have complete dominion over their babies, 'we neither help her nor her children, nor society in general' (Welldon, 1992, p 83). (SS, 0214c, L3)

The essay from which this extract is taken is about sexual inequality in society. The shell content (underlined) is presented as new information since there is no corresponding mention in the earlier segment. Viewing it in a larger context, it is noted that this extract has highly linear cross-referential links from the rheme of one clause to the theme of the next clause. The writer presents complex arguments in which each successive idea expands upon the previous sentence's idea. Furthermore, it is important to note that such a linear progression of thematic development may reflect an aspect of the tightly and linearly concatenated organization of SS L3 writing, as shown in Table 5.33 below:

Table 5-33

Table 5.33: Textual structure of example (5.63)

Clause 1	→	R1	→	T1'	→	R1'
T1 Welldon (1992) however		illustrates that		the.. subordin ation		lies..society
						↓ Clause 2 T2 → R2



Speculatively, what is suggested in this typical new information status of *idea* occupying a focus position in the Np-V-SN-CL pattern is that SS L3 writers may excel more at constructing a discussion of subsumption leading to a vertically apparent feature within a text. New knowledge, then, is more likely to be constructed on a basis of cohesive logical reasoning rather than contextualized plausible interpretation compared to SS L1 writing.

One possible explanation might be that the first year SS texts allow for a gradual easing into the field with a lot of iteration and interpretation, while the third year SS texts, on the other hand, efficiently and technically put forward their heuristic and experiential aspect of meaning in a manner of linearity. The robustness of this interpretation is supported by the findings on the use of evidential shell nouns (e.g. *evidence*). This thesis discusses the results and findings in the following section.

5.2.2.2. *evidence* in evidential shell-noun use

When establishing general cognitive setup of entities and state of affairs, SS L3 writing uses more evidential shell nouns (e.g. *evidence*), giving a view on the acceptance of knowledge as more frequently associated with 'what is known' rather than 'what is thought' (Hunston, 2013, p.628). Further corroboration of this initial

quantitative impression (see table 4.37 in 4.3.3) comes from a contextualized analysis of ‘*evidentiality*’¹ (Bednarek, 2006), which is one of the epistemic considerations that are linguistically encoded (shell noun *evidence* in this case).

Before moving on, it would be useful to briefly discuss the notions of ‘averral’ and ‘attribution’. An attribution (i.e. attributed statement) is a statement presented as deriving from ‘someone other than the writer’ (Hunston, 2004, p.16), while an averral (i.e. averred statement) is made by ‘the current writer’ (ibid.), who assumes the responsibility for the propositions’ veracity. However, of course, ‘all attribution is also a form of averred’, as it is embedded within an averred statement.

When accounting for the ways of attribution in academic writing, citations are mainly used by the writer as they ‘provide evidence for a proposition’ and ‘indicate where the ideas come from’ (Thompson and Tribble, 2001, p.95). There are two types of citation patterns: integral citation and non-integral (Hyland, 1999; Thompson and Tribble, 2001) (see section 5.1.2.2). Integral citations are those that play explicit roles as clause constituents, while non-integral citations are those outside the sentence. A distinction between these two types of citation patterns is that the former is typically associated with a focus on the researcher, while the latter emphasizes the research or knowledge claim itself.

In terms of averral, which is a category of ‘self’ as source, a distinction is made

¹ It might be helpful to make a distinction between epistemic modality and evidentiality. Epistemic modality refers to ‘matters of knowledge or belief on which basis the speakers express their judgments about states of affairs, events or action’ (Hoye, 2014, p.42). In the case of epistemic modality, the speaker makes a knowledge claim based on his/her ‘world of knowledge and beliefs’ while evidentiality concerns the situations when speakers make a claim base on the ‘sources of information other than the speaker’ (Narrog, 2012, p.11).

between self-sourced with human subjects (e.g. *I claim that...*), hidden averral with non-human subjects (e.g. *this evidence [the preceding text] shows that...*) and hidden averral with unnamed attributee² (e.g. *there is evidence that...*) (see Hunston (2000) for more discussion). This distinction between self-sourced averrals are important for the exploration of the way the ideology is positioned. This is because when writers report their own works (averral), their epistemological stance becomes particularly apparent, since in this case they can ‘manipulate the source to which they attribute a proposition in order either to emphasize or obscure their own responsibility’ (Charles, 2007, p.494).

Furthermore, previous studies have also shown that the decision to use either attribution or averral is influenced by the primary means by which knowledge is achieved. To a certain extent, disciplinary conventions also play a part here (e.g. Hyland, 1999; Thompson and Tribble, 2001). In soft fields, knowledge is ‘advanced recursively’ (Charles, 2007, p.493), while in hard fields, knowledge is primarily advanced through ‘cognitive progression’ (Hyland, 2018, p.16). This progression is constructed on experimentation, which is both manipulated and described by the writer.

5.2.2.2.1. SS L1 writing (*evidence*)

In terms of the use of *evidence* in L1 writing, this thesis confirms a remarkable preference for attribution in the form of integral citation, as observed in the SS L1

² As Hunston (2000, p.178) notes, in general a writer assumes responsibility for what is averred, but delegates responsibility for what is attributed to the *attributee*.

sub-corpus. Specifically, SS L1 writing tends to include the cited author within the sentence structure as the subject (5.64), agent (5.65), part of noun phrase (5.66) or adjunct (5.67).

- (5.64) Verspagen found strong economic **evidence** that catching up in the golden age did indeed involve some convergence in technological competence and reduced technology gaps between countries. (SS, 0053a, L1)
- (5.65) In addition to this, a very recent publication by Yoel Rak, Avishag Ginzburg and Eli Geffen seem to provide further **evidence** that Neanderthals are not closely related to us . (SS, 3016d, L1)
- (5.66) The case of Boart Longyear gives **evidence** that employees tend to work harder in pursuing goals that they have helped set than those that have been assigned to them. (SS, 0212c, 1)
- (5.67) As Chris Stringer illustrates the oldest fossil **evidence** for modern humans in Africa is that 130,000, with the oldest outside being 100,000 years from specimens at Skhul and Qafzeh in Levant. (SS, 3016d, L1)

The upshot of these findings is: the shell-noun uses of *evidence* in SS L1 writing are strongly related to the writer's preference for subordinating their own voice to that of the cited authors in the construal of knowledge attribution. On the other hand, the primary audience/reader of students' writing is, of course, the lecturer who grades the text and expects students to demonstrate 'due diligence, critical-thinking ability to interpret, connect, and synthesize sources appropriately' (Swales, 2014b, p.119). L1 students may also cite to demonstrate their sustained research-based engagement in undertaking assignment, i.e., citation as a 'performance' (Harwood and Petric, 2012). Moreover, sources can serve as 'shorthanded references to procedures' (Hyland, 1999, p.347). Thus, it is not surprising to find that citations are prevalent in the use of

evidence in L1 writing.

5.2.2.2.2. SS L3 writing (*evidence*)

In quite the opposite way, SS L3 writing prefers averral over attribution, and greater emphasis has been placed on the writer's arguments. Perhaps more importantly, SS L3 writers tend to construe the experience of averral in a hidden way through *evidence*'s shell-noun uses. This thesis finds that hidden averral in SS L3 writing mainly occurs in two types: a.) hidden averral with non-human subjects, where the writer delegates responsibility for the proposition to the entity (usually the preceding text) appearing in the subject position (as in (5.68) and (5.69)) b). hidden averral with an unnamed attributee, which usually but not always occurs in the existential *there* construction (as in (5.70) and (5.71)). Notably, there is no occurrence of emphasized averral in L3 writing.

(5.68) ...From this **evidence** submitted throughout this essay it is indicated that the inception of Bretton Woods agreement itself engendered inevitable collapse. (SS, 0399c, L3)

(5.69) I have also conducted interviews with both a police officer and a criminal solicitor and hope to use this **evidence** to provide an insight into the practical use of the right to silence within the trial process. (SS, 0411c, L3)

(5.70) Whilst researchers agree on the benefits learning can have on health as a whole, there is **evidence** that intense activity in study can be detrimental to the health. (SS, 0320c, L3)

(5.71) This is a vast enhancement of tourists to the country no doubtly increasing the visitor spend in Australia by a great figure relishing in the **evidence** that mega events can create economical benefits to a city/country, not within the event itself but within other aspects and areas.

To further explore the modes of knowing of L3 writers, this thesis distinguishes between integral and non-integral citations in L3 writing. The results show that citations are far more likely (all but three) to be placed outside the clause within a parenthesis (as in (5.72) through (5.74)).

- (5.72) The majority of the **evidence** is that the driving force behind the grade shift in ergaster was not meat only vs. plant foods only, but rather a mix of the two elements (Milton/1999; Foley/2001; Vasey & Walker/2001) (SS, 3016b, L3)
- (5.73) Theft remains the common crime between the two groups, being the highest offence committed by both groups (National Statistics Office 2005:2). This **evidence** shows a clear difference between the rates and types of crime committed by girls and boys. (SS, 0320h, L3)
- (5.74) ...member governments made clear that they would resist the gradual transfer of sovereignty to the Community, and that EC decision-making would reflect the continuing primacy of the nation state. (Pollack, 'Theorizing EU Policy-Making' p.17). Therefore Hoffman used this **evidence** to argue that... (SS, 0244j, L3)

Several conclusions may be drawn from this analysis of 'evidentiality' (Bednarek, 2006, p.635). The increasing preference for hidden averrals in *evidence*'s uses in SS L3 writing is consistent with Hyland and Jiang's (2018b, p.375) diachronic corpus study of disciplinary writings, which found that self-citation tends to play a more visible role in social science. This shift towards shell-noun use may be the result of increasing specialization as L3 subject studies become more, for example, ethnological convoluted, methodological collaborative and cases focused. Another interpretation may be that certain social science disciplines, such as economics, have peculiar characteristics influenced by the archetype of the natural sciences, making it 'a largely self-referential intellectual project that is mostly inaccessible to disciplinary

or paradigmatic outsiders' Aistleitner et al.'s (2019, p. 362). Furthermore, the rationality of these explanations is accentuated by the earlier findings of mental shell-noun uses (see section 5.2.2.1), suggesting that the need to build persuasive efficacy through shell-noun uses from 'idea' family seems to be less urgent or compelled in SS L3 writing.

However, since it is the engagements with experimental evidence that are increasing, SS L3 writers are more likely to be constrained to eliminate, as far as possible, references to their own role in the operation of empirical practice and linguistic marking of evidentiality. Therefore, hidden averrals with non-human subjects are more likely to be applied by them. As for hidden averrals with an unnamed attributee, this is a point we will return to in our discussion of the colligation of the shell-noun phrase *evidence that*.

Another feature worth noting is the inclination of non-integral citations in SS L3 writing. Whether or not SS L3 writing uses more non-integral citation than L1 writing needs further investigations, but it is a possibility. In fact, the increasing trend of non-integral citation was also been found in other corpus studies, such as those by Nesi (2014; 2021) and Hyland (1999). Nesi (2014, 2021) examined students' uses of citations in the BAWE corpus and found that non-integral citation was much more common than integral citation in every domain, except for physical science,. Hyland (1999) noted a similar general preference for non-integral citation across science and social science research articles, with the exception of philosophy (see section 5.1.2). Furthermore, this finding also echoes the results of Swales' (2014b) study on

undergraduate and graduate writings in the MICUSP corpus, which found that non-integral citation increases steadily across the levels of study examined in the writings.

Since non-integral citations do not require a ‘reporting signal’ carried by an adjunct (e.g. according to X...) or a clause (e.g X claims that...), they are less explicit in representing the relationship between citing and the cited texts. The increasing tendency of non-integral citation in *evidence*’s shell-noun uses might indicate that SS writers have progressed from being inexperienced academic writers to becoming relatively more experienced community insiders. These writers tend to lay emphasis on the cited contents about the topic under discussion instead of the authoritativeness of the agent of the previous research. Thus, they are more likely to write for the same purpose as research writers, ‘to convince the readers that ... knowledge claims are justifiable and significant’ (Bloch, 2010, p.222).

Another possible explanation is that shell noun uses containing non-integral citations ‘can be less grammatically complex’ (Nesi, 2021, p.13). The SS L3 writers’ rhetorical choice of non-integral citation may be a result of its simplicity and ease in contributing to intertextuality. What is perhaps more interesting is that this choice may reflect some aspect of hard science orientation in maintaining the legitimacy of disciplinary knowledge. Previous studies (e.g. Hyland, 1999; Thompson and Tribble, 2001) show that hard fields tend to show a relatively higher use of non-integral forms, which is helpful in highlighting knowledge claims that are based on non-contingent attestations with strict procedures rather than the persuasive efficacy brought by

academic authorities in the disciplinary community.

However, this is not to say that SS L3 writing has lost the discursive characteristics of soft knowledge fields. In fact, SS L3 writing tends to strengthen the persuasiveness of their knowledge claims in a more implicit way. This feature will be discussed in the following section. First, Table 5.34 below presents the distributions of syntactic patterns for *evidence* in SS L3 writing:

Table 5-34

Table 5.34: Distribution of syntactic patterns for *evidence* in SS L3 sub-corpus

SS L3		Freq. (%)
Shell-noun syntactic pattern		
SN- <i>that</i> post nominal		82%
<i>th</i> -SN		18%

It is clear that the most pronounced feature concerns the dominant popularity of the *evidence-that* post nominal pattern in SS L3 writing. As noted earlier (see section 4.2), the *evidence-that* pattern is syntactically versatile and can function as ‘all types of clause constituents’ (Schmid, 2000. p.330). It is for this reason that *evidence-that* is very useful for creating logical coherence, forming discursial cohesion, and shaping texts to be familiar, plausible and persuasive to the maximum extent possible. Moreover, as discussed earlier in section 4.2, the *evidence-that* pattern allows writers to create arguments that allow them to take ownership of their knowledge claims while remaining sensitivity to the disciplinary preferred ways of modes of knowing. The frequency and ubiquity of this syntactic pattern in SS L3 writing testify to the heightened awareness of these useful functions that it offers them, a typical example of this point is shown below in (5.75):

- (5.75) However, the **evidence** that the court is motivated in its decision by a distaste for tax avoidance helps to bolster claim that the nature of debates surrounding the tax avoidance has mutated since Re Chapman. (SS, 0397b, L3)

The essay from which this extract is taken is about the mega event's influence on the host city. There are two shell-noun uses in the form of SN-*that* syntactic pattern, respectively. The post-nominal *that*-clause headed by the shell noun *evidence* (*the evidence that ...*) can be interpreted as a case of hidden averral, conveying an indication of logic basis for the knowledge claim (proof with caused modalization: this makes it certain that X). This choice of shell noun can be understood as rhetorical since the writer could have chosen projection (see section 4.2) by using shell nouns such as *belief*, *idea* or *viewpoint*. In this case, presenting the information as 'what is known' rather than 'what is thought' (Hunston, 2013, p.625) may have seemed more effective and reasonable to the writer. By contrast, the other post-nominal clause headed by the shell noun *claim* (*claim that ...*) can be interpreted as a case of embedded projection of ideas (with the highest level of mediation on the part of the writer). The grammar of these two shell-noun uses is adopted by the writer to express succinctly the development of a proposition (*claim that...*) from suspicion ('what is thought') to something that is solidified/supported/confirmed by evidence ('what is known').

It might be plausible to hypothesize a connection between the writer's implied epistemological stance towards 'what is known' (e.g. *evidence*) and 'what is thought' (e.g. *claim*), as well as the grammatical form surrounding their expressions (e.g.

SN-that). However, predicting what that connection might be is less easy. The underlying mechanism entailed by such a flow of *SN-that* pattern accompanied with *evidence* and *claim* respectively might indicate an acceptance of the self-sourced argument made by the writer as no longer open to questioning. This explanation is probable, as Hyland and Jiang (2019b, p.77) observe that writers in soft knowledge fields largely use verbs that refer to discourse activities in referring to knowledge source, involving the expression of arguments and claims, such as *claim*, *discuss*, *argue* and so on. Thus, this allows writers to ‘discursively explore issues’.

In addition, if we look back at example (5.75), it is clear that the logico-semantic relation between these two *SN-that*-clauses can be interpreted as a case of expansion (elaboration), as these ‘two entities are clearly related to one another’ (see section 2.5 and section 4.2.1) (Halliday and Matthiessen, 1999, p. 117). This type of elaborated clausal orientation typically reveals the epistemology and ideology of soft knowledge disciplinary domains (e.g. Gray, 2015a, p.60; Staples et al., 2016, p.168; Benitez-Castro, 2021, p.133).

Therefore, what these observations suggest seem clear that the prevalence of *evidence-that* in L3 writing is also related to its discourse- and argument-oriented behaviour of *evidence*. The aim is to support/confirm/corroborate ideas, claims and so on by presenting evidence and thus in turn developing a body of new knowledge whereby writers demonstrate their critical thinking skills and persuasive endeavors in rhetorically crafting well-reasoned academic arguments.

Furthermore, the persuasive and elaborated style of writing can also be found in

L3 writing's colligation of *evidence that*. A close analysis of the colligation of *evidence that* in L3 writing shows that the existential *there* construction frequently colligates with the two-word shell-noun phrase *evidence that* (7 out of 9). This word, *there*, behaving as the grammatical subject, positions *evidence* further into the sentence as an notional (actual) subject. It has often been seen as 'non-referential, dummy and empty' (Celce-Murcia and Larsen-Freeman, 1999, p.444), in that it merely 'fills the space in a sentence without contributing to its meaning' (Hall and Birkerts, 2007, p.393).

One aspect of the colligation of shell-noun phrase *evidence that* that has interested the researcher is what kind of rhetorical functions are realized by this colligation. Therefore, while not a central part of this data analysis, it is also worth looking at the rhetorical function of existence *there* structure in academic writing, as it is useful in understanding the function of the shell noun *evidence*.

In terms of the pragmatic function of the existential *there* structure, Lambrecht (1996, p.177) suggests it as 'that of presenting or introducing a referent into the 'place' or 'scene' of the discourse and thereby of raising it into the addressee's consciousness, rather than of asserting its mere existence'. This claim corresponds to Thompson's (2014, p.166) observation that 'existential *there* functions as a 'pass option, pointing forward to the topic of the clause, and, in many cases, of the following stretch of text'. In addition to this main function, studies have recognized two additional functions.

One is the function of enumeration (e.g. Huckin and Pesante, 1988; Celce-Murcia and Larsen-Freeman, 1999; Biber et al, 1999; Carter-Thomas and Rowley-Jolivet,

2001) This signalling role is described as ‘a springboard in developing the text when it is used to introduce a series of elements’ (Biber et al, 1999, p. 952) and contributes to the organization of a cohesive and coherent flow of information. The other is to summarize or reiterate the given information in the previous discourse (e.g. Carter-Thomas and Rowley-Jolivet, 2001; Hiltunen, 2010; Palacios-Martínez and Martínez-Insua, 2006). That is to say, while the main function of the existential *there* structure is to state the (non)existence of certain things, the status of the information expressed in the notional subject is not necessarily always new.

Interestingly, the shell-noun use of *evidence* in SS L3 writing shows that apart from the most prominent discourse function (emphasis on the existence), the other two relatively peripheral functions are fully exploited as well. In other words, all three discourse functions in the context of the shell-noun use of *evidence that* can be simultaneously realized, as shown in (5.76) drawn from the SS L3 sub-corpus:

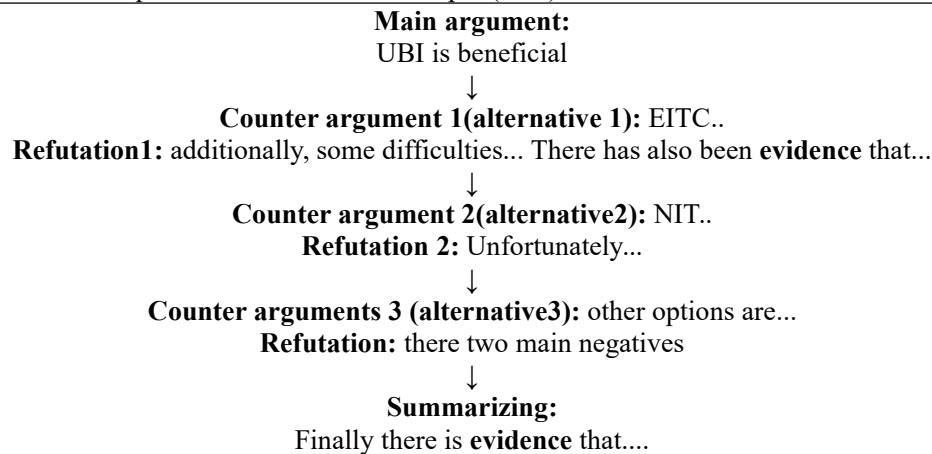
(5.76) I am going to compare the merits of my argument with those of other supporters and then attempt to defend a UBI against the many objections and alternatives which have been suggested.An alternative which America has chosen to use is the Earned Income Tax Credit or EITC.....Additionally the scheme has encountered difficulties.....There has also been **evidence** that the increase in labor supply which it has affected has led to a fall in hourly wages among those eligible for the credit. Along similar lines is the negative income tax scheme (NIT)...Unfortunately this scheme also means ...Other options are a wage cost subsidy or a marginal wage cost subsidy.....There are two main negatives to these proposals.....Finally there is **evidence** that ‘creating an extra job is more expensive than paying an unemployment benefit or social assistance, so a rise of the tax burden would be necessary to balance the government budget’ (De Beer, P. (2000). (SS, 0244k, L3)

The essay from which this extract is taken is about the arguments for and against

Unconditional Basic Income (henceforth UBI), a political issue. The text starts with a statement of its aim and main argument, stating that UBI is irreplaceable and beneficial to society and individuals. This is followed by several sub-counterarguments, which are refuted by the writer's arguments from various aspects of problems related to these counterarguments. This, in turn, is followed by a conclusion that expresses a piece of final evidence that resonates with the writer's main argument. The sequence of the discourse is shown in Table 5.35 below:

Table 5-35

Table 5.35: Sequence of discourse of example (5.76)



Regarding the two occurrences of shell-noun uses of *evidence*, the first use occurs in the refutation section, and the second appears when the writer summarizes the given information, marking a temporary termination of the discourse. Specifically, the first shell-nouns use of *evidence* in existential *there* construction marks the enumeration of a series of refutations, which is indicated by *also*. After a series of refutations with counterarguments, in the conclusion part of the paragraph, the second *evidence* colligating with existential *there* construction shows a function of

summarizing, which is indicated by *finally*.

More importantly, the lexical realization of these two *evidence* takes the form of *that*-complement clause, colligating with the existential *there* construction. However, the evidentiality conveyed by these two types of shell-noun uses is different. This subtle difference can be seen from what Groom (2000) refers to as ‘propositional responsibility’. Speaking of which, the shell noun use of the first *evidence* allows the writer to delegate responsibility for the proposition (e.g. *that the increase in labor supply...*) to an unnamed attributee (i.e. hidden averral with an unnamed attributee), which is implied in a general consensus. Thus, although this form constructs an objective and impersonal tone, the role of the writer is not entirely concealed. By contrast, the shell noun use of the second *evidence* leads the reader to understand that it is the preceding text that is responsible for the proposition advanced (*that ‘creating an extra job is...’*), and the propositional responsibility is delegated to the author³, through the use of non-integral citation.

In conclusion, the skilled exploitation of the complex interplay between averral, attribution and *evidence*’s shell-noun uses reflects the writer’s ‘more personal construction of knowledge’ (e.g. Charles, 2007, p.499). This discursive and interpretative feature is part of the ideology and epistemology that is acceptable and appropriate to their disciplinary community (SS). Table 5.36 below shows the rhetorical function of *evidence* used in colligation with existential *there* construction.

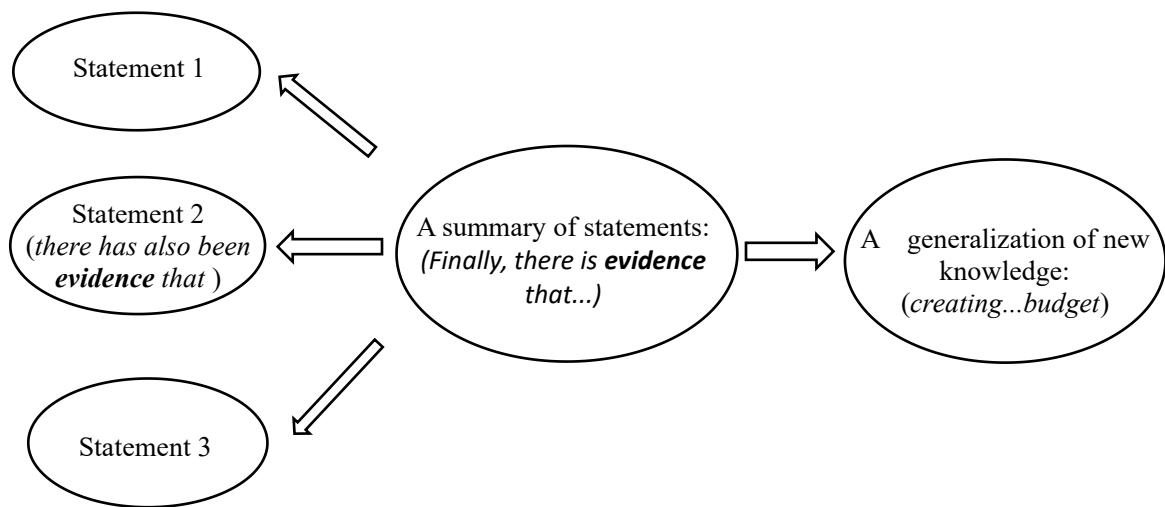
This usage may illustrate certain aspects of what constitutes knowledge (e.g. the self

³ Following Thompson and Yiyun (1991), Hyland (1999) and Thompson and Tribble (2001), this thesis uses the terms ‘writer’ and ‘author’ to refer to the current student writer and the cited antecedent author, respectively.

source/basis of knowledge: evidentiality), and the way in which it is constructed (e.g. knowledge is advanced recursively in an obscure or hidden way).

Furthermore, these two shell-noun uses of *evidence*, in combination with existential *there* construction, enable the writer to make knowledge-claims more clearly and intelligible by adhering to the ‘end-weight principle’ in English (e.g. Biber et al., 1999). This means that the writer may feel a greater need to anticipate the readers’ processing needs at critical points and guide them towards a cohesive and coherent interpretation of message.

Table 5-36: *Evidence’s* uses in example (5.76):



Based on these analyses, it has been shown that the colligation of *evidence that* and the existential *there* construction substantially helps writers meet the adequacy and acceptability conditions of persuasion (Hyland, 1998). This colligation not only assists writers in processing and interpreting the text in accordance with their epistemological understanding and expectations, but also enable them to construct a

seemingly objective and impersonal stance towards what is proposed in the discourse.

In this regard, this thesis considers that L3 texts show a tendency towards a more objective and impersonal style of scientific writing that gives less prominence to the authors. However, the shell-noun use of *evidence* nevertheless remains an important rhetorical tool for writers to construct new knowledge through implicit elaboration, reference and interpretation. In L3 writing, *Evidence*'s persuasive power is related not only to the semantic property of shell-noun hood (e.g. Schmid's (2000) 'Evidence' family), but also to the choice of colligations accompanying the noun and their discursal expansion in the argument. This may be a result of the development of L3 students' linguistic repertoires.

5.2.2.3. *solution* in volitional shell-noun use

Within the volitional shell noun category, it is the use of *solution* and *concern* which have changed remarkably markedly (see table 4.37 in section 4.3.3).

5.2.2.3.1. SS L1 writing (*solution*)

Regarding the shell-noun use of *solution* in SS L1 writing, it primarily occurred in the anaphoric *th*-SN pattern preceded by the demonstrative determiner *this*. Two notable features of the anaphoric shell-noun use of *solution* in SS L1 writing are: 1). it often refers to an argument cited by other researchers, and 2). the shell noun phrase *this solution* makes an abrupt shift of focus, as shown in (5.77):

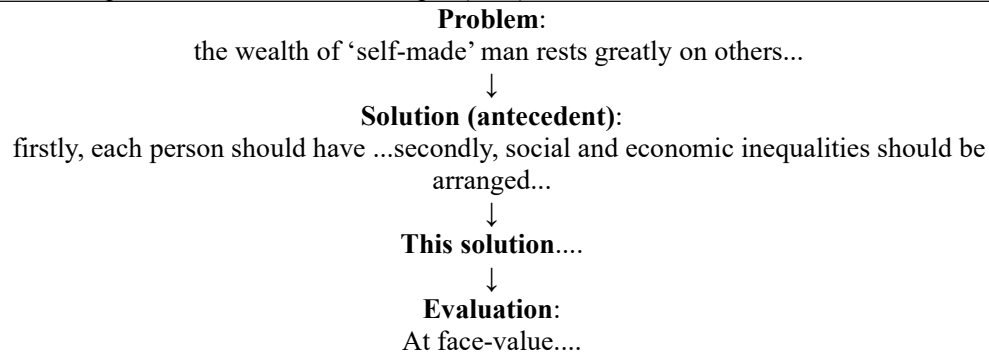
(5.77). However as Goodwin (1999, p.232) perceptively points out the wealth of a so-called 'self-made man' rests greatly on the hard work of his employee's, on the knowledge bequeathed to him by others, on the technological and social advancements attributable to others and partly on fortuitous fluctuations in the market. Thus if he is not solely responsible for what he has yielded but instead much of his success is rooted in the contributions of wider society, it seems questionable whether he should be entitled to the totality of it.

Rawls' conception of justice is apparently derived from a thought experiment in which a group of rational individuals are put behind a 'veil of ignorance' rendering them unable to determine what social or economic position they occupy. If asked to choose the principles for the design of a society that they would most like to find themselves in, Rawls argues they would all arrive at the same conclusion crystallised in the two statements below. Firstly and most importantly, that as a 'primary good', each person should have as much freedom as is compatible with freedom for all. Secondly, social and economic inequalities should be arranged in such a way that the greatest benefit goes to the least advantaged, and attached to offices and positions open to all under fair conditions of equal opportunity. This **solution** is the maximin solution to the problem of social justice as it calls for the maximising of the minimum level of wealth in society. At face-value Rawls' proposal seems to be more workable than that of Nozick's for the simple reason that....However, it seems important to interpret poverty in relative terms if the crux of one's theory is based on the maximisation of freedom. (SS, 0075j, L1)

The essay from which the extract is taken, is about social justice and liberal political theory. The meaning of *solution* is expressed as the reported comment (underlined) in the preceding discourse. The shell noun phrase *this solution* has anaphoric reference and is used in juxtaposition with problem (*problem of social justice*) in the same sentence. Viewed in larger context, it should be noted that the shell noun phrase *this solution* occurs in the upper body part of the essay and functions to temporarily terminate the discourse. A termination can be inferred since the sentence where *this solution* is thematised is used as a kind of topic sentence which shifts the discourse to a lengthy stretch of evaluation for Rawl's proposition, namely, the solution in the preceding discourse. The writer seems to use what Flowerdew (2008) terms a 'Problem and solution text pattern'. This sequence of discourse is shown in Table (5.37) below:

Table 5-37

Table 5.37: Sequence of discourse of example (5.77)



However, in this case, although the writing is crafted in a clear Problem-Solution pattern, it 'doesn't sound quite right' (Flowerdew, 2008, p. 132), for want of a clear indication of the 'problem' or better expressing of their concerns. Strictly speaking, the problem segment (wavy-lined) is a hypothetical situation based on Goodwin's perceptible arguments and the writer's inference, rather than a generally acknowledged problem based on evidence. Even if the argument was valid, the problem that needs to be addressed seems to be social inequality, which is economically presented as the disproportionately distributed social wealth. However, it is only the phrase *the problem of social justice* that refers to the phenomena of social inequalities. In other words, there is an ambiguity caused by inconsistency in meaning. This ambiguity thus conveys an impression that the essay's development of the argument unfolds less logically. Clearer phrases such as *the problem of the existence of inequality* or *the problem of how to realize social justice* would be helpful in eliminating such ambiguity, which in turn renders the anaphoric shell noun phrase

that is juxtaposed with the problem to form a natural sequence of problem-solution text pattern (Flowerdew, 2008).

5.2.2.3.2. SS L3 writing (*solution*)

In SS L3 writing, on the other hand, the most common type of pattern is cataphoric, where *solution* is thematised and has a referent in the succeeding context (SN-*be-to*-clause). The cataphoric shell-noun use of *solution* is almost entirely L3-corpus specific. This use in SS L3 writing is characterized by three features: 1). used at the end of the text, the writer makes a summary by presenting a generalized proposition based on inferences drawn from specific sub-solutions, and the text thus forms a clear Problem-Solution text pattern; 2). used in the body part of the text, the writer specifies criteria for a solution to be met; 3). the source of a solution is always the writer themselves. The first type of cataphoric shell-noun use of *solution*, which is placed at the end of the text, is shown in example (5.78) below:

(5.78) In fact, human-wildlife conflicts are multifaceted and assume a variety of forms.This occurred in Uganda when the Virunga Mountains were opened up to tourist groups for gorilla viewing. The gorillas became habituated to the tourist groups and no longer harbouring a fear for humans they came down the mountain, during the night to raid the crops of the local farmers and often became aggressive and violent when approached.....human - elephant conflict in the Waza-Logone Region in Cameroon, West Africa..foxes apparent tendencies for feeding frenzies in hencoops, upset this balance and incur the anger of the farmer and establish a conflict.idea called 'commensalism,' where people-wildlife conflicts have adapted to perform as interfaces, where it seems some wildlife have been able to adapt successfully to humanised environments.....wildlife damage manager [as] a professional 'buffer' between wildlife and humans, protecting humans form animals, while at the same time protecting wildlife from humans.....The **solution** is to adopt community-based management schemes that involve the local population in conservation projects. Finally,it appears

that human-wildlife conflicts are a consequence of wider, far ranging conflicts between human populations whether local, national or international and certainly not restricted to the tropics. (SS, 3001h, L3)

The essay from which this extract is taken, is about the problem of human-wildlife conflict that occurs in different regions. It is first stated that the problem (first wavy-lined) is that human-wildlife conflicts are multifaceted and assume various forms, which is further comprehensively explained by providing several distinct examples based on solid evidence. Then, the essay enters into a solution segment where several solutions are provided by the writer, ranging from new ideas (second wavy-lined) in psychological state of mind to feasible practice in the real-world (third wavy-lined). At the end of the essay, which is the penultimate paragraph, the writer adopts a cataphoric shell-noun use of *solution* to summarize the previously proposed sub-solutions that are placed distantly in a long stretch of discourse and smoothly leads the discourse to a conclusion. This sequence is shown in Table 5.38 below:

Table 5-38

Table 5.38: Sequence of discourse of example (5.78)

Problem:
human-wildlife conflicts are multifaceted and assume a variety of forms
→ Example 1: human-gorilla conflicts in Uganda
→ Example 2: human-elephant conflicts in the Waza-Logone Region in Cameroon, West Africa
→ Example 3: human-fox conflicts
Solution:
→ Solution 1: commensalism
→ Solution 2: wildlife damage manager [as] a professional 'buffer' between wildlife and humans
Summarizing: The solution is to adopt...
Conclusion:
Finally...

The feature of this type of cataphoric shell-noun use of *solution* is that the source

of the solution is always the writer themselves, and the content of the solution is a summary of the discussion in the whole of the preceding discourse with a clearly focused topic.

Furthermore, another type of cataphoric shell-noun use of *solution* occurs in the body part of the text. It is featured by a focus on exploring a compromise between the two argumentative sides to arrive at a solution discussion, rather than actually solving the issue. This type of lexicalization is shown in example (5.79) below:

(5.79) Up to this stage, it has been fully established that capital punishment is justified on its own terms. ...To impose the death penalty on petty crimes such as theft will only diminish the respect for such a penalty that can be abused by certain states....Hence, the death penalty will be more justified when applied to severe crimes like treason, murder, arson, rape, multiple homicides. Beyond these and some others, it is extremely questionable whether there is any necessity or expediency of applying so great a severity to every crime.With all that in mind, it is evident that the capital punishment can be justified on various grounds, and should not be abolished....At the very least, the **solution** should be to remedy the defect, not abolish the system. Thus far, we have established that the murdered victim has the right to exact punishment & there are beneficial consequences of imposing the death penalty, this section will follow on to defend the rightful authority of the state to impose a just capital punishment. (SS, 0137l, L3)

The essay, from which this extract is taken, is a discussion on whether or not it is justified to impose the capital punishment. This cataphoric shell-noun use of *solution* is placed in the body part of the text and focuses on specifying the criteria for a solution to meet, conveying an eclecticism to both sides of the argument. The sequence of the discourse is shown in Table 5.39 below:

Table 5-39

Table 5.39: Sequence of discourse of example (5.79)

Main arguments:capital punishment is justified on its own terms.

↓

Counterargument 1: To impose the death penalty on petty crimes such as theft will only diminish the respect for such a penalty

Refutation 1: ...the death penalty will be more justified when applied to severe crimes...

↓

Counterargument 2: it is extremely questionable whether there is any necessity or expediency of applying so great a severity to every crime

Refutation 2: ...should not be abolished...

↓

Conclusion: the **solution** should be to remedy the defect, not abolish the system

↓

Further defence of the main argument: Thus far, ...this section will follow on to defend the rightful authority of the state to impose a just capital punishment

In this case, the writer seems to use cataphoric shell-noun use of *solution* to temporarily terminate the discussion, in that it is followed by apparent discourse shift markers such as *Thus far...* and *this section will follow on...*

In a nutshell, it is the epistemic and circumstantial contexts of these *solution*'s uses that render SS L3 writing different from SS L1 writing and AH writing. To a certain extent, these uses uncover a discursively- but more technically-oriented characteristics in SS L3 writing. By 'discursively', it is meant that the metadiscursive function of *solution* is shaping an explicit argumentative sequence, labelling text stages, summarizing conclusions, and indicating topic shifts, which is close to what Hyland (2004a, p.138) refers to 'discourse marker'. By 'technicality', it is meant that the source of a solution is the writer's inferences that are drawn from evidence and reasoning. Besides, the focus of solution is not just 'a way of solving a problem or dealing with a difficult situation' (LDCE) but also the epistemic criteria that a solution should meet.

To recapitulate, whilst acknowledging that the evidence is limited, arguably, this thesis proposes that L3 writing is characterized by presenting a clear and persuasive

‘problem-solution’ text pattern, and the source of a solution is usually based on the writer’s logical inferences. By contrast, L1 writing is more likely to be featured by a seemingly unclear and unpersuasive ‘problem-solution’ text pattern. Additionally, a nuanced difference is identified in terms of SS L3 writing’s evaluation of solution. Specifically, in SS L1 writing, an evaluation of a *solution* is independently and elaborately given after a solution had been presented in the preceding segment (as in (5.80)). By contrast, in SS L3 writing, shell noun phrases in which *solution* collocates with an evaluative adjective are integrated into the text in the lexico-grammatical phrase where *solution* occurs, as shown in (5.81) below:

(5.80)This **solution** to the dualistic problem does indeed appear a least, to be a neat all encompassing structural tool. (SS 3016c, L1)

(5.81) One pragmatic **solution** is to carefully revise and redefine a genus category in gradistic terms of the first interpretation above, so that paraphyletic taxa are inadmissible. (SS, 3028d, L3)

This difference reflects that, to a greater or lesser degree, SS L3 writers tend to downplay and circumvent their evaluative roles to highlight the content of *solution* proposed in their writings, subtly conveying an empiricist position (see section 5.2.2.2.2), where their study outcomes are of a practical use.

Although these findings are mainly confined to the *solution*’s uses, they may indicate some important changes of epistemological orientations in SS L3 writing: the first-year texts are perhaps more discourse-oriented, exhibiting a greater tendency to discuss, illustrate, present, interpret the ‘Problem-Solution’ pattern, and evaluate solutions, while the third-year texts are more epistemic-oriented, concentrating on

discovering, exploring, detecting possible solutions, as well as examining criteria for these solutions.

5.2.2.4. Summary: shell-noun uses in *mental* process

The SS L1 writing's overall preference, and conversely, the SS L3 writing's general disfavour of mental process, was quite clear from neat decreases of several sub-categories of mental shell-noun uses, such as conceptual, volitional and emotive shell-noun uses.

Delving deeper into each shell noun, the SS L1 writing's shell-noun use of *idea* indicated the given status of the lexicalized meaning of the noun in the **SN-that** pattern, where it generally would be expected to convey new information. In SS L3 writing, however, *idea* in **SN-that** pattern neatly reveals new status of shelled information, which in turn endows an impression of tightly and hierarchically concatenated textual structure. The use of *evidence* in SS L3 writing was found to be predominantly accompanied by the **SN-that** post-nominal pattern. The prevalence of *evidence*'s uses in this pattern may be explained by the distinct potential of the **SN-that** pattern for strengthening persuasive efficacy. Furthermore, SS L3 writing's preference for non-integral citation and hidden averral perhaps reflects a more hard-science orientation in knowledge construction. Finally, the shell-noun use of *solution* in L3 writing shows less volitional but a higher epistemic aspect in comparison to SS L1 writing.

5.3. Changes in NS

It is noted that the *evidential* shell-noun use under the *mental* process records the most significant rise ($LLV= 12.99$, $p<0.001$) of a massive 220% and the attitudinal shell-noun use under *attributive relational* process shows an approximately 50% rise ($LLV= -8.27$, $p<0.01$) since level 1 of study. The fact that the two most frequent shell nouns in these two types of transitivity process are *problem* and *evidence*, respectively (see section 4.3), makes it worthwhile to focus on these two shell nouns in the following section. Starting from the shell noun *problem*.

5.3.1. *problem* in the construal of *attributive relational* process.

While NS was a disciplinary domain that has been a relatively low user of attitudinal shell nouns in the primary level of study, *problem* in particular has grown (140 percent) in NS L3 writing. This change seems to be an interesting starting point in the exploration of changes in NS L3 writers' modes-of-knowing, as it is possible that the shell noun *problem* is not just used to highlight the comprehensibility of the 'disconformity of a particular fact or event with something desired' (Schmid, 2000, p.122) but also to enhance persuasive efficacy in some way by denoting the writer's negative evaluation.

5.3.1.1. NS L1 writing

Table 5.40 below outlines the frequent lexico-grammatical patterns for *problem*

and their distributions in NS L1 writing. It is worth drawing attention to the concentration of anaphoric shell-noun uses of *problem* in NS L1 writing, as they occupy nearly 70% of all the syntactic patterns.

Table 5-40

Table 5.40: Frequent syntactic patterns for *problem* and their distributions in NS L1 writing.

NS L1		Freq. (%)
Lexico-grammatical patterns		
	<i>th-SN</i>	41%
	<i>SN-be-clause</i>	33%
	<i>th-be-SN</i>	26%

With regards to the shell-noun use of *problem* in the *th-SN* pattern, it generally functions as a discourse marker of the problem-solution pattern (see section 5.2.2.3). An examination of the contexts of *this problem* indicates that where it occurs, the focus tends to be on reviewing classic ways of handling undesirable situations, where they have either failed to solve the problem (as in (5.82)), or have been successful enough to be well-known to the disciplinary insiders (as in (5.83)).

- (5.82) However this design is not always successful. Many people do not use computers, some because they find computers difficult to use others because they do not find computers user friendly. In an attempt to overcome this **problem** software companies like Microsoft have tried to portray their products with a certain image. They have lots of bright colours to appeal to one type of person, claim ease of use to appeal to another, and still say that their product is the most powerful tool to date. The result is that many people get frustrated by too much help, whilst others cannot find the right help for the right places.(NS, 0228c, L1)
- (5.83) The impractical clothing places the NHS at risk of legal proceedings, as an employer it has the responsibility of protecting employees from injury at work, back injuries affect a large proportion of nurses and midwives (Dimond, 2005). This **problem** appeared to be tackled with the recommendation and gradually the introduction of a new uniform recommended by the Royal Collage of Nursing. The option of tunic and trousers is much more practical and gives far greater freedom of movement as does the introduction of scrubs on delivery suite (NS, 3034d, L1)

It turns out that problems marked as belonging to a previous time, either by the tense (such as *have tried* in (5.82)) or aspect of the verb phrase (such as *appeared to be* in (5.83)), are generally used to construe the experience of reviewing and discussing previous cases and solutions. The focus appears to be on establishing a set of well-reasoned arguments in order to build a body of new knowledge through these heuristic reviews and discussions. This kind of NS L1-specific behaviour of *problem* could perhaps reflect NS L1 texts' greater emphasis on presenting problematic situations, discussing and modeling solutions, which is similar to SS L1 texts (see section 5.2.2.3).

Turning now to the *problem* in *th-be-SN*, it is worth drawing attention to the observation that *problem* marked as belonging to the present or a future time is generally associated with circumstances wherein undesirable phenomena occur. These uses of *problem* seem to denote the writer's concern with revealing the importance of the research field under study (as in (5.84)), identifying or defining the problem (as in (5.85)), and presenting or illustrating in detail the nature of the problem (as in (5.86)).

(5.84) The concern is that the developing countries farmers will not be able to afford the seeds for the genetically modified crop. This is a big **problem** because one of the benefits of genetically modified food is it will ensure that there is enough food for the world's population, especially in the developing countries. (NS, 6037e, L1)

(5.85) Just as the familiarity of the nurses' uniform appears the give practitioners rights to the patients it also appears to give the wearer passport to all areas of the hospital without question by staff, security or patients. This is a **problem** when considering the ease with which one can obtain a nurses outfit or a white coat. (NS, 3034d, L1)

- (5.86) Following on from this, a GP may find this group difficult to work with due to service users who have little or no communication or limited knowledge about health issues. This can be a **problem** for both the service user and the GP if the latter has no skilled support. (NS, 3064e, L1)

Furthermore, if we look back to (5.84), for example, the meaning of the *problem* is marked as belonging to a future time by the aspect of the verb phrase *will not be able to*, indicating the hypothetical or potential undesirable situation that has not yet surfaced. In other words, circumstances or situations wherein problems occur are typically unrealis: they are hypothetical and potential cases that are not subject to immediate verification, rather than an actual ones. Unlike aforementioned shell-noun phrases of *this problem* in (5.82) and (5.83), which are both observable negative situations influencing the real world, the *problem* in (5.84) can perhaps be described as occupying a subjective point along a continuum, with the subjective end at one pole and the objective one at the other. It seems, then, that *problem* in *th-be-SN* pattern can be of some use in manipulating the reader's interpretation of preceding discourse and guiding their attention towards the succeeding discourse. It is a useful supplement to rhetorical persuasiveness and textual cohesion.

This observation along with the aforementioned finding concerning *this problem* leads this thesis to draw conclusion that shell-noun use of *problem* in NS L1 writing reflects the more discursive patterns of argument in the soft knowledge fields (e.g. Flowerdew and Forest, 2015, p.180; Staples et al., 2016, p.165; Benitez-Castro, 2021, p.142).

5.3.1.2. NS L3 writing

Turning now to the NS L3 writing, Table 5.41 below outlines the frequent referential functions of *problem* and its distribution in NS L3 writing.

Table 5-41

Table 5.41: Frequent syntactic patterns of *problem* and their distributions in NS L3 writing.

	NS L3	Freq. (%)
Referential function		
SN- <i>be-that</i>		55%
SN- <i>that</i>		15%
<i>th-SN</i>		30%

As shown in Table 5.41, the cataphoric shell-noun use of *problem* is salient in the NS L3 sub-corpus. It is apparent that the shell-noun use of *problem* differs in its preference for the cataphoric referential function. For the sake of a comprehensive comparison between NS L1 and L3 writings, the following analysis focuses not only on the cataphoric but also the anaphoric shell-noun use of *problem* in NS L3 writing.

In respect of anaphoric shell-noun uses of *problem*, while the uses of *problem* in NS L1 writing are almost always associated with a problem-solution pattern (see section 5.3.1.1), in NS L3 writing, they are more likely to be related to epistemic or circumstantial contexts. In other words, the uses of *problem* in L3 writing tend to construe the experience of pinpointing and uncovering the underlying causes behind certain problems and shifting the discourse from a result to a cause, as shown in (5.87):

(5.87) the only movement of the mandible in a lateral movement will cause trauma, and is occasionally so severe that the coronoid process engages the zygomatic arch, locking the jaws in the depressed position. This **problem** occurs due to that the mouth of evolved

carnivores, ...(NS, 6181d, L3)

Similar to *This problem*, for example, in (5.83), the shell noun phrase occurs at a sentence initial position, but otherwise, they are quite different. In this case, the use of ‘*This problem*’ shifts the discourse to a cause rather than a response, as indicated by *due to*. In this case, *This problem* plays an important role in establishing a Result-Cause discourse pattern. It is to be noted that the anaphoric *this problem* functioning as discourse marker of the Result-Cause discourse pattern exclusively occurred in the L3 sub-corpus.

As noted earlier (see section 5.3.1.1), when a *problem* is marked as belonging to a present or future time through verb phrase tense or aspect, it was most frequently then said to be concerned with identifying, defining, interpreting and evaluating problems (see examples (5.84) through (5.86)). However, this is not the case for anaphoric uses of *problem* in NS L3 writing. None of the instances of *problem* in NS L3 writing show any sign of interpretation or evaluation, but instead offer applicable solutions, as shown in (5.88):

(5.88) Should the gene be harmful, to the plant or indeed to humans, the damage this would cause is indeed reason to be cautious. This **problem** can be managed, albeit with difficulty, such as to minimise the risks of contamination. (NS, 0181b, L3)

The essay from which this extract is taken discusses Genetically Modified (GM) food. Similar to aforementioned *problem*’s anaphoric uses (see example (5.84) in section 5.3.1.1), the antecedent of *problem* is marked as belonging to a future time or

hypothetical world, indicated by *should* and *would*, however, is quite different from NS L1's uses, in that the focus is not on identifying the problem itself, but making efforts in offering detailed accounts of how a potential problematic situation could be averted or tackled efficiently and urgently.

There is one additional point that should be made. While NS L1 and L3 *problem*'s uses have one thing in common, which is functioning as discourse markers to signal discourse shifts from a Problem to Solution, a nuanced but important difference is that L1 writing tends to construe the experience of discussing and reviewing solutions that are drawn from others (see example (5.82) and (5.83)), while L3 writing actively construes the experience of offering solutions based on the writer's part (see example (5.88)).

Turning now to the cataphoric shell-noun use of *problem*, which is chiefly related to the SN-*be-that* syntactic pattern. Although this syntactic pattern highlights the shell content in the *that*-clause, the shell noun phrase is also placed in a prominent position through topicalization. Therefore, this specific syntactic pattern has been perceived by some as the 'focus formula' (Tuggy, 1996, p.725) and 'powerful construction' (Hyland and Tse, 2005b, p.124), serving as a useful resource to allow the writer to foreground their evaluation and guide the readers to a preferred interpretation of the information.

A major distinction reflected by *problem*'s uses in the SN-*be-that* syntactic pattern is that the entity marked as a *problem* almost always occurs in the context of its link to established research approaches (as in (5.89)) and a specific aspect of

methods (as in (5.90)).

- (5.89) This mathematical approach to jury deliberation does appear to be an effective way of predicting the effect deliberation will have on the initial set of references. However, a **problem** is that the assumptions of the leniency bias could be explained by the fact that most of the data was collected using students, (NS, 0014e, L3)
- (5.90) Simply having four wavelength specific cones available is known as weak tetrachromacy, but being able to use them is strong tetrachromacy. This however is very difficult to test whilst the person is alive, and behavioural tests do not appear to be sensitive enough to accurately discriminate between tetrachromatic vision and trichromatic vision (Jameson et al, 2001). Regarding the sensitivity of behavioural tests, a **problem** is that the shifted fourth cones are possibly too close to original trichromatic cones to be able to accurately distinguish between them. (NS, 0014d, L3)

This relatively substantial use of *problem* in NS L3 writing reflects the fact that NS L3 writers tend to place considerable emphasis on precision, particularly to ensure the accurate understanding of empirically practical procedures and results. If we look back to (5.90), for example, the essay from which the extract is taken is a study of whether there are any tetrachromats, and this paragraph is a discussion regarding the specific method of testing tetrachromats. The text starts with a situation, stating the difference between weak and strong tetrachromacy. This is followed by a problem segment which actually presents two problematic issues related to the ways of identifying tetrachromats (wavy-lined): 1). the difficulty of executing tests on people who are alive; and 2). the inaccuracy of the behavioural tests. Next, the discourse delves deeper into one element (*sensitivity*) of an aspect of the issues (*behavior tests*). The meaning of *problem* is lexicalized in the *that*-clause (underlined), which further explains the reason for why the sensitivity in behaviour tests is viewed as being

problematic. The writer's emphasis on the specificity and precision of research method can be displayed by the thematic progression below in Table (5.42):

Table 5-42

Table 5.42: Discourse sequence of example (5.90)

T1	→	R1			
Simply...as weak tetrachromacy		...but strong tetrachromacy .			
		↓			
		T2	→	R2	→
		This		however is very difficult to test whilst the person is alive...	→
					↓
					T3
					→
					R3
				A problem regarding the sensitivity...,	is that...

Therefore, another distinction reflected by *problem's* uses in the SN-*be-that* syntactic pattern is the supplements of partitive meaning regarding problem, as seen in example (5.90). As mentioned earlier, this syntactic pattern is featured by its powerful rhetorical function of guiding readers to the preferred interpretation of the upcoming statement and foregrounding the writers' evaluative meaning towards the following information. However, the prepositional phrase containing a partitive meaning creates a 'relation of particularization' (Schmid, 2000, p.122) and thus allows the writers' evaluation to be more impartial and levelheaded (e.g. Liu and Deng, 2017, p.37; Schmid, 2000, p.334).

To corroborate this observation, a closer examination of the noun structures of *problem* in the SN-*be-that* syntactic pattern seems to be helpful here, as the aim is to test whether the results for the *problem* uncover a partitive meaning in shell noun

phrases. Table 5.43 outlines the frequent noun structures and their distribution in NS L3 sub-corpus. Unsurprisingly, the cataphoric shell noun *problem* has a high tendency to be followed by a prepositional phrase. It is in this respect that this thesis argues that evaluations in such a strong construction tend to be more rational and precise.

Table 5-43

Table 5.43: Frequent post-modifiers of cataphoric *problem* and their distribution in NS L3 writing (Search Terms Position: On Right):

NS L3		
Rank	Types of post modifier	Freq.(%)
1	Prepositional phrase	71.4%
2	No postmodifier	28.5%
3	Other	0.1%

Further examination of these post-modifiers suggests that this rationality and precision is almost exclusively realized by specifying the method/approach/procedure to which the problem belongs, as seen in examples (5.91) and (5.92):

(5.91) The **problem** with the methodology is that parents may not be objective when observing and recording their child's behaviour and sleep problems. (NS, 0421a, L3)

(5.92) A **problem** with CBT approach is that it requires a radical change in behaviour which is often interpreted as neglectful, abusive and inappropriate by parents and as such, their negative beliefs about the method affect its efficacy (NS, 0421a, L3)

To round off the analysis of *problem*'s cataphoric uses in this section, it is helpful to briefly examine certain attitudinal shell-noun uses that denote positive evaluations. This in turn mainly revolves around two shell noun uses: *advantage* and *importance* in cataphoric uses (e.g. **SN-be-that**; **SN-that**). It must be acknowledged that the subset of examples picked up by a search for *the advantage that* or *the importance is that*

may not be representative of the whole. It is also true that not all positive evaluations are explicitly labelled as such. On the other hand, it can be argued that what this thesis is looking for is not *advantage* or *importance* per se, but the information which is explicitly labelled as being positive. Possible searches, then, centre on the noun *advantage* and *importance*.

The aim was to test whether the results for the *advantage* and *importance* were borne out on this wider scale: that the statements labelled *advantage* and *importance* occurred in the contexts of their link to empirical procedures/approaches/methods/results. In both cases, the answer is affirmative. The study of *advantage* finds eight instances (over half of all the returns including non shell-noun uses) where the evaluation is focused on practical approaches (as in (5.93)) and experimental/engineering methods (as in (5.94)):

- (5.93) Introduction of a uniform national business rate has the **advantage** that it does not distort the location of industrial, commercial and financial undertakings. (NS, 3134g, L3)
- (5.94) The main **advantage** to open architecture is that it allows the design of a core that verifies system specifications and is compatible with other developments, sharing know-how and experiences with other designers. (NS, 6101c, L3)

There are six instances (over half of all the returns including non shell-noun uses) where *importance* is used as an evaluation towards other empirical results and practices, as shown in examples (5.95) and (5.96):

- (5.95) The **importance** of Cohn et al's results is that it demonstrates the expression of two

different types of cell, those containing one type of photopigment and those containing a slightly shifted variant. (NS, 0014d, L3)

(5.96) ... the **importance** of heterogeneity is that the factors of happiness are neither the same nor equally important for everybody. (NS, 0383a, L3)

To summarize this brief study of the attitudinal shell-noun uses that indicate the attributes of *advantage* and *importance*: defined and specific procedures and methods do seem to be a site of evaluation—indeed, it seems that the significance of specific empirical procedures and methods is quite likely to be evaluated and labelled as *advantage* and *importance*. In addition, the shell-noun uses of *advantage* and *importance* are probably related to rhetorical persuasion, as the desideratum of these uses is the writer's evaluation of an assumption in the conventional experimental activities in terms of their applicability, reasonableness, correctness and so on. The reason might be related to the writer's guidance to the reader on what the writer acknowledges as reasonableness and correctness or considers worth attending to. This is not always the case, however. All in all, this investigation has provided corroboration for the aforementioned observations based on the close study of *problem*.

5.3.1.3. Summary: shell nouns in *attributive relational* process

The variation between L1 and L3 writings is most evident in the colligations of *problem* in this case. Specifically, the shell noun *problem* in NS L1 writing is strongly primed for the *th-SN* syntactic pattern for anaphoric uses, while the cataphoric shell-noun uses of *problem* predominates in NS L3 writing, typically *problem is that*.

In NS L1 writing, the shell noun *problem*, through its typical anaphoric realizations, functions chiefly to shift discourse topics and prefers to construe the experience of reviewing and recapitulating past problems and solutions. This seems to distance itself from the purely empirically-based, procedure-focused meanings evident in hard science (e.g. Hyland, 2008; Jiang and Hyland, 2015), and instead shows uses closer to an argumentative ‘discursive marker’ (Botley, 2006, p.98) that aims to alink one discourse to the next. Therefore, the interpretative and discursive rationale indicated by NS L1 writing leads itself closer to soft field, as *problem*’s uses are aimed at discussing the circumstances wherein particular problems might occur and cohesively linking.

In NS L3 writing, *problem*’s anaphoric uses reflect writers’ practical know-how knowledge concerning the underlying causes/reasons behind these problems and methods to prevent and mitigate problems in their routine empirical practices. Furthermore, the shell noun *problem* in NS L3 writing shows a strong association with the SN-*be-that* pattern. More importantly, these shell-noun uses of *problem* are almost always linked to the reference of established procedures, defined methods or specific aspects of objects, models, equipment and materials. The focus appears to be on assessing their utility and applicability to reveal pitfalls in highly standardized practices. A possible explanation for such an emphasis might be related to the intrinsic nature of hard science fields or perhaps the types of modules in students’ assignments. Additionally, *problem*’s uses show that rationality and precision of assessment are highly valued by L3 writers in deploying this powerful evaluative syntactic pattern.

The increased shell-noun uses of *problem* in the **SN-be-that** syntactic pattern may, in turn, indicate NS L3 writers' greater confidence in making efforts to build alignments with their readers.

To conclude, while disciplinary writing in hard fields remains the prototypical exemplar for representing disciplinary knowledge and expressing meanings in an objective and impersonal way, it also circumvents affective expression and subordinats the authority of the individual to the authority of the text (e.g. Charles, 2007; Liu and Deng, 2017; Hyland and Jiang, 2019b; Nesi, 2021), NS writers in L3 appear increasingly inclined to exploit attitudinal nouns, such as *problem*, to critically evaluate the disciplinary knowledge nucleus, such as methods and approaches, with greater confidence and willingness than in their primary level of study.

5.3.2. *evidence* in the construal of *mental process*

It can be seen from Table 4.37 in section 4.3 that within the mental process category, NS writing has followed a similar path to SS writing. That is, NS L3 writing increasingly adopted evidential shell-noun uses, such as *evidence*, with a massive 220% increase ($LLV=-6.16, p<0.01$). The increase of shell-noun use of *evidence* is from a relatively low base, reflecting a level of study where the NS L1 writers were in their infancy and lacked sufficient reserves of manipulable resources of evidential knowledge they had in higher levels of study.

Moreover, it might be an indication that NS L3 writers are inclined to present certainty that can withstand the rigors of falsifiability and thus convey a positivist detachment sense of scientific writing. This increase is straightforward to explain, since research in the hard sciences, particularly in higher levels of study, is more

empirically-oriented. Thus, emphasis on evidence plays a more important role in conveying the grounded, experimental basis of academic arguments.

In terms of the syntactic patterns, while the *SN-that* post nominal syntactic pattern predominates NS L3 writing with the highest frequency, the anaphoric syntactic pattern *th-SN* records the most significant increase, a massive 337%. The Table 5.44 below shows the changes in *evidence* patterns.

Table 5-44

Table 5.44: Changes of shell noun item *evidence* syntactic patterns across levels of study in NS (normalized per million words, LLV based on raw frequency)

		NS		
	L1 freq.	L3 freq.		LLV
SN- <i>that</i> nominal clause	65	154		-3.59
SN <i>be</i> clause	19	36		-0.5
<i>th-SN</i>	19	83		-4.3*

5.3.2.1. NS L1 writing

An examination of the concordance lines of NS L1 writing suggests five main categories of discourse function in the context of the uses of *evidence*. These main discourse functions are shown in Table 5.45:

Table 5-45

Table 5.45: Discourse functions of *evidence* in NS L1 writing

Type	Description	Example
Existence	assertion that the evidence exists, and are typically realized in the existential there construction.	(5.97) There is also evidence that <u>such changes in housing wealth are a key determinant of consumer spending.</u> (NS,3109c, L1)
Evaluation	evaluation of the evidence either positively or negatively.	(5.98) <u>After a series of paradoxical arguments, since the beginnings of modern science, the Expanding Universe Theory does seem to be well supported by the evidence produced since the 1920's.</u> Most important of this evidence is Hubble's law, and the correspondence of CMBR, in the Big Bang

model. (NS, 6091a, L1)

Cause	identification of the evidence as the cause of something else.	(5.99) <u>Further support comes from evidence that authoritarianism is more pronounced among people with less education (Chrishe, 1954).</u> (NS, 0020b, L1)
Result	identification of the evidence as the result of something else	(5.100) What Hubble did, with the help of his assistant Milton Humason in 1929, was put the findings into a mathematical form to create evidence that <u>the universe is expanding.</u> (NS, 6097e, L1)
Confirmation	recognition of confirmation of evidence.	(5.101) Conway & Berkerian (1987; cited by Oatley, 1996) also hold the opinion that emotionally relevant information is organised into groups of basic emotions. They obtained evidence that <u>between priming an unrelated word and priming an emotion term within a certain emotion group, the latter produced faster reaction times in a lexical decision task.</u> (NS, 0033B, L1)

Next, the following Table 5.46 below outlines the frequent discourse functions and their distributions in NS L1 writing.

Table 5-46

Table 5.46: Frequent discourse functions and their distributions in NS L1 writing

NS L1		
Rank	Type	Frequency(%)
1	evaluation	39%
2	confirmation	23%
3	existence	15%
4	cause	15%
5	result	8%

Of the five discourse functions, NS L1 writing shows a greater preference for the evaluation category, followed by the confirmation category, existence and cause successively. The result category occupies the least proportion of discourse function. Interestingly, the most frequent category, evaluation, is an important feature of writing in the soft knowledge field, such as AH, as shown in (5.102) and (5.103):

(5.102) The first major and probably the most significant piece of **evidence** is that of red shift; this proves that all galaxies are moving away from each other in proportion to their separation distance. (NS, 6094b, L1)

(5.103) Another very convincing piece of **evidence** regarding the expanding universe is that the existence of cosmic microwave background radiation was first observed by Penzias and Wilson in 1965. (NS, 6097e, L1)

Based on examples such as (5.102) and (5.103), it can be proposed that evaluation in NS L1 writing mainly focuses on an assessment of the value of the evidence, and these assessments often indicating agreements or positive appreciations.

In terms of the second most frequent category, confirmation, most of the uses of *evidence* are associated with attribution. That is, the source and reliability of the confirmation of *evidence* are from another author or previous studies (see section 5.2.2.2), as shown in (5.104) and (5.105):

(5.104) This means that the child recognises two rows of 10 coins are the 'same' if they're spaced equally, but if one is stretched then it will contain 'more'. Piaget (1964) saw this as **evidence** that the child lacked the concept of conservation. (NS, 0190d, L1)

(5.105) Conway & Berkerian (1987; cited by Oatley, 1996) also hold the opinion that emotionally relevant information is organised into groups of basic emotions. They obtained **evidence** that between priming an unrelated word and priming an emotion term within a certain emotion group, the latter produced faster reaction times in a lexical decision task. (NS, 0033b, L1)

In terms of citation practice, this shell-noun use of *evidence* is akin to the shell-noun use in soft disciplinary domains, where greater emphasis being placed on integral type of citation (e.g. Hyland, 1999; Thompson and Tribble, 2001). This resemblance can also be observed in NS L1 writing's preference for the existence

category (e.g. existential *there* construction) (as in example (5.97)), which is also strongly associated with the shell-noun use of *evidence* in SS writing (see section 5.2.2.2).

5.3.2.2. NS L3 writing

Drawn from the above set of categories of discourse function, Table 5.47 below outlines the frequent discourse functions and their distributions in the NS L3 writing.

Table 5-47

Table 5.47: Frequent discourse functions and their distributions in NS L3 writing

NS L3		
Rank	Type	Frequency(%)
1	cause	30%
2	existence	25%
3	confirmation	20%
4	evaluation	15%
5	result	10%

The most notable difference concerns the concentration of the cause category in NS L3 writing compared to its second-to-last rank in NS L1 writing. It is further observed that there is a nuanced difference in the cause category between NS L1 and L3 writings. In NS L3 writing, in the majority of cause instances where *evidence* occurs, it is construed as being the basis or cause for other empirical activities and encapsulated by quantitative observations or empirical findings. For example, consider (5.106) drawn from NS L3 sub-corpus first:

(5.106) European data from the Euro-Barometer Survey Series (1975-01, as cited in Frey & Stutzer, 2002) provided data to support the economic view of utility; reporting that 88% of people in the upper quartile of the income range defining themselves as fairly or very happy, in comparison with only 66% in the lower quartile. It would therefore appear to

be the case that the economic belief has failed to account for some significant findings.
This **evidence** has led psychologists to investigate why the happiness-income relationship is not as straightforward as it was originally thought. (NS, 0383a, L3)

In this extract, the writer presents a piece of evidence supported by statistics to show a counter-intuitive phenomenon, and thus initiates a discussion of the reasons for this phenomenon. The meaning of *this evidence* is expressed in the preceding discourse (underlined), which presents a series of quantitative results from an empirical study. *This evidence* picks up on preceding information and connects it to the succeeding result segment, which is another empirical research pertaining to the investigation of the reason behind the non-straightforward relationship between happiness and income. Schematically, it may be presented as shown in Table 5.48:

Table 5-48

Table 5.48: The sequence of discourse in example (5.106)

Cause (quantitative empirical results): ...88% of people...

→This **evidence**

Result (another research activity): has led psychologists to investigate why the happiness-income relationship is not as straightforward as....

However, a distinguishing feature of the cause category in NS L1 writing is that it points a cognitive activity to another cognitive activity. That is, *evidence* is often construed as theoretical concepts or individual arguments, and is linked to other cognitive activities, as shown in (5.107):

(5.107) This calls into question the causality between authoritarianism and prejudice. Further support comes from **evidence** that authoritarianism is more pronounced among people with less education (Chrishe, 1954). (NS, 0020b, L1)

The essay from which this extract is taken discusses the theory of the authoritarian personality. The extract describes one of the reasons why authoritarianism may not arise from social attitudes but from types of upbringing. The text starts with a result, which is the writer's cognitive activity of questioning the causality between authoritarianism and prejudice. The referent is a reported statement from the cited author (underlined), and it function as the basis/cause, as indicated by the phrasal verb *comes from*, which carries a meaning of causality. Shown below in Table 5.49 is the sequence of the discourse:

Table 5-49

Table 5.49: Sequence of discourse in example (5.107)

Result (idea on the writer's part): This calls into question the causality between authoritarianism and prejudice

→Further support comes from **evidence** that

Cause (idea on the other's part): authoritarianism is more pronounced among people with less education (Chrishe, 1954)

One plausible explanation for this subtle difference is that the emphasis of scientific ideology in the two levels of study is probably slightly different. In the case of an account of new knowledge, NS L3 writers attempt to combine a number of scientific piecemeal features or evidence to make available to the readers a picture of the knowledge-making process 'as a path like a sequence of logical steps towards the revelation of a hitherto unknown phenomenon' (Woolgar, 1981, p.262). In addition, the knowledge-making process in L3 writing may be increasingly tied to an emphasis on the empirical over the interpretative, maximizing the importance of the empirical basis and contributing to reliable and strong knowledge claims in the hard sciences. By contrast, the knowledge-making process in NS L1 writing could perhaps be

understood as a rational response to the evidence of the writer's senses, which is based on the say-so of particular individuals. In other words, readers of L1 writing might be tricked into believing in surrounding, convincing explanations. Another alternative explanation might be that the types of module in these two levels are different. NS L1 writing's modules usually concern the theoretically oriented fundamentals of disciplinary knowledge, such as *Topics in Food and Biotechnology* and *Computer Science Roadmap*. Whereas, modules in NS L3 writing often focus on the utility- and applicability-oriented methodological nature of the assignments, such as *applied psychology of aging* and *Food Manufacturing and the Environment*.

Furthermore, to extend the study of NS L3's writing and to examine more detail at *evidence*, this thesis identifies three common types of *evidence*'s anaphoric reference in NS L3 writing: 1). the shell-noun phrase, in sentence-initial subject position, refers anaphorically to the preceding long segment of discourse; 2). the shell-noun phrase, in subject position and preceded by connectives, such as *and*, is lexically realized within the sentence anaphorically; 3). the shell-noun phrase, in an adverbial phrase, is preceded by prepositions, such as *with*, and the shell-noun phrase is encapsulated anaphorically across the sentence boundaries. It is observed that these three realizations of anaphoric *evidence* do not serve the same functions with each of the syntactic patterns concerned. The following section starts with the first type of anaphoric shell-noun use of *evidence*.

From a textual point of view, compared with other syntactic patterns, especially SN-*that* post nominal patterns, the *th*-SN pattern is highly interwoven with the preceding linguistic context, providing writers with a handy means of encapsulating

preceding information and incorporating it into the succeeding ongoing discourse, as shown in example (5.108):

(5.108) It was found from this that heterozygous carriers of the two types of red opsin, delineated more bands of colour in the diffracted spectrum than a control group of non heterozygous trichromatic females. This removes any differences that may be attributed to gender (Jameson et al, 2001) and also implies that the two groups of females are experiencing different perceptions of colour. This also suggests that the heterozygous carriers are experiencing strong tetrachromacy and that their visual system is utilising the fourth photopsin. This **evidence** demonstrates a stronger tendency for heterozygous females to exhibit tetrachromatic colour vision than Jordan and Mollen (1993) found, and it is likely that this is a more sensitive measure of colour perception. (NS, 0014d, L3)

The essay from which the extract is taken is about a discussion on whether there are tetrachromats. *This evidence* is lexicalized in the preceding discourse, which is comprised of three aspects pertaining to evidentiality. The text starts with an overall empirical quantitative result. This is followed by the first finding, which states that any gender differences may be eliminated (first underlined). This is followed by a second finding, which expresses that within the same gender, females experience different perceptions of color (second underlining). This is followed by the third and final finding that strong tetrachromacy is carried by heterozygous females. It is clear that these three empirical findings are logically linked and progressed, from general to specific.

The demonstrative *this* with an attending shell noun *evidence* encapsulates the prior piecemeal information and turns it into generalized knowledge, that is, a general notion of evidence is evoked in the surrounding context. One explanation for *evidence*'s shell-noun behaviour might be related to the ways in which the knowledge is typically constructed. As Becher and Trowler (2001, p.177) point out, the value

attached to the hard knowledge is related to the ways in which information is presented 'to be spread across a wide front rather than clustered round a limited number of problems'. The epistemological focus in the hard fields is on the generalization and utilitarian application of knowledge. Furthermore, if we look back at example (5.108), what is fascinating here is that a comparison of linguistic structures between preceding text (underlined sentences) and *evidence*'s use provides some insights into the epistemic status in the process of scientific knowledge construction. Before moving on, a brief introduction about Latour and Woolgar (1979)'s 5 point schema in which linguistic structures they encountered are transformed into a classification of 'statement types' is in order. They are:

Type 1: statements comprise conjectures or speculations; Type 2: statements contain modalities which draw attention to the generality of available evidence (or the lack of it) sometimes taking the form of tentative suggestions; Type 3: a statement where the modality is constituted by the included reference; Type 4: deletion of modalities leaves a type four statement of fact; Type 5 statements corresponding to a taken-for-granted fact (Latour and Woolgar, 1979, pp.76-81).

Differing from the preceding segments that make less strong knowledge claims, which are comprised of type 1, type 2 and type 3 statements in the form of citations, involving explicit hedging device, such as *may be attributed*, or linguistic devices to

signal model meaning, such as *implies that...* or *suggests that...*, such an *evidence*'s use contributes to the construction of a 'type 4 statement'. This allows the writer not just to present the preceding information as a 'figure of being' rather than 'a figure of sensing' (Halliday and Matthiessen, 1999) (see section 2.5), but also to convey an apparent certainty that is still worth mentioning, rather than having reached the stage of being an established fact that it can go unsaid. The strength of such a knowledge claim is a result of the rhetorical choice, since the writer could have chosen projection by using shell noun *suggestion* (*This suggestion...*), or a hedged knowledge claim by using a hedging device such as *might* (*this evidence might demonstrate that...*). Viewing this flow of linguistic structures from an epistemology perspective (e.g. Lakatos, 1968; 1976ab; 1978), what seems to be clear is that the new knowledge-making process in NS L3 writing is characterized by the activity of predicting and finessing the hypothesis. In other words, what exactly the writer does to signal new knowledge is to constantly predict what will be discovered/suggested and for those predictions to be correct, making a theory or a set of practices to be scientific. Perhaps more importantly, such an *evidence*'s use is consistent with Plappert's (2019, p.173) observation of the scientific discourse (genetics in this case), where 'the most typical strength of knowledge claim found...was type four statement of fact'. What this suggests seems clear that such an *evidence*'s use in NS L3 writing plays an important role in professionalizing the force of a generalized knowledge claim.

It is further observed that the second type of *evidence*'s anaphoric use is related to

the cognitively-oriented transitional links between two types of ‘inferring evidence⁴’ (Willett, 1988, p. 57), as shown in example (5.109):

(5.109) Averaging out over the year and across all foxhunts, each fox killed costs £ 930. Looking at it by region, the cost per fox killed varies from less than £ 100 in six hunts that participate in Wales, but over £ 3,000 for each fox that is killed in seven hunts in the South of England. This is an un-economical cost to culling in the South, and this **evidence**, together with collections and memorabilia that is available, reflects to a certain extent that hunting with dogs can be seen as an unnecessary recreational activity. (AH, 6181b, L3)

The essay from which the extract is taken is a discussion of the relationship between the animal hunting and the social class. The antecedent of *this evidence* is expressed in the preceding discourse (underlined). At first glance, it seems to be a proposition which might be questioned, but it is not true in this case. Viewed in a larger context, it is clear that the meaning of referent is based on statistical results. The statements following the shell-noun phrase (e.g. *together with...reflects to a certain...*) are based on logical reasoning. That is, shell-noun phrase *this evidence* links two types of inferring evidence coherently and thus builds up an interpretative frame for how readers should comprehend the previous message, essentially informing them what implicatures the previous statement carries, and effectively aligns the readers’ comprehension with the writer’s own.

Based on examples such as (5.109), it seems that the writer’s logical inference plays an important role regarding the acceptable perspective on what counts as knowledge. An explanation for this could be that in higher levels of hard fields, an

⁴ In terms of types of inferring evidence, one is specifically marked as involving ‘results’ and the other involves ‘a mental construct only reasoning’ (Willett, 1988, p.57).

account of knowledge is increasingly reliant on ‘logic, the process of proofs and refutations’ (Lakatos, 1976ab). Constantly reinforcing the ties with prior discourse plays an important role in either illustrating the compositions of data, clarifying the rationale of scientific operations, or guiding the readers to appropriate inferences, thus strengthening the reliability of their knowledge claims.

However, while these *evidence*’s uses may play a major role in nuancing knowledge claims, another interesting and increasingly typical role has been identified in NS L3 writing. Speaking of which, whilst at the formal level, *evidence* is consistent with ‘facts or signs that show clearly that something exists or is true’ (LDEC). However, it contains implicature which is not always equated to the truth conditions of the sentence meaning present, as shown in the following example (5.110):

(5.110) They found that verdicts returned were consistent with the order in which the evidence was presented, thus implying that stories are a mediating mechanism in jury decision making. There is however a large problem with this **evidence** regarding its generalisability to real jury decision making processes. In real trials the evidence is not presented in "story order" but is presented in "witness order", which produces a complex disorganization of the evidence. (NS, 0014e, L3)

As this example shows, at first glance, the adverbial phrase (e.g. *with this evidence*) where *evidence* occurs seems to form a statement revealing some ‘facts’ surrounding an entity in order to convey a richness of ‘factivity’ (Latour and Woolgar, 1979) that has been assumed. However, if we look at a larger context, it should be noted that *evidence* does not bring in any epistemic certainty but rather acts as an entity that is evaluated by the writer. The true conditions of the meaning presented by the sentence, in fact, are a doubt about *this evidence*. This is highly significant in

terms of the encoding of knowledge since it would appear that what is expressed in the preceding text is the propositions that these messages may play some form of evidential role but that the exact nature of this falls short of a fully evidential role. What is important about this finding is that this L3-specific *evidence*'s use shows some developments in the knowledge-making process in NS. That is, NS L3 writers tend to be more precise in comprehending scientific facts and more critical in extending the domain of their validity. This reflects a continuous and typical way that knowledge in hard fields accumulates and develops. In addition, this rhetorical choice might reflect NS L3 writers heightened proficiency in the control of linguistic resources necessary for them to engage critically and persuasively with texts and academic communications.

5.3.2.3. Summary: shell-noun uses in *mental* process

To conclude, first, the hard field's overall increasing use of *evidence*, which strengthens the infallibility and solidity of their knowledge claims, corresponds to the results of Poole et al.'s (2019) diachronic study of experimental science articles that identified an increasing use of epistemic markers indexing a greater degree of certainty. A possible explanation might be that as students' knowledge in a particular domain develops and matures, the need for and expectation of knowledge claims indexing whether knowledge claims are valid and justifiable to the disciplinary community becomes more urgent as their levels of study advances. However, this is not to say that there is a simple and straightforward relation between the form of a statement in which *evidence* occurs and the level of certainty it expresses, in fact, NS L3 writers are more likely to consider the acceptability of evidence or

challenge/dispute the theoretical basis of evidence. This convoluted and L3-specific use of *evidence* indicates some epistemic developments in terms of disciplinary knowledge construction.

Furthermore, a contextualized analysis of shell-noun use of *evidence* in NS L1 writing points to a possible tendency that NS writing in the primary level, whilst pertaining to the hard fields, shares some of the interpretative and discursive orientation in soft fields. By contrast, shell-noun uses of *evidence* in NS L3 writing suggest an inclination towards a positivist approach to academic writing. This change is probably due to NS L3 writers' maturation of knowledge base pertaining to a particular domain of study and their increased confidence in making academic arguments based on empirical, complex, incremental knowledge and logical induction/deduction.

5.4. Summary: chapter 5

Chapter 5 has shown how the uses of shell nouns have changed in disciplinary undergraduate writing across levels of study. This chapter has, moreover, discussed variations in the functional work that shell nouns have been asked to do across the levels of study. It has mainly investigated the shell-noun uses in the construal of human experiences into disciplinary specific knowledge and partly the textual behaviors straddling the linguistic levels of noun phrase structure and syntactic functions. These analyses are focused on shell nouns that have statistically significant differences in their frequency figures across levels of study in each disciplinary domain (e.g. *fact, idea, problem, attempt, evidence*).

The chapter has offered evidence that shell-nouns uses are affected by levels of

study from various perspectives, such as ways of knowledge-making, developments of linguistic repertoires, confidence in their path toward disciplinary expertise, change of readerships, types of modules, and so on. All of the findings in this chapter will be discussed further in the next chapter, to suggest pedagogical implications.

Chapter 6 Conclusion

6.1. Introduction

The main aim of this corpus-based study has been to compare three different disciplinary domains in light of a linguistic feature: the deployment of shell nouns in performing metadiscursive and knowledge construal functions. As mentioned above, this research is based on the premises that, firstly, distinct epistemological orientations of certain academic communities can be reflected by the ways in which disciplinary knowledge is construed and communicated through certain linguistic features, and secondly, that linguistic characteristics exhibited by certain academic communities can indicate discipline-related commonalities and differences. Perhaps more importantly, it should be noted here that while this study has foregrounded disciplinary domains as significant sources of institutional influence on communicative practices, at the same time, the present study is aware of the arguments concerning the development of interdisciplinarity, and the dwindling significance of disciplinary boundaries (e.g. Muguero, 2019; Thompson and Hunston, 2019). In addition, other factors, such as local ideologies (e.g. Manathunga and Brew, 2012) and digital technologies in research paradigms and approaches to research work (e.g. Oliver, 2012) are also crucially important to knowledge development. However, as Becher and Trowler et al. (2012, p.246) argue, ‘disciplines have real epistemological characteristics’, which lead to different ways of knowledge

construction and development. The EAP research and practice involved in helping students, especially those who are international, either before or during a programme, to realize their full potential in achieving academic success largely involves acquiring knowledge in a specialized field. Therefore, the disciplines act as not just sources of knowledge, but the foundations for acknowledged and shared communicative practices. This is why sensitivity to understanding how disciplinary specificity is enacted gradually through semantically abstract shell nouns is valuable to student writers rather than the examination of just the disciplinary specificity or technicality per se. In other words, the kernel of academic literacy is not simply a distilled set of cognitive or technical abilities, but a communicative competence which varies with the contexts.

An interesting question emerges from these ideas: that is, how do student writers as members of disciplinary communities, make their linguistic decisions about how to best construct new disciplinary knowledge and present their academic work? Such a question is important because it allows ESAP practitioners to distinguish themselves from other EAP practitioners by understanding how the concept of academic writing as a disciplinary-specific practice, rather than just as a linguistically autonomous object is firmly embedded in students' experiences, lives, disciplines and contexts. In an attempt to find linguistic evidence that might describe disciplinary student writing, this thesis has analyzed the use of shell nouns, by examining three specific research questions:

Question 1: What is the distribution characteristic of the shell nouns in the six

shell-noun syntactic patterns across the three disciplinary domains?

Question 2: How, if at all, do the three disciplinary domains vary, in their construals of disciplinary knowledge according to the distribution of grammatical metaphors manifested by shell-noun constructions across the three corpora?

Question 3. What differences, if any, have occurred in the shell nouns' deployments in the disciplinary student writings to perform metadiscursive and knowledge construal functions across levels of study? Have these differences been consistent across the three disciplinary domains?

To foreshadow the main results found herein, disciplinary variation investigated in this study was highly related to shell-noun uses overall. A summary of the main findings and answers to the proposed research questions are presented in section 6.2. Then, whilst this thesis may hesitate to categorically account for these research results, some possible explanations as regards the causes of these variations and differences in the use of shell nouns are introduced in section 6.3. Following that, suggestions about pedagogical implications of the study are provided in section 6.4. Section 6.5 acknowledges the most important limitations of the study. While several limitations are acknowledged, areas and directions in need of further exploration are proposed in section 6.6. Some final thoughts are presented in section 6.7. This chapter of the thesis concludes the present study.

6.2. Summary of the main findings and answers to the research questions

In chapter 4, the study first examined the distributions of shell-noun items across the three sets of data representing three different disciplinary domains: Arts and Humanities (AH), Natural Science (NS) and Social Science (SS). Regarding question 1, on the distribution characteristics of shell nouns and shell-noun syntactic patterns by disciplinary domain, first, the hard field (e.g. NS) and the soft field (e.g. AH and SS) were found to be remarkably distinct in terms of overall cataphoric shell noun uses. The AH and SS sub-corpora together accounted for approximately 88 percent of the total number of cataphoric shell nouns in the corpus, whereas the NS accounted for only 12 percent. In addition, there was a greater diversity of cataphoric shell nouns in AH and SS sub-corpora than that of NS corpus. This distribution characteristic is an evidence that corroborated the typical soft versus hard divide identified in previous research findings on the study of shell nouns (e.g. Benitez-Castro, 2021; Flowerdew, 2015; Staples et al., 2016). In addition, this thesis has added nuance to this finding: it was found that the three disciplinary domains demonstrated a certain degree of lexical similarity in the use of cataphoric shell nouns and the two soft fields showed a significant similar preference in the choice of cataphoric shell noun types.

Regarding the anaphoric shell nouns, the NS corpus accounted for approximately 50 percent of the total anaphoric shell-noun tokens of the three corpora, whereas the AH and SS sub-corpora combined contributed another half of the anaphoric shell-noun tokens (around 25% for each). Furthermore, the diversity of the anaphoric

shell noun types was roughly equal in the three corpora. This said, the NS corpus was in fact instantiated from a small set of anaphoric shell nouns covering a remarkably larger portion of the total uses (see section 4.2.1). In terms of variations across levels of study, it was found that there was a modest 3% decline in shell nouns across the levels of study without any statistical significance.

Moving now to the syntactic level, the study has yielded significant findings about the distributions of shell-noun syntactic patterns across the three corpora and levels of study. For complement constructions, the NS corpus preferred the use of *to*-infinitival clauses, including **SN-*to*-clause** and **SN-*be-to*-inf** while the use of finite *that* clauses including **SN-*that*-clause** and **SN-*be-that*-clause**, was most significantly preferred in the AH and SS corpora. For anaphoric syntactic patterns, there was a highly uniform preference for *th*-**SN** syntactic pattern across the three corpora, while the *th-be*-**SN** syntactic pattern showed marginal frequency rates in AH and was completely absent from the NS and SS corpora. In regard to the changes across levels of study, the most pronounced and interesting finding is that the **SN-*that*** syntactic pattern increased significantly in the NS corpus, while it showed a remarkable decline in the SS corpus.

Furthermore, notable divergences were found at the semantic cognitive (epistemological) level in the second part of chapter 4, which examined the different distributions of types of knowledge construal instantiated by shell nouns across the three corpora (Question 2). Briefly, for cataphoric shell nouns, *attributive relational* processes were saliently favoured in the NS corpus, while *mental* processes were

favoured in AH and SS corpora. This is expected and in line with the findings about the distribution characteristics of shell-noun complement constructions, as the construal of psychological states tends to derive from the employment of finite *that*-clauses, while *to*-infinitive clauses are adopted to embody modality of varying degrees where associated shell nouns serve to construe the judgments of possibilities and necessities of potential occurrences of the empirical activities. By contrast, in terms of anaphoric shell nouns, the three corpora examined in this study showed a general similarity in the preference of *mental* processes. In respect to changes across levels of study, the two most significant changes of all the types of transitivity process were the decreases of *identifying relational* (decreased by 24%) and *mental* processes (decreased by 41%). While there was considerable variation across disciplinary domains in how frequently shell nouns were used and the types of associated lexicogrammatical patterns and so on, an in-depth analysis was conducted mainly on the ways in which shell nouns serve to reconstruct human experience into knowledge through the disciplinary-specific construals of different transitivity processes within disciplinary domains, as well as across levels of study. This aspect is recognized as central to identifying the types of epistemological roles that the shell nouns played in the process of knowledge construction (see section 2.2).

Subsequently, by focusing on the qualitative results obtained by analyzing examples in the following chapter, this study found that differences related to levels of study were not consistently distributed across the three disciplinary domains. Specifically, chapter 5 delves into an investigation of these inconsistencies in

shell-noun uses through the exploration of their deployment in each of the three disciplinary student writings to perform the metadiscursive and knowledge construal functions (Question 3). The content of chapter 5 stems from shell nouns that occurred in the types of knowledge construal that show statistically significant changes in either of the three corpora. The first part of chapter 5 examined two remarkable changes in the AH corpus: the remarkable decline of the *identifying relational* process construed by *fact* and the increase of the *material* process construed by *attempt*. By examining the shell-noun uses of *fact*, the study found that the shadow of negotiability constantly hangs over the occurrences of *fact* in AH L1 writing. By contrast, AH L3 writing was more interested in construing an objective reality and showed less explicit efforts to shift the conceptual status of the shell contents. Next, a close analysis of *attempt* has also provided several meaningful findings: on the one hand, by examining the verbs in the infinitive *to*-clauses, this study found that AH L1 writing showed a discursive and interpretative tendency while the approach to shell-noun uses of *attempt* in AH L3 writing was likely to be comparatively more pragmatically- and empirically based. On the other hand, the shell-noun use of *attempt* in L3 writing seemed to serve as an important part of the writers' rhetorical armoury to strengthen the persuasiveness of their arguments.

The second part of chapter 5 examined several significant changes that occurred in SS: the decrease of the *identifying relational* process construed by *fact* and the *mental* process construed by *idea*, *evidence* and *solution*. SS L1 writing showed a preference for using the combination of *fact* colligated with the *th-SN* syntactic

pattern to construe the experience of presenting writers' evaluations, demonstrating their logical reasoning and well-reasoned academic arguments, while SS L3 writing favored using *fact-that* clauses to construe the experience of illustrating factual aspect pertaining to the evidential relations or contradictory state of affairs. Interestingly, it was found that the two cognate soft fields diverged in the use of *fact* in the syntactic pattern of **SN-that**: AH L3 writing reflected a trend toward a more scientifically oriented knowledge structure, as the construals of knowledge based on objective condition had increased substantially, while SS L3 writing showed a greater circumspection in the construals of knowledge based on objective conditions.

We now turn to the changes of *mental* process construed by *idea*, *evidence* and *solution* in SS writing. The most prominent and interesting finding in this case concerned the shell noun *evidence* in relation to citation practices. Briefly, the shell-noun use of *evidence* in SS L1 writing was strongly related to the use of integral citations and the construal of knowledge attribution in which the writers preferred to subordinate their own voice to those of the cited authors. By contrast, SS L3 writing preferred non-integral citations and hidden averrals through *evidence*'s shell-noun uses. Moreover, the existential *there* construction was frequently colligated with the two-word shell-noun phrase *evidence that* in SS L3 writing. By analyzing the metadiscursive functions performed by *evidence* in this colligation, this study found that while *evidence*'s uses in L3 writing showed a tendency towards a more scientific style of writing, it was nevertheless still an important rhetorical tool for writers to build interactions with their readers and claim credits for their propositions.

Finally, in the last part of chapter 5, changes in the NS corpus were analyzed focusing on the two shell nouns *problem* and *evidence*. In NS L1 writing, shell noun *problem*, through its typical anaphoric realizations, its chief functions of shifting discourse topics, and its preference for construing the experience of reviewing and recapitulating past problems and solutions, seemed to distance itself from the purely empirically-based, procedure-focused meanings evident in hard science writing, a feature previously recognized as central to professional scientific writing (e.g. Liu and Deng, 2017, Jiang and Hyland, 2018; 2019a; Kim and Crosthwaite, 2019; Benitez-Castro, 2021), and to show instead uses closer to a situation reference (Fraurud, 1992, p.4) or argumentative discursive marker (Botely, 2006, p. 98). By contrast, in NS L3 writing, *problem*'s anaphoric uses reflected writers' practical know-how knowledge concerning the underlying causes/reasons behind these problems and methods in their routine empirical practices.

Turning now to the shell-noun uses of *evidence*, this study examined the different categories of discourse function in its context. The study found that NS L1 writing showed a greater preference for the evaluation and confirmation categories, while NS L3 writing favored the cause category. Additionally, it was observed that, in the confirmation category, the majority of the *evidence* in NS L1 writing was associated with attribution. Perhaps more importantly, these attributions were often took the form of integrated citation, which is akin to the shell-noun uses in soft disciplinary domains, where greater emphasis is placed on integral type of citation (e.g. Hyland, 1999; Thompson, 2001). In NS L3 writing, in the majority of cause instances where

evidence occurred, were construed as being the basis or cause for other empirical activities and encapsulated by quantitative observations or empirical findings. Besides, NS L3 writers were more likely to consider the acceptability of evidence or challenge/dispute its theoretical basis.

6.3. Possible explanations for the differences in disciplinary student writing

The findings of this study have offered substantial evidence that the ways in which knowledge is constructed, produced, communicated, and negotiated are shaped by the intellectual boundaries drawn by different disciplinary frameworks. It is clear that the differences in shell-noun uses across disciplinary domains are mainly a result of the divergence between soft and hard modes of knowing and ways of viewing the world, namely, epistemological variation. Although this explanation may be somewhat of an oversimplification, it is nevertheless the most important and influential one. To put it briefly, hard science is more likely to be oriented towards a scientism, focusing on laboratory methods, empirical approaches, and grounded data and so on. As a result, the textual behaviors of shell nouns in NS writing contribute to a vertical and compressed style of writing, and the knowledge construed by them is related to the descriptions of research objects or contexts, or the specification of logical aspects of models, materials, and problems in the research environment. By contrast, new knowledge in soft science is often based on interpretations and understandings, making shell-noun uses in soft-field writing more discursive and

interpretative in nature (e.g., Hyland and Jiang, 2019a; Omidian and Siyanova-Chanturia, 2021; Eckstein et al., 2022).

However, the explanations for the differences in the uses of shell noun across levels of study within each disciplinary domain are less straightforward. While acknowledging that there are no categorical accounts for these differences, this thesis proposes the following factors that might influence the different uses of shell nouns across levels of study within each disciplinary domain: 1). the various topics of different modules (section 5.2.2 and section 5.3.2); 2). student writers' heightened confidence as members of disciplinary communities over the years (see section 5.3.1 and section 5.2.2); 3). student writers' growing familiarity of disciplinary-specific research methods over the years (see section 5.3.1); 4). increasing scientism in certain soft disciplines, such as applied linguistics, due to a more hard science orientation in their dominant methods and approaches (see section 5.1.1 and section 5.2.1); 5). increasing specialization in social sciences as subject study becomes more methodologically collaborative, ethnologically convoluted, and cases focused (see section 5.2.2); 6). Student writers' proficiency in controlling the linguistic resources necessary for them to critically engage with texts and persuasively participate in academic communications (see section 5.1.3, section 5.2.2 and section 5.3.2); 7). different degrees of disciplinary professionalism possessed by student writers (see section 5.2.2 and section 5.3.2), and different readerships (see section 5.2.2 and section 5.2.1).

6.4. Implications for EAP practitioners and curriculum designers

By investigating disciplinary conventions through the lens of their linguistic realizations in student academic writing, with a focus on shell-noun uses, this thesis has argued that highlighting the disciplinary distinctions that arise from the ways in which student writers construct and communicate knowledge opens up invaluable way to facilitate EAP practitioners' and a particular group of language users' (e.g. potential L2 undergraduate students; pre-tertiary groups of learners) understanding of disciplinary writing. Based on this view, this section proposes some pedagogical suggestions on how to prepare English language learners who are en route to pursuing their academic degrees in English-medium universities and help novice undergraduate students at the beginning stages of their academic lives acquire distinctive disciplinary features in their academic writings. Regarding semantically abstract units such as shell nouns, suggestions for EAP practitioners and ESAP practices are as follows:

Designing Data Driven Learning (henceforth DDL) activities based on authentic corpora (Charles et al., 2009; Charles, 2012; 2014; Chen and Flowerdew, 2018; Crosthwaite 2019) is encouraged for EAP practitioners. For example, to set up a DDL environment in which students are encouraged to adopt the role of active language detectives, discovering frequent phraseologies and syntactic patterns in the exploration of the occurrences of the same shell noun across different disciplines. Besides, this study highlights the importance of creating exercises easily from the concordances, such as cloze exercises. Taking this point into consideration,

WordSmith Tools (version 8.0) (Scott, 2004) might be a good choice, in that it has a function of blanking out the search-words, allowing practitioners to produce several types of cloze tests. For example, Figure 6.1 below shows a screenshot of the function of blanking out the search word (an abstract noun *way* in this case) *WordSmith tools* (version 8.0) in the analysis of BAWE corpus.

Despite the clear empirical evidence of significant language learning gains offered by the use of these traditional concordancers, an aspect that may derail an EAP writing class is that they tend to focus students' attention on certain superficial forms of words (e.g. shell nouns) because they present language patterns in a decontextualized, tabular form (Ädel, 2010). To address this challenge in terms of shell noun instruction in an EAP classroom, two offline traditional corpus analysis tools offer one useful solution: *LancsBox* corpus analysis toolkit (<http://corpora.lancs.ac.uk/lancsbox/>) and *FireAnt* corpus analysis toolkit (<https://www.laurenceanthony.net/software/fireant/>) (Anthony and Hardaker, 2017). The former one incorporate a new visualization method: *GraphColl* (Brezina et al., 2015), to provide students with more contextual information on shell noun uses, while the latter is able to produce network visualization, geositional maps, and time-series plots. In a DDL classroom aimed at teaching shell noun uses, it is essential for EAP practitioners to encourage students to use such toolkits when trying to understand where, why and how, for example, particular shell noun patterns are used, as they allow students to form hypothesis about shell-noun uses based on their first-hand observations.

In addition to traditional offline corpus analysis tools, online corpus analysis tools also provide a useful solution. Popular online corpus analysis tools, such as *CQPWeb* (<https://cqpweb.lancs.ac.uk/>), *English Corpora* (<https://www.english-corpora.org/corpora.asp>) and *Sketch Engine* (<https://www.sketchengine.eu/>) allow students to get a more complete picture of the context surrounding a particular shell noun use example.



Figure 6.1: The function of blanking out the search word (e.g. *way*) in the analysis of BAWE corpus.

Perhaps equally important is to be aware that while DDL has been potentially beneficial in many aspects, it may also present certain challenges, especially for second-language learners who are not familiar with shell noun use and thus may struggle to know what shell nouns to search for or may be overwhelmed by the sheer number of examples of different shell noun uses presented to them. Furthermore, they may also struggle to interpret the shell noun patterns and instances that the corpus

consultation presented to them. This requires EAP practitioners to be well-prepared for the implementation of DDL activities by, for example, systemically selecting which set of categories of shell-noun patterns to focus on. With regard to this aspect, this study provides the basis for what Flowerdew (2002) has referred as ‘pedagogic grammar’ of shell nouns. Context-sensitive analysis of individual shell nouns specific to genres (e.g. essay genre writings), and particular disciplinary domains (e.g. Arts and Humanities, Natural Sciences, Social Sciences) can be useful for the development of more finely tuned local grammars of shell noun patterns in EAP classroom.

1. While inductive pedagogy, such as corpus consultations in aforementioned suggestion, is vital for raising students’ awareness of features in disciplinary writing as a whole, deductive pedagogy should not be dismissed. It is also important for EAP practitioners to build explicit hands-on EAP courses and teaching materials about key disciplinary textual trends and syntactico-semantic features. This is the case, for example, with the direct use of corpora (e.g. Yoon and Jo, 2014; Öztekin and Candan, 2019), which focuses on the influence of corpora on syllabus design and linguistic materials for the purpose of EAP teaching, relying on researchers as providers of relevant linguistic materials. With regard to this aspect, I would argue that corpus studies of this kind have great potential to inform EAP research and practice. First, the findings demonstrated here in this study can provide both an overall picture of how the use of shell nouns contributes to a distinctive disciplinary flavour in texts and a qualitative characterization of similarities and differences along dimensions of variation. Besides, this study provides lists of shell nouns that can be used as a basis

for the development of teaching aids, and also frequency for disciplinary domains that could be immediately incorporated into a specific-purpose syllabus. However, while the lists of shell nouns might be useful, it is important for EAP practitioners to make learners aware that abstract nouns which can function as shell nouns, can also be exophoric. EAP pedagogical treatment of shell nouns is thus required to raise learners' consciousness not just of searching within texts for the referent of a potential shell noun but also outside the texts. Learners need to be guided to have an overall perspective that disciplinarity is actively reflected by shell-noun usage.

2. In previous suggestions, we have talked about the incorporation of DDL methodology into the classroom teaching, corpus-informed learning/teaching materials, syllabus design and so on, it is crucially important for EAP teaching to adopt a text-centered method in a genre-based approach. This is because a genre-based approach provides learners with 'texts of various genres and the grammatical resources needed to pull apart and put together the meanings of these texts. Students take on the role of apprentices, as teachers, in the role of expert text users, model, deconstruct and jointly construct texts' (Cullip, 2000, p.95).

Specifically, it is suggested that EAP practitioners should teach one genre across different disciplines by not only exposing students to canonical models of academic writing but also deconstructing them with particular attention to shell-noun clauses and their textual behaviors. For example, in an EAP class where the teacher uses the genre pedagogy to teach the essay genre, the teacher is encouraged to provide students with three short essay writings from different disciplines (e.g. medical, applied

linguistics, chemistry), in which shell nouns are used effectively. Then, the students, as readers when approaching the texts are requested to specify or think about the intended and or potential readers when reading these texts based on the factors such as shell noun uses. Subsequently, students are asked to identify the referents for shell nouns and discuss the function of these shell nouns. After that, students are asked what texture effect (e.g. less cohesive, less precise or ambiguous) might be achieved if shell nouns were removed. EAP practitioners then explicitly teach students the textual behaviors of these shell noun items in different disciplinary essay genre writings. This is because explicit instruction on nominalization is of clear benefit to the high quality of L2 student academic writing (Hu and Perez, 2022). Besides, the results of Bychkovska's (2021) study demonstrate that targeted instruction on noun phrases may indeed benefit L2 students' syntactic complexity development and lead to the production of genre-appropriate high quality texts.

Taking one step further, students as writers are requested to work either as a group or individually to reconstruct these three essay writings by filling in missing shell nouns from a given list. Next, in order to familiarize students with shell noun uses, they are asked to do a second filling task in which shell nouns are omitted again but this time without any alternatives provided. Finally, feedback is provided by the EAP practitioner on students' work.

At the same time, this study has demonstrated that grammatical patterns are a useful heuristic in the EAP investigation and instruction of disciplinarity in academic writing. Taking shell noun complementation patterns as a heuristic, the analysis of

these patterns reveals some aspects which have yet not been adequately discussed. To be specific, it has been observed that SN-*that* complement pattern plays an equally important role in conveying writer's attitudes as its cohesive function in connecting segments for the forming of a unity of passages. Both of these two roles tend to be linguistic reflections of epistemic perspectives and conventions accepted by writers' particular community concerning what counts as knowledge. This further suggests that disciplinarity would better be interpreted as simultaneous choice made in terms of grammatical patterns and lexical items.

3. Furthermore, another suggestion is the employment of student-centered pedagogy in raising students' awareness of disciplinary features reflected by shell nouns. For example, in a heterogeneous EAP class of students from a range of specialist disciplines, students can be divided into groups according to their disciplinary backgrounds: group 1 contains students from hard disciplinary domains, such as chemistry, biology and so on, while group 2 includes students from soft disciplinary domains, such as applied linguistics, history and so on. Next, the EAP practitioner provides each group with text materials written by writers from disciplines differing from their group members' own disciplines. Within each group, the allocated text contains a range of sentences using a specific shell noun (e.g. *solution, thing, result* and so on) in ways that diverge significantly from typical use in their own disciplinary domains. Then, students as group members are asked first to identify the shell nouns and their corresponding referents and then to discover aspects that might differ from expected patterns and in what ways. Based on their discussions,

group members are required to conclude some main points about the shell noun's function, clausal and textual behavior from the material presented. Lastly, the EAP practitioner highlights key disciplinary syntactico-semantic and textual trends by presenting students with more shell noun examples and facilitating students to discover how the deployment of these nouns diverges in the function of, for example, signalling text patterns across disciplines.

4. Last but not the least, in recent years, Digital Multimodal Composing (DMMC) has been increasingly common in tertiary education due to the advance of educational technology. Kim and Belcher (2020, P.86) define DMMC in a second language writing classroom as 'teaching writing as the social practice of meaning-making using various semiotic tools'. Incorporating DMMC into EAP teaching is useful in helping students to compose texts with an improved level of clarity through images, graphs, drawings and photographs (Kohnke et al., 2021); enabling students to develop critical literacy (Jiang, 2017; Yi et al., 2019; Kohnke et al., 2021); and enhancing vocabulary (Vandommele et al., 2017). This is the reason why it may be useful in helping students become better disciplinary readers and writers. With regard to the issue of shell nouns, it is suggested that EAP practices in teaching academic writing should enact multimodally responsive pedagogies. For example, applying infographics as a way of explaining, say, the grammatical structures, referential directions and the encapsulated segments of shell nouns. This is because infographics, as spatial representation of linear text or information where ideas and concepts are visibly highlighted by graphic devices such as maps and diagrams, have the potential to promote learners'

knowledge on shell nouns' clausal and textual behaviors.

It is worth noting that the use of infographics can also be incorporated into the students' reflective EAP activities in a disciplinary writing classroom. For example, in an EAP course aiming to develop the linguistic repertoire necessary for students to critically and effectively engage in future professions in business design, students can first be provided with two short texts in which shell nouns are efficiently used and inefficiently used, respectively. Then, students can be asked to produce infographics for these two texts by paying particular attention to the functions of shell nouns in signalling textual patterns, referring to information, and so on. Next, students can be required to present and compare these two infographics, and discuss the significance and impact of shell nouns in the process of infographic designing. It is the process of designing and reflecting on infographic that enables EAP practitioners to provide students with heuristics and disciplinary-specific impressions of typical textual patterns influenced by shell-noun uses.

6.5. Limitations of the study

Inevitably, this study has had some limitations. In the following paragraphs, this thesis acknowledges some contentious points related to the study, mainly in regard to the methodology. First and foremost, perhaps the most noticeable limitation of this work stems from the fact that this comparative study has been primarily concerned with disciplinary student writing from three broad disciplinary domains, that is, the patterns of variation across specific sub-disciplines within each disciplinary domain

were left aside without interpretation. However, it turned out to be noteworthy that some of the most interesting findings in this study were given by particular disciplines such as applied linguistics, business and computer science. In this case, an investigation of disciplinary specificity across specific disciplines would have been needed so as to make it possible to analyze more factors contributing to differences in disciplinary student writing and thus provide more well-grounded evidence for the preliminary conclusions reached.

A second major limitation is rooted in the fact that this study chiefly applied automated means of corpus linguistic methods to investigate the shell noun phenomenon, relying on structural tests of shell noun status. Specifically, the data analyzed in this study was based on the retrieval of predefined lexical-grammatical patterns that have already been identified in the literature, these patterns include *SN-that-clause*, *SN-to-infinitive clause*, *SN-be-that-cl*, *SN-be-to-infinitive clause*, *this-SN* and *this-be-SN*. At the same time, however, this reliance has an unfortunate consequence that it fails to reflect shell noun instances that do not occur in typical patterns but still have a prototypical shell noun role. Examples of these minor patterns which are not covered in the literature i). instances typically signalled by punctuation (usually colon, although not always) where shell nouns find their specifics juxtapose in a relation of identity (as in (6.1) to (6.2)). ii). comparative constructions where the specifics of shell nouns are presented by analogy or exemplification (as in (6.3) and (6.4)). iii). adjunct groups where shell nouns occur as part of adjuncts with their specifics in the main body of the clause (as in (6.5) and (6.6)).

- (6.1) It was a pragmatic solution to a commonplace **problem**: families rarely create formal agreements; illustrating Denning as grounded and understanding: not the stereotypical judge. (SS,0209a)
- (6.2) The conquest of Mexico poses a complex **problem** in history: how is it possible that a small band of Spanish explorers were able to conquer a Mesoamerican warlike and powerful society, on their own territory, in the space of just two years? (AH, 6211j)
- (6.3) Finally, it should recruit and select only those candidates who are fit in the desired **criteria**. For example, if an organisation has to survive competition in the market, its strategy will be to make sure that its products are low-cost, better quality, and innovative in comparison to its competitors. (SS, 0136b)
- (6.4) Although black athletes fulfil one element of the hegemonic masculinity (physical prowess), they are prevented from fulfilling other **aspects** such as becoming professionals...(SS,0004c)
- (6.5) However, as a **result** of the growing interest in cultural history following the 1980s 'linguistic turn', recent study, particularly in the so called 'Constructivist' school, has challenged Cold War historians to spread their wings methodologically, and explore the culture of the conflict. (AH, 0005c)
- (6.6) On the other hand, 'the Cusabo welcomed the English...as a welcome alternative to ineffectual Spanish' signifying an overriding difference in their colonies in that the Indians showed a preference for the English. (AH, 0129a)

The occurrence of these additional patterns may be due to the complex nature of the shell noun phenomenon, which is shaped by their context-specific senses in particular texts (Benitez-Castro, 2014, p.467). Furthermore, this is also because shell noun use is a matter of degree which occurs on a cline of context-dependency, which precludes any generalizations about their grammatical, semantic and textual features, as also observed by Ivanic (1991, p.109) and Schmid (2000, p.85). Therefore it is not yet clear how many lexical items and syntactic patterns give rise to a shell-nounhood.

With regard to the present study, the fact that it has not captured all instances

associated with shell noun uses would not pose many problems for the observation drawn about the association between shell noun use and disciplinary variations (see chapter 4 and 5), as the current study has explicitly stated that the purpose of this study is not so much linguistic description of shell nouns, but the disciplinary trends reflected in the ways of knowledge-construction and knowledge-communication through certain linguistic features (shell noun in this case) in academic genres. This particular type of structural test was taken as a starting point for the fulfillment of that purpose. However the restriction of these structural tests raises serious concern for the shell noun lists developed in the current study (see chapter 4), because this restriction indicates the incomplete inclusion of shell noun instances. This requires a more comprehensive analysis of shell noun uses, including applying corpus-driven methodology, analyzing as many potential shell noun instances and patterns as possible. This would contribute essentially to the development of sufficient shell noun tool box, which can be used as a basis for the development teaching aids.

A third major limitation is that all the interpretations of the findings were reliant exclusively on the researcher's judgments. This limitation arises from the fact that this is a corpus-based study relying on existing resources and, due to this, the researcher did not have the opportunity to conduct interviews with student writers. Since it is the student writers who are informed about disciplinary norms and who practise them in their writing, the inclusion of interviews with the writers of the essays would have been ideal, as they could have provided more reliable reasons for their purposes and choices as well as underlying cognitive factors for the basis of quantitative data

interpretations in this study.

6.6. Directions for further research

This thesis aimed to investigate the different epistemological orientations across academic disciplinary domains and levels of study, as reflected in the ways of knowledge construction and communication through certain linguistic features in disciplinary student writing. This was approached from the perspective of the use of shell nouns, a type of ‘highly general units often treated as belonging to allegedly all-purpose lexical repositories that remain unaffected by disciplinary way of meaning’ (Benitez-Castro, 2021, p.146). Despite the aforementioned limitations, the study has provided substantially reliable evidence that the use of shell nouns reveals numerous disciplinary associations straddling the dimensions of lexis, syntactic patterns and semantic categories. This points to some important and yet uncharted or under-researched areas of enquiry.

Firstly, the present study has provided evidence that substantiates the typical broad soft-hard division of disciplinary discourse. A subsequent inquiry could be extended to finer-grained micro-disciplines and sub-disciplines that are perceived to share a number of similarities from the outside, such as the area of *condensed matter physics*, *optics* and *medical physics* (three disciplines in physics). This may demonstrate some interesting and significant differences.

Furthermore, this study can be described as text-centered or data-based, as the focus was placed on the study of linguistic features as encountered in the texts, which

is based on specific structural predefined patterns. As mentioned previously (see section 6.5), the current investigation of shell-noun use shows that there are some minor grammatical patterns that are associated with, but not accounted for in terms of, the lexicalizations of shell nouns. This indicates the necessity to investigate and consider more variables, such as formal structure, syntactic function, and the semantic structure of the shell noun (phrase) and so on, in order to provide a more detailed description of shell-noun phenomenon. Therefore, it is worthwhile for further research to consider a data-driven scrutiny of shell noun phenomenon in particular academic sub-genres and disciplines, where shell noun instances can be discovered thoroughly and investigated systemically. This direction of future research enables the researcher to provide more evidence to support, problematise, challenge or refine the existing theoretical framework, which relies on phrase and clause level patterns as the primary identifying features of shell nouns in discourse.

This thesis has presented a comparative and diachronic study focusing on identifying differences between and within disciplinary student writing, and due to this nature, the developmental characteristics features were not taken into consideration. Following the present study's discussion, it can be said that another important area of future corpus-based research might be to focus on the observation of developmental patterns of shell-noun uses in L2 students' disciplinary writing, in comparison with L1 student writings. Taking a longitudinal and cross-contextual perspective, it is believed that the corpus exploration of L1 and L2 student writing collected from different sources could contribute to a more comprehensive description

of distinctive developmental trajectories in different institutional contexts

Provided that the present study has demonstrated the usefulness of shell-noun uses in reflecting disciplinary ways of meaning, future studies are thus expected to explore the impact of teaching of shell nouns in EAP classroom on the performances of student academic writing, especially in disciplinary writings produced by non-native students who are en route to pursuing entrance to English-medium universities with various disciplinary academic backgrounds.

6.7. Final thoughts

In his book discussing the current tensions between the ‘three cultures’, that is, the culture of humanities, the culture of natural sciences and the culture of the social sciences, Kagan (2009, p.275) suggests, ‘it is time for the members of the three cultures to adopt a posture of greater humility for, like tigers, sharks, and hawks, each group is potent in its own territory but impotent in the territory of the other’. Furthermore, as acknowledged before, disciplinary intellectual boundaries are neither ‘entirely fixed nor fluid’, By contrast, they are ‘relational and informational’ (Barry and Born. 2013, p.20). Therefore, it is an obvious and important need for ESAP practitioners to highlight the fact that there is a greater mutuality of understanding among the members of different disciplinary domains.

Furthermore, interdisciplinary work, which ‘integrates knowledge and modes of thinking from two or more disciplines’ (Barry and Born, 2013, p.24), is becoming/has become as important and popular outside academia, as well as within. Such an

academic trend in the 21st century requires ESAP practitioners to actively explore the interdisciplinarity reflected by certain linguistic features, such as shell nouns, in order to help students who choose to engage in interdisciplinary work achieve academic success in their path towards expertise.

Perhaps more importantly, what has doing this thesis taught me about the discipline of corpus linguistics itself? Is it a soft or hard knowledge field? Although Teubert (2005, p.13) claims emphatically that the answer is the former, in saying that ‘corpus linguistics localizes the study of language, once again, firmly and deliberately, in the Geisteswissenschaften, the humanities’, I would like to propose that corpus linguistics is by no means merely restricted to the soft knowledge field, as evidenced by the frequent recourse that this thesis has made to the methodological and discursive apparatuses of the hard fields, such as statistical procedures, tables of quantitative data, computerized data, and so on. On the other hand, it is a soft field when it concerns the business of interpretation and contextualized explanations. Last but not the least, throughout the development of this thesis, I hope to have humbly contributed to the endeavour of understanding more about the relationship between the disciplinary shell-noun use and epistemological orientations. For all its many imperfections, it has gone some way towards achieving this goal.

I would like to end with a quote from the final poem *little Gidding* of T. S. Eliot's (1941) *Four Quartets* to mark the conclusion of the current investigation and the beginning of my future exploratory journey of corpus studies and the exploitation of corpus resources in language education:

*We shall not cease from exploration,
and the end of all our exploring,
will be to arrive where we started,
and know the place for the first time.*

References

- Adams, J., Khan, H.T., Raeside, R. and White, D.I. (2007) *Research methods for graduate business and social science students*. SAGE publications India.
- Ädel, A. (2010) 'Using corpora to teach academic writing: Challenges for the direct approach', *Corpus-based approaches to English language teaching*, pp.39-55.
- Aistleitner, M., Kapeller, J. and Steinerberger, S. (2019) 'Citation patterns in economics and beyond', *Science in context*, 32(4), pp.361-380.
- Alsop, S. and Nesi, H. (2009) 'Issues in the development of the British Academic Written English (BAWE) corpus', *Corpora*, 4(1), pp.71-83.
- Akimoto, M. (1990) 'Deverbal nouns in grammar and discourse', *Lacus Forum*, pp. 281-290.
- Anthony, L. (2017) *AntFileConverter (Version 1.2.1) Computer Software, Tokyo, Japan, Waseda University*. Available at: <http://www.laurenceanthony.net/software>. (Accessed: 20 March 2022)
- Anthony, L. (2018) *Introducing English for Specific Purposes*. Routledge.
- Anthony, L. and Hardaker, C. (2017) *FireAnt (Version 1.1.4) Computer Software, Tokyo, Japan, Waseda University*. Available at: <http://www.laurenceanthony.net/software>. (Accessed: 20 March 2022)
- Baert, P. (2005) *Philosophy of the social sciences: Towards pragmatism*. Polity.
- Barnbrook, G. (1996) *Language and Computers: A practical Introduction to the Computer Analysis of Language*. Edinburgh; Edinburgh University Press.
- Barry, A., and Born, G. (2013) *Interdisciplinarity. Reconfigurations of the social and natural sciences*. London: Routledge.
- Basturkmen, H., East, M. and Bitchener, J. (2014) 'Supervisors' on-script feedback comments on drafts of dissertations: socialising students into the academic discourse community', *Teaching in Higher Education*, 19(4), pp.432-445.
- Bauer, H.H., (1992) *Scientific literacy and the myth of the scientific method*. University of Illinois Press.
- Bazerman, C. (1988) *Shaping written knowledge: Rhetoric of the human sciences*. The University of Wisconsin Press. The University of Wisconsin Press.
- Becher, T. (1987) 'Disciplinary discourse', *Studies in Higher Education*, 12(3), pp. 261-274.
- Becher, T. (1989) 'Academic Tribes and Territories: Intellectual Enquiry and the Cultures of Disciplines', *Milton Keynes: Society for Research into Higher Education*.
- Becher, T. and Trowler, P. (2001) *Academic tribes and territories*. McGraw-Hill Education (UK).
- Bednarek, M. (2006) 'Epistemological positioning and evidentiality in English news discourse: A text-driven approach', *Text & Talk*, 26 (6), pp. 635-660.
- Bednarek, M. (2008) *Emotion talk across corpora*. Springer.
- Benítez-Castro, M.Á. (2014) 'Formal, syntactic, semantic and textual features of

- English shell nouns', *PhD thesis*, Universidad de Granada.
- Benitez-Castro, M.A., and Thompson, P. (2015) 'Shell-nounhood in academic discourse: A critical state-of-the-art review', *International Journal of Corpus Linguistics*, 20(3), pp. 378-404.
- Benitez-Castro, M.A. (2021) 'Shell-noun use in disciplinary student writing: A multifaceted analysis of problem and way in third-year undergraduate writing across three disciplines', *English for Specific Purposes*, 61, pp.132-149.
- Berliner, D.C. (2002) 'Comment: Educational research: The hardest science of all', *Educational researcher*, 31(8), pp.18-20.
- Bhatia, V.K. (2014) 'A generic view of academic discourse', *Academic discourse*. pp. 31-49. Routledge.
- Biber, D., Conrad, S., and Reppen, R. (1998) *Corpus Linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E. (1999) *Longman Grammar of Spoken and Written English*. Pearson Education, Harlow.
- Biber, D. (2006) 'University language: A corpus-based study of spoken and written registers', *TESL-EJ*, 11(2).
- Biber, D., Gray, B. and Staples, S. (2016) 'Predicting patterns of grammatical complexity across language exam task types and proficiency levels', *Applied Linguistics*, 37(5), pp. 639-668.
- Biber, D. and Gray, B. (2016) *Grammatical complexity in academic English: Linguistic change in writing*. Cambridge: Cambridge University Press.
- Biglan, A. (1973) 'The characteristics of subject matter in different academic areas', *Journal of applied Psychology*, 57(3), p.195.
- Bloch, J. (2010) 'A concordance-based study of the use of reporting verbs as rhetorical devices in academic papers', *Journal of Writing Research*, 2(2), pp.219-244.
- Bloor, M. and Bloor, T. (1986) 'Languages for specific purposes: practice and theory', *Occasional Paper-Centre for Language and Communication Studies*, (19), pp.1-35.
- Bodin-Galvez, J. and Ding, A. (2019). 'Interdisciplinary EAP: Moving Beyond Aporitic English for General Academic Purposes', *The Language Scholar*, 4, pp.78-88.
- Botley, S.P. (2006) 'Indirect anaphora: Testing the limits of corpus-based linguistics', *International Journal of Corpus Linguistics*, 11(1), pp.73-112.
- Brezina, V., McEnery, T. and Wattam, S. (2015) 'Collocations in context: A new perspective on collocation networks', *International Journal of Corpus Linguistics*, 20(2), pp. 139-173.
- Bruce, I. (2005) 'Syllabus design for general EAP writing courses: A cognitive approach', *Journal of English for Academic Purposes*, 4(3), pp.239-256.
- Bruffee, K.A. (1986) 'Social construction, language, and the authority of knowledge: A bibliographical essay', *College English*, 48(8), pp.773-790.
- Bychkovska, T. (2021) 'Effects of explicit instruction on noun phrase production in L2 undergraduate writing', *Journal of English for Academic Purposes*, 54,

- p.101040.
- Caldwell, Candice. (2009) *Lexical vagueness in student writing: Are shell nouns the problem?*. Saarbrücken: VDM, Verlag Dr. Müller.
- Campion, G.C. (2016) 'The learning never ends: Exploring teachers' views on the transition from General English to EAP', *Journal of English for Academic Purposes*, 23, pp.59-70.
- Canagarajah, A.S. (2013) *Critical academic writing and multilingual students*. University of Michigan Press.
- Carter-Thomas, S. and Rowley-Jolivet, E. (2001) 'Syntactic differences in oral and written scientific discourse: the role of information structure', *ASp. la revue du GERAS*, (31-33), pp.19-37.
- Carter, R. (2004) *Language and creativity: The art of common talk*. London: Routledge.
- Celce-Murcia, M., Larsen-Freeman, D. (1999) *The Grammar Book: An ESL/EFL Teacher's Course, 2nd ed.* Heinle & Heinle, Boston.
- Charles, M. (2003) "'This mystery...': a corpus-based study of the use of nouns to construct stance in theses from two contrasting disciplines", *Journal of English for Academic Purposes*, 2(4), pp.313-326.
- Charles, M. (2007) 'Argument or evidence? Disciplinary variation in the use of the Noun that pattern in stance construction', *English for Specific Purposes*, 26(2), pp.203-218.
- Charles, M., Hunston, S. and Pecorari, D. eds. (2009) *Academic writing: At the interface of corpus and discourse*. Bloomsbury Publishing.
- Charles, M. (2012) 'Proper vocabulary and juicy collocations: EAP students evaluate do-it-yourself corpus-building', *English for Specific Purposes*, 31(2), pp.93-102.
- Charles, M. (2014) 'Getting the corpus habit: EAP students' long-term use of personal corpora', *English for Specific Purposes*, 35, pp.30-40.
- Chen, M. and Flowerdew, J., (2018) 'A critical review of research and practice in data-driven learning (DDL) in the academic writing classroom', *International Journal of Corpus Linguistics*, 23(3), pp.335-369.
- Cheng, W. and Ching, T. (2018) "'Not a guarantee of future performance': The local grammar of disclaimers", *Applied Linguistics*, 39(3), pp.263-301.
- Cochran, W.G. (1952) 'The χ^2 test of goodness of fit', *The Annals of mathematical statistics*, pp.315-345.
- Cohen, L., Manion, L. and Morrison, K. (2002) *Research methods in education*. routledge.
- Conrad, S. (2002) 'Corpus linguistic approaches for discourse analysis', *Annual review of applied linguistics*, 22, p.75.
- Cressie, N. and Read, T.R. (1989) 'Pearson's X^2 and the loglikelihood ratio statistic G^2 : a comparative review', *International Statistical Review/Revue Internationale de Statistique*, pp.19-43.
- Crosthwaite, P. ed. (2019) *Data-driven learning for the next generation: Corpora and DDL for pre-tertiary learners*. Routledge.
- Crosthwaite, P. and Cheung, L. (2019) *Learning the language of dentistry:*

- Disciplinary corpora in the teaching of English for specific academic purposes.* John Benjamins Publishing Company.
- Cullip, P.F. (2000) 'Text technology: The power-tool of grammatical metaphor', *RELC Journal*, 31(2), pp.76-104.
- Dahl, T. (2004) 'Textual metadiscourse in research articles: a marker of national culture or of academic discipline?', *Journal of Pragmatics*, 36(10), pp.1807-1825.
- Ding, A. and Evans, M. eds. (2022) *Social Theory for English for Academic Purposes: Foundations and Perspectives*. Bloomsbury Publishing.
- Donald, J.G. (1990) 'University professors' views of knowledge and validation processes', *Journal of Educational Psychology*, 82(2), p.242.
- Dong, J. and Lu, X. (2020) 'Promoting discipline-specific genre competence with corpus-based genre analysis activities', *English for Specific Purposes*, 58, pp. 138-154.
- Dong, M., Fang, A.C. and Qiu, X. (2020) 'Shell nouns as grammatical metaphor in knowledge construal: Variation across science and engineering discourse', *Lingua*, 248, p.102946.
- Dunning, Ted. (1993) 'Accurate methods for the statistics of surprise and coincidence', *Computational Linguistics*, 19(1), pp.61-74.
- Durrant, P. (2014) 'Discipline and level specificity in university students' written vocabulary', *Applied Linguistics*, 35(3), pp.328-356.
- Durrant, P. (2017) 'Lexical bundles and disciplinary variation in university students' writing: Mapping the territories', *Applied Linguistics*, 38(2), pp.165-193.
- Eckstein, G., Rawlins, J., Taylor, H., Briggs, H., Candland, A., Hanks, E. and Hill, S. (2022) 'Reporting verb variation across disciplines', *Journal of Academic Language and Learning*, 16(1), pp.59-75.
- Eliot, T.S. (1941) *The Four Quartets: Burnt Norton, East Coker, The Dry Salvages, Little Gidding*. Faber and Faber, London.
- Fang, A.C. and Dong, M. (2021) 'Shell nouns as register-specific discourse devices', *International Journal of Corpus Linguistics*, 26(2), pp.219-247.
- Feak, C.B. (2008) 'Culture shock? Genre shock', *British association of lecturers in English for academic purposes*. University of Reading, Whiteknights, UK.
- Feak, C.B. (2019) 'EAP support for post-graduate students', *The Routledge handbook of English for Academic Purposes*, pp.513-525.
- Fish, S. (1980) *Is there a text in this class?: The authority of interpretive communities*. Harvard University Press.
- Flowerdew, J. (2000) 'Discourse community, legitimate peripheral participation, and the nonnative-English-speaking scholar', *TESOL quarterly*, 34(1), pp. 127-150.
- Flowerdew, J. (2002) 'A pedagogic grammar of signalling nouns in discourse', *Revista Canaria de Estudios Ingleses*, (44), pp.141-156.
- Flowerdew, J. (2003) 'Signalling nouns in discourse', *English for Specific Purposes*, 22(4), pp. 329-346.
- Flowerdew, L. (2004) 'The argument for using English specialized corpora to understand academic and professional language', *Discourse in the professions:*

- Perspectives from corpus linguistics*, 11, p.33.
- Flowerdew, J. (2006) 'Use of signalling nouns in a learner corpus', *International Journal of Corpus Linguistics*, 11(3), pp.345-362.
- Flowerdew, L. (2008) *Corpus-based analyses of the problem-solution pattern: A phraseological approach* (Vol. 29). John Benjamins Publishing.
- Flowerdew, J. and Forest, R.W. (2015) *Signalling nouns in academic English*. Cambridge University Press.
- Francis, G. (1986) *Anaphoric nouns* (No.11). English Language Research, Department of English, University of Birmingham.
- Francis, G. (1994) 'A corpus-driven approach to grammar—principles, methods and examples', in *M. Baker, G. Francis and E. Tognini-Bonelli (eds): Text and Technology: In Honour of John Sinclair*. John Benjamins, pp.137-56.
- Francis, G. (1994) 'Labelling discourse: An aspect of nominal-group lexical cohesion', *Advances in Written Text Analysis*. pp.83-101. London: Routledge.
- Francis, Gill, Susan Hunston., Elizabeth Manning. (1998) *Collins Cobuild grammar patterns 2: Nouns and adjectives*. London: HarperCollins Publishers.
- Fraurud, K., (1992) *Processing noun phrases in natural discourse*. (Doctoral dissertation, Stockholm University).
- Gardner, Sheena, Nesi, Hilary, and Douglas, Biber (2019) 'Discipline, level, genre: Integrating situational perspectives in a new MD analysis of university student writing', *Applied Linguistics*, 2018, pp.1-30.
- Gergen, K.J. and Thatchenkery, T.J. (1996) 'Organization Science as Social', *Journal of Applied Behavioral Science*, 32(4), pp. 356-377.
- Gilbert, S.F., (1995) 'Introduction postmodernism and science', *Science in Context*, 8(4), pp.559-561.
- Glynos, J. and Howarth, D. (2007) *Logics of critical explanation in social and political theory*. Routledge.
- Granger, S. (1998) 'The computer learner corpus: A versatile new source of data for SLA research', *Learner English on Computer*, pp.3-18.
- Gray, B. (2021) 'On the Complexity of Academic Writing: Disciplinary Variation and Structural Complexity', *The Register-Functional Approach to Grammatical Complexity*, pp.93-112.
- Gray, B. (2015a) *Linguistic variation in research articles: when discipline tells only part of the story* (Vol. 10). John Benjamins Publishing Company.
- Gray, B. (2015b) 'On the complexity of academic writing: Disciplinary variation and structural complexity', *Corpus-based Research in Applied Linguistics*, pp.49-78.
- Gray, B. (2013) 'More than discipline: Uncovering multi-dimensional patterns of variation in academic research articles', *Corpora*, 8(2), pp.153-181.
- Groom, N. (2000) 'Attribution and averral revisited: Three perspectives on manifest intertextuality in academic writing', *Patterns and perspectives: Insights into EAP writing practice*, pp.14-25.
- Groom, N.W. (2007) *Phraseology and epistemology in humanities writing: a corpus-driven study*. (Doctoral dissertation, University of Birmingham).
- Halliday, M.A. and Hasan, R. (1976) *Cohesion in English*. London: Longman, 41.

- Halliday, M.A.K., Matthiessen, C.M.I.M. (1999) *Construing Experience Through Meaning: A Language-Based Approach to Cognition*. Cassell, London/New York.
- Halliday, M.A.K., Matthiessen, C.M. (2014) *An introduction to functional grammar*. Routledge.
- Hall, D., and Birkerts, S., (2007) *Writing Well, 9th ed.* Longman, New York.
- Hartley, J., Sotto, E. and Fox, C. (2004) 'Clarity across the disciplines: An analysis of texts in the sciences, social sciences, and arts and humanities', *Science Communication*, 26(2), pp.188-210.
- Harwood, N. (2017) 'The EAP practitioner as researcher and disseminator of knowledge', *Teaching EAP*.
- Harwood, N., Petrić, B. (2012) 'Performance in the Citing Behavior of Two Student Writers', *Written communication*, Vol.29 (1), p.55-103.
- Heuboeck, A., Holmes, J. and Nesi, H. (2010) *The BAWE Corpus Manual*. for the project entitled An investigation of genres of assessed writing in British Higher Education.
- Hiltunen, T. (2010) "There are good reasons for this': Disciplinary variation in the use of existential there constructions in academic research articles', *Constructing Interpersonality: Multiple Perspectives on Written Academic Genres*, pp. 181-204.
- Hinkel, E. (2001) 'Matters of cohesion in L2 academic texts', *Applied language learning*, 12(2), pp.111-132.
- Hinkel, E. (2003) *Teaching academic ESL writing: Practical techniques in vocabulary and grammar*. Routledge.
- Hoey, M. and Winter, E.O. (1983) *On the surface of discourse*. London: Allen & Unwin.
- Hoey, M. (1997) 'Some text properties of certain nouns', *Plenary paper presented at the Colloquium on Discourse Anaphora and Reference Resolution*, Lancaster, England.
- Hoye, L., (2014) *Adverbs and modality in English*. Routledge.
- Hu, G. and Perez, MR (2022) 'Effects of Explicit Instruction in Nominalisation on ESL Learners' Academic Writing', *Revista Iberica*, 43, pp.27-54.
- Hu, G. and Wang, G. (2014) 'Disciplinary and ethnolinguistic influences on citation in research articles', *Journal of English for Academic Purposes*, 14, pp.14-28.
- Huckin, T.N. (2003) 'Specificity in LSP', *Ibérica, Revista de la Asociación Europea de Lenguas para Fines Específicos*, (5), pp.3-18.
- Hu, G. and Perez, M.R. (2022). 'Effects of explicit instruction in nominalisation on ESL learners' academic writing'. *Ibérica: Revista de la Asociación Europea de Lenguas para Fines Específicos (AELFE)*, (43), pp.27-54.
- Huckin, T.N., Pesante, L.H. (1988) 'Existential there', *Written Communication*. 5(3), pp. 368-391.
- Hudson, R. (1984) *Word Grammar*. Oxford: Blackwell.
- Hunston, S. and G. Francis. (2000). *Pattern Grammar: A Corpus-driven Approach to the Lexical Grammar of English*. John Benjamins.

- Hunston, S. (2000) 'Evaluation and planes of discourse: status and value in persuasive texts', *Pattern Grammar: A Corpus-Driven Approach to Lexical Grammar of English*, pp.176–207.
- Hunston, S., and Sinclair, J. (2000) 'A local grammar of evaluation', *Evaluation in text*. Oxford: Oxford University Press, pp. 74-101.
- Hunston, S. and Thompson, G. eds. (2000) *Evaluation in text: Authorial stance and the construction of discourse: Authorial stance and the construction of discourse*. Oxford University Press, UK.
- Hunston, S. and Francis, G. (2000) *Pattern grammar: A corpus-driven approach to the lexical grammar of English* (No. 4). John Benjamins Publishing.
- Hunston, S. (2002a) *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Hunston, S. (2002b) *Pattern grammar, language teaching, and linguistic variation*. Using corpora to explore linguistic variation, pp.167-183.
- Hunston, S. (2004) "It has rightly been pointed out...". Attribution, consensus and conflict in academic English', *Academic Discourse, Genre and Small Corpora*, pp.15-33.
- Hunston, S. (2011) *Corpus Approaches to Evaluation: Phraseology and Evaluative Language*. Routledge.
- Hunston, S. (2013) 'Systemic functional linguistics, corpus linguistics, and the ideology of science', *Text & Talk*, 33(4-5), pp.617-640.
- Hunston, S. (2015) 'Lexical grammar', *Handbook of Corpus Linguistics*, pp.201-215.
- Hunston, S. and Su, H. (2019) 'Patterns, constructions, and local grammar: A case study of 'evaluation'', *Applied Linguistics*, 40(4), pp.567-593.
- Hyland, K. (1998) 'Persuasion and context: The pragmatics of academic metadiscourse', *Journal of Pragmatics*, 30(4), pp.437-455.
- Hyland, K. (1999) 'Academic attribution: Citation and the construction of disciplinary knowledge', *Applied Linguistics*, 20(3), pp.341-367.
- Hyland, K. (2000) *Disciplinary discourses, Michigan classics ed.: Social interactions in academic writing*. University of Michigan Press.
- Hyland, K. and Hamp-Lyons, L. (2002) 'EAP: Issues and directions', *Journal of English for Academic Purposes*, 1(1), pp.1-12.
- Hyland, K. (2004a) *Disciplinary discourses, Michigan classics ed.: Social interactions in academic writing*. University of Michigan Press.
- Hyland, K. (2004b) 'Disciplinary interactions: Metadiscourse in L2 postgraduate writing', *Journal of Second Language Writing*, 13(2), pp.133-151.
- Hyland, K. and P. Tse. (2005a), 'Evaluative that constructions: Signalling stance in research abstracts,' *Functions of Language*, 12(1), pp.39-63.
- Hyland, K. and P. Tse. (2005b), 'Hooking the reader: A corpus study of evaluative that in abstracts,' *English for Specific Purposes*, 24(2), pp.123-39.
- Hyland, K. (2006) *English for Academic Purposes: An Advanced Resource Book*. Routledge.
- Hyland, K. and Bondi, M. eds. (2006) *Academic Discourse Across Disciplines* (Vol. 42). Peter Lang.

- Hyland, K. and Tse, P. (2007) 'Is there an "academic vocabulary"?', *TESOL quarterly*, 41(2), pp.235-253.
- Hyland, K. (2007) 'Applying a gloss: Exemplifying and reformulating in academic discourse', *Applied Linguistics*, 28(2), pp.266-285.
- Hyland, K. (2008) 'As can be seen: Lexical bundles and disciplinary variation', *English for Specific Purposes*, 27(1), pp.4-21.
- Hyland, K. (2011) 'Disciplines and discourses: Social interactions in the construction of knowledge', *Writing in knowledge societies*, pp.193-213.
- Hyland, K. (2012) *Disciplinary identities: Individuality and community in academic discourse*. Cambridge University Press.
- Hyland, K. (2015) 'Corpora and written academic English', *The Cambridge Handbook of English Corpus Linguistics (Cambridge Handbooks in Language and Linguistics)*, pp.292-308.
- Hyland, K. and Shaw, P. eds. (2016) *The Routledge handbook of English for academic purposes*. London: Routledge.
- Hyland, K., (2018) 'Sympathy for the devil? A defence of EAP', *Language Teaching*, 51(3), pp.383-399.
- Hyland, K. and Jiang, F.K. (2018a) 'In this paper we suggest: Changing patterns of disciplinary metadiscourse', *English for Specific Purposes*, 51, pp.18-30.
- Hyland, K. and Jiang, F.K., (2018b) 'Changing patterns of self-citation: Cumulative inquiry or self-promotion?', *Text & Talk*, 38(3), pp.365-387.
- Hyland, K. and Jiang, F.K. (2019a) *Academic discourse and global publishing: Disciplinary persuasion in changing times*. Routledge.
- Hyland, K. and Jiang, F.K. (2019b) 'Points of reference: Changing patterns of academic citation', *Applied Linguistics*, 40(1), pp.64-85.
- Hymes, D. (1972) 'On communicative competence', *Sociolinguistics*, 269293(1972), pp.269-293.
- Ivanič, R. (1991) 'Nouns in search of a context: A study of nouns with both open-and closed-system characteristics', *International Review of Applied Linguistics in Language Teaching*, 29(1), pp.93-114.
- Johns, A. M. (1997) *Text, role and context: developing academic literacies*. Cambridge: Cambridge University Press.
- Jiang, F.K. and Hyland, K. (2015) 'The fact that: Stance nouns in disciplinary writing', *Discourse Studies*, 17(5), pp.529-550.
- Jiang, F.K. and Hyland, K. (2017) 'Metadiscursive nouns: Interaction and cohesion in abstract moves', *English for Specific Purposes*, 46(1), pp.1-14.
- Jiang, L. (2017) 'The affordances of digital multimodal composing for EFL learning', *Elt Journal*, 71(4), pp.413-422.
- Jiang, F.K. and Hyland, K. (2018) 'Nouns and academic interactions: A neglected feature of metadiscourse', *Applied Linguistics*, 39(4), pp.508-531.
- Jiang, F.K. and Hyland, K. (2020) 'There are significant differences...: the secret life of existential there in academic writing', *Lingua*, 233(2), p.102758.
- Jwa, S. (2022) 'L2 College Students' Disciplinary Literacy Development: The Mediating Role of Writing Practice', *Journal of Language, Identity & Education*,

- pp.1-17.
- Jiang, F.K. and Hyland, K. (2022) ‘the datasets do not agree’: Negation in research abstracts’. *English for Specific Purposes*, 68(2022), pp. 60-72.
- Kagan, J. (2009) *The three cultures: Natural sciences, social sciences and the humanities in the 21st century*. New York: Cambridge University Press.
- Kilgarriff, A., Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý, Vít Suchomel. (2014) ‘The Sketch Engine: ten years on’, *Lexicography*, 1, pp.7-36.
- Kim, C. and Crosthwaite, P. (2019) ‘Disciplinary differences in the use of evaluative that: Expression of stance via that-clauses in business and medicine’, *Journal of English for Academic Purposes*, 41, p.100775.
- Kim, Y. and Belcher, D. (2020) ‘Multimodal composing and traditional essays: Linguistic performance and learner perceptions’. *RELC Journal*, 51(1), pp. 86-100.
- Koester., A . (2010) ‘Building small specialised corpora’, *The Routledge handbook of corpus linguistics*, Routledge, pp.66-79.
- Kohnke, L., Jarvis, A. and Ting, A. (2021) ‘Digital multimodal composing as authentic assessment in discipline-specific English courses: Insights from ESP learners’, *TESOL Journal*, 12(3), p.e600.
- Kolb, D.A. (1981) ‘Learning styles and disciplinary differences’, *The modern American college*, 1(January1981), pp.232-235.
- Kuhn, T.S. (2011) *The Essential Tension: Selected Studies in Scientific Tradition and Change*. University of Chicago press.
- Lambrecht, K. (1996) *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge Cambridge University Press.
- Lakatos, I. (1968) *The problem of inductive knowledge*. Amsterdam, North Holland.
- Lakatos, I. (1976a) ‘Falsification and the methodology of scientific research programmes’, *Can theories be refuted?*, pp.205-259.
- Lakatos, I. (1976b) *Proofs and refutations*. Cambridge, Cambridge University Press.
- Lakatos, I. (1978) ‘Science and pseudoscience’, *Philosophical papers*, 1, pp.1-7.
- Latour, B., Woolgar, S. (1979) *Laboratory Life: the Social Construction of Scientific Facts*. Sage Publications, Beverly Hills; London
- Lea, M.R. and Street, B. (2014) ‘Understanding textual practices in higher education’, *Writing: Texts, Process and Practices*, pp.62-81.
- Liu, Q. and Deng, L. (2017) ‘A genre-based study of shell-noun use in the N-be-that construction in popular and professional science articles’, *English for Specific Purposes*, 48(1), pp.32-43.
- Lyons, J. (1977) *Semantics: Volume 2* (Vol. 2). Cambridge university press.
- Mahlberg, M. (2005) *English General Nouns*. Amsterdam/Philadelphia: John Benjamins.
- Manathunga C., and Brew, A. (2012) ‘Beyond tribes and territories: New metaphors for new times’, *Tribes and territories in the 21st-century: Rethinking the significance of disciplines in higher education*, London: Routledge, pp.44-56.

- Mauranen, A. and Bondi, M. (2003) 'Evaluative language use in academic discourse', *Journal of English for Academic Purposes*, 2(4), pp.269-271.
- Muguiro, N.F. (2019) *Interdisciplinarity and academic writing: a corpus-based case study of three interdisciplines* (Doctoral dissertation, University of Birmingham).
- Martin, J.R. (1992) Genre and literacy-modeling context in educational linguistics. *Annual review of applied linguistics*, 13, pp.141-172.
- Martin, J. (2011) 'Bridging troubled waters: Interdisciplinarity and what makes it stick', *Disciplinarity: Functional linguistic and sociological perspectives*, pp.35-61.
- Mbodj, N.B. (2021) 'Writing in the Disciplines and Within-discipline Variations: A Comparison of the Formulaic Profiles of the Medical Research Article and the Medical Case Report', *PhD thesis*, Georgia State University.
- McEnery, T. and Kifle, N.A. (2002) 'Epistemic modality in argumentative essays of second-language writers', *Academic Discourse*, pp.182-195.
- McEnery, T. and Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.
- Michigan Corpus of Upper-level Student Papers. (2009). Ann Arbor, MI: The Regents of the University of Michigan.
- Moore, T. (2002) 'Knowledge and agency: a study of 'metaphenomenal discourse' in textbooks from three disciplines', *English for Specific Purposes*, 21(4), pp. 347-366.
- Myers, G. (1991) 'Lexical cohesion and specialized knowledge in science and popular science texts', *Discourse processes*, 14(1), pp.1-26.
- Narrog, H. (2012) *Modality, subjectivity, and semantic change: a cross-linguistic perspective*. Oxford University Press.
- Nesi, H., Sharpling, G. and Ganobcsik-Williams, L. (2004) 'Student papers across the curriculum: Designing and developing a corpus of British student writing', *Computers and Composition*, 21(4), pp.439-450.
- Nesi, H. and Gardner, S. (2012) *Genres across the disciplines: Student writing in higher education*. Cambridge University Press.
- Nesi, H., (2014) 'Corpus query techniques for investigating citation in student assignments', *Corpus analysis for descriptive and pedagogical purposes*, pp.85-106.
- Nesi, H. (2021) 'Sources for courses: Metadiscourse and the role of citation in student writing', *Lingua*, 253, p.103040.
- Oliver, M. (2012) *Technology and change in academic practice*. Routledge.
- Omidian, T., Shahriari, H. and Siyanova-Chanturia, A. (2018) 'A cross-disciplinary investigation of multi-word expressions in the moves of research article abstracts', *Journal of English for Academic Purposes*, 36, pp.1-14.
- Omidian, T., Siyanova-Chanturia, A. and Biber, D. (2021) 'A new multidimensional model of writing for research publication: An analysis of disciplinarity, intra-textual variation, and L1 versus LX expert writing', *Journal of English for Academic Purposes*, 53, p.101020.
- Omidian, T. and Siyanova-Chanturia, A. (2021) 'Parameters of variation in the use of

- words in empirical research writing', *English for Specific Purposes*, 62, pp. 15-29.
- O'Keeffe, A., McCarthy, M., and Carter, R. (2007) *From corpus to classroom: Language use and language teaching*. Cambridge: Cambridge University Press.
- Öztekin, E. and Candan, E. (2019) 'The Role of Corpora in Language Teaching', *Research Trends in English Language Teacher Education and*, p.327.
- Palacios-Martínez, I. and Martínez-Insua, A. (2006) 'Connecting linguistic description and language teaching: native and learner use of existential there', *International Journal of Applied Linguistics*, 16(2), pp.213-231.
- Palmer, F. R. (1986) *Mood and Modality*. Cambridge University Press.
- Plappert, G. (2019) 'Not hedging but implying: Identifying epistemic implicature through a corpus-driven approach to scientific discourse', *Journal of Pragmatics*, 139, pp.163-174.
- Pearson, K. (1900) 'On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling', *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302), pp. 157-175.
- Podgórecki, A. (1997) *Higher Faculties: A Cross-National Study of University Culture*. Praeger Publishers, 88 Post Road West, Box 5007, Westport, CT 06881.
- Poole, R., Gnann, A. and Hahn-Powell, G. (2019) 'Epistemic stance and the construction of knowledge in science writing: A diachronic corpus study', *Journal of English for Academic Purposes*, 42, p.100784.
- Prior, P. (2013) *Writing/disciplinarity: A sociohistoric account of literate activity in the academy*. Routledge.
- Quirk, R. (2010) *A comprehensive grammar of the English language*. Pearson Education India.
- Rayson, P. and Garside, R. (2000) 'Comparing corpora using frequency profiling', *The workshop on comparing corpora*, pp.1-6.
- Rayson, P. (2003). *Matrix: A statistical method and software tool for linguistic analysis through corpus comparison*. Lancaster University (United Kingdom).
- Rayson, P. (2008). 'From key words to key semantic domains', *International journal of corpus linguistics*, 13(4), pp.519-549.
- Rayson, Paul. (2016). *Log-likelihood and effect size calculator*. Available at: <http://ucrel.lancs.ac.uk/llwizard.html> (Accessed: September 2021)
- Rayson, P. and Chapelle, C.A. (2019). 'Corpus analysis of key words', *The concise encyclopedia of applied linguistics*, pp. 320-326.
- Rezaei Keramati, S., Kuhl, D. and Saeidi, M., (2019) 'Cross-sectional diachronic corpus analysis of stance and engagement markers in three leading journals of applied linguistics', *Journal of Modern Research in English Language Studies*, 6(2), pp.25-1.
- Rezaei Keramati, S., Kuhl, D. and Saeidi, M., (2021) 'Diachronic corpus analysis of stance markers in research articles: The field of applied linguistics', *Cogent Arts & Humanities*, 8(1), p.1872165.

- Römer, U. and Swales, J.M. (2010) 'The Michigan corpus of upper-level student papers (MICUSP)', *Journal of English for Academic Purposes*, 9(3), p.249.
- Ruppert, E. (2013) 'Rethinking empirical social sciences', *Dialogues in Human Geography*, 3(3), pp. 268-273.
- Samraj, B. (2002) 'Introductions in research articles: Variations across disciplines', *English for Specific Purposes*, 21(1), pp.1-17.
- Samraj, B. (2016) 'Discourse structure and variation in manuscript reviews: Implications for genre categorization', *English for Specific Purposes*, 42, pp. 76-88.
- Sarte, K.M. and Gnevsheva, K., (2022) 'Noun phrasal complexity in ESL written essays under a constructed-response task: Examining proficiency and topic effects', *Assessing Writing*, 51, p.100595.
- Schmid, H.J. (1997) 'Constant and ephemeral hypostatization: thing, problem and other shell nouns', *Proceedings of the 16th International Congress of Linguists*, Paris.
- Schmid, H.J. (2000) *English abstract nouns as conceptual shells*. Berlin; New York: Mouton de Gruyter.
- Schmid, H.J. (2007) 'Non-compositionality and emergent meaning of lexico-grammatical chunks: A corpus study of noun phrases with sentential complements as constructions', *Zeitschrift für Anglistik und Amerikanistik*, 55(3), pp.313-340.
- Scott, M. (2004) *Wordsmith Tools version 4*. Oxford: Oxford University Press.
- Simon-Vandenberg, A.M., Taverniers, M. and Ravelli, L.J. eds. (2003) *Grammatical metaphor: Views from systemic functional linguistics* (Vol. 236), John Benjamins Publishing.
- Sinclair, John M., Gwyneth Fox, Stephen Bullon, Ramesh Krishnamurthy, Elizabeth Manning., John Todd. (1990) *Collins Cobuild English grammar*. London: HaperCollins Publishers.
- Sinclair, J. (1991) *Corpus Concordance Collocation*. Oxford: OUP.
- Sinclair, John M. (2004) *Trust the text: Language, corpus and discourse* (Edited with Ronald Carter). London: Routledge.
- Sing, C.S. (2013) 'Shell noun patterns in student writing in English for specific academic purposes', *Twenty years of learner corpus research: Looking back, moving ahead*, pp.411-422.
- Sosa, E. (1997) 'Reflective knowledge in the best circles', *The Journal of Philosophy*, 94(8), pp.410-430.
- Staples, S., Egbert, J., Biber, D. and Gray, B. (2016) 'Academic writing development at the university level: Phrasal and clausal complexity across level of study, discipline, and genre', *Written Communication*, 33(2), pp.149-183.
- Starfield, S. (2001) "'I'll go with the group': Rethinking 'discourse community' in EAP", *Research perspectives on English for academic purposes*, pp.132-147.
- Steinberg, D.D.C. and Osgood, C.E. (1971) *Semantics: An interdisciplinary reader in philosophy, linguistics and psychology*. CUP Archive.
- Su, H. (2017) 'Local grammars of speech acts: An exploratory study', *Journal of*

- Pragmatics* 111, pp.72-83.
- Su, H. and Wei, N. (2018) 'I'm really sorry about what I said: A local grammar of apology', *Pragmatics*, 28(3), pp.439-462.
- Su, H., Zhang, Y. and Lu, X. (2021) 'Applying local grammars to the diachronic investigation of discourse acts in academic writing: The case of exemplification in Linguistics research articles', *English for Specific Purposes*, 63(2), pp.120-133.
- Swales, J. M. (2014a) 'The concept of discourse community', *Wardle and Downs*, pp. 215-28.
- Swales, J. M. (2014b) 'Variation in citational practice in a corpus of student biology papers: From parenthetical plonking to intertextual storytelling', *Written Communication*, 31(2), pp.118-141.
- Tadros, A. (1994) 'Predictive categories in expository text', *Advances in Written Text Analysis*, pp. 66-82. New York: Routledge.
- Teubert, W. (2005) 'My version of corpus linguistics', *International Journal of Corpus Linguistics*, 10(1), pp. 1-13.
- Thompson, G. (2014) *Introducing Functional Grammar, 3rd ed.* Routledge, London.
- Thompson, G. and Yiyun, Y. (1991) 'Evaluation in the reporting verbs used in academic papers', *Applied linguistics*, 12(4), pp.365-382.
- Thompson, P. (2001) 'A pedagogically-motivated corpus-based examination of PhD theses: Macrostructure, citation practices and uses of modal verbs', *PhD thesis*, Reading, UK: University of Reading.
- Thompson, P. and Tribble, C. (2001) 'Looking at citations: Using corpora in English for academic purposes', *Language Learning & Technology*, 5(3), pp.91-105.
- Thompson, P. (2005) 'Points of focus and position: Intertextual reference in PhD theses', *Journal of English for Academic Purposes*, 4(4), pp.307-323.
- Thompson, P. (2014) 'Exploring Hoey's notion of text colligation in a corpus of student writing', *Diachrony and Synchrony in English Corpus Linguistics*, Peter Lang, pp.347-371.
- Thompson, P. and Hunston, S. (2019) *Interdisciplinary research discourse: Corpus investigations into environment journals*. Routledge.
- Thompson, P. (2022) 'Corpus Analysis of Disciplinary Variation and the Teaching of Esp/Eap', *The Routledge Handbook of Corpora and English Language Teaching and Learning*, pp.177-192.
- Thoreau, M. (2013) *Write on track: a guide to academic writing*. Pearson Education New Zealand.
- Tognini-Bonelli, Elena. (2001) *Corpus linguistics at work*. Amsterdam, Philadelphia, PA: John Benjamins Publishing Company.
- Trowler, P., Saunders, M. and Bamber, V. eds. (2012) *Tribes and territories in the 21st-century* (pp. 5-29). London: Routledge.
- Trowler, P. (2013) 'Depicting and researching disciplines: strong and moderate essentialist approaches', *Studies in Higher Education*, 39(10), pp.1720-1731.
- Tuggy, D. (1996) 'The thing is is that people talk that way. The question is is why. Cognitive Linguistics in the Redwoods', *The Expansion of a New Paradigm in*

- Linguistics*, pp.713-752.
- Vandommele, G., Van den Branden, K., Van Gorp, K., and De Maeyer, S. (2017) 'In-school and out-of-school multimodal writing as an L2 writing resource for beginner learners of Dutch', *Journal of Second Language Writing*, 36, pp.23-36.
- Vendler, Z. (1968) *Adjectives and Nominalizations*. The Hague: Nouton.
- Viana, V. and O'Boyle, A. (2021) *Corpus linguistics for English for academic purposes*. Routledge.
- Warburton, N. (2020) *The basics of essay writing*. Routledge.
- Widdowson, H.G. (1983) *Learning purpose and language use*. Oxford University Press.
- Widdowson, H.G. (2000) 'On the limitations of linguistics applied', *Applied Linguistics*, 21(1), pp.3-25.
- Willett, T. (1988) 'A cross-linguistic survey of the grammaticization of evidentiality. Studies in Language', *International Journal sponsored by the Foundation "Foundations of Language"*, 12(1), pp. 51-97.
- Winter, E.O. (1977) 'A clause-relational approach to English texts: A study of some predictive lexical items in written discourse', *Instructional Science*, 6(1), pp.1-92.
- Winter, E.O. (1992) 'The notion of unspecific versus specific as one way of analysing the information of a fund-raising letter', *Discourse description: Diverse linguistic analyses of a fund-raising text*, pp.131-170.
- Woodward-Kron, R. (2004) "'Discourse communities' and 'writing apprenticeship': an investigation of these concepts in undergraduate Education students' writing", *Journal of English for Academic Purposes*, 3(2), pp.139-161.
- Woolgar, S. (1981) 'Discovery: logic and sequence in a scientific text', *The social process of scientific investigation*, pp. 239-268.
- Yi, Y., Shin, D.-S., and Cimasko, T. (2019) 'Multimodal literacies in teaching and learning English in and outside of school', In L. C. de Oliveira (Ed.), *The handbook of TESOL in K-12*, pp.163-177.
- Yoon, H. and Jo, J. (2014) 'Direct and indirect access to corpora: An exploratory case study comparing students' error correction and learning strategy use in L2 writing', *Language Learning & Technology*, 18(1), pp.96-117.
- Zhang, Y., Yu, S. and Yuan, K. (2020) 'Understanding Master's students' peer feedback practices from the academic discourse community perspective: A rethinking of postgraduate pedagogies', *Teaching in higher education*, 25(2), pp.126-140.
- Zou, H.J. and Hyland, K. (2022). 'How the medium shapes the message: Stance in two forms of book reviews'. *Journal of Pragmatics*, 193(2022), pp. 269-280.
- Zipf, G. (1936) *The Psychobiology of Language*. London: Routledge.
- Zipf, G. (1949) *Human Behavior and the Principle of Least Effort*. New York: Addison-Wesley.
- Zagzebski, L.T. (1996) *Virtues of the mind: An inquiry into the nature of virtue and the ethical foundations of knowledge*. Cambridge University Press.

Appendices

Appendix 1

Table 4.2: A word frequency list of cataphoric shell nouns found in AH (normalized per million words)

fact	406	drive	9	step	3	misconception	1	criterion	1
attempt	239	concept	9	recognition	3	metaphor	1	contention	1
ability	197	reluctance	8	question	3	message	1	conspiracy	1
idea	152	result	8	part	3	measure	1	competence	1
belief	99	realization	8	option	3	reality	1	assessment	1
need	80	potential	8	necessity	3	insight	1	allegation	1
desire	79	interpretation	8	motivation	3	inference	1	irony	1
notion	68	duty	8	impulse	3	incapacity	1	innovation	1
view	62	criticism	8	impetus	3	inaptitude	1	information	1
failure	56	role	7	illusion	3	impact	1	indicator	1
claim	52	response	7	hint	3	illustration	1	inclination	1
argument	50	quest	7	guarantee	3	ignorance	1	entitlement	1
inability	48	proposition	7	discovery	3	ideal	1	dogma	1
effort	43	observation	7	difficulty	3	hesitation	1	devotion	1
point	42	objection	7	demonstration	3	grace	1	denial	1
right	40	mistake	7	change	3	force	1	demand	1
conclusion	40	incentive	7	challenge	3	flaw	1	action	1
evidence	39	ground	7	responsibility	3	fight	1	appeal	1
decision	39	example	7	authority	3	fiction	1	campaign	1
assumption	35	thesis	7	assurance	3	faith	1	commitment	1
suggestion	32	wish	5	approach	3	success	1	consciousness	1
reason	32	knowledge	5	age	3	element	1	doctrine	1
possibility	28	effect	5	movement	3	discussion	1	proclamation	1
theory	23	conviction	5	policy	3	discourse	1	fallacy	1
problem	20	consequence	5	tenet	1	desperateness	1	focus	1
impression	20	condition	5	warning	1	defence	1	initiative	1
purpose	20	concern	5	version	1	danger	1	key	1
assertion	19	advantage	5	reference	1	crisis	1	line	1
power	17	account	5	use	1	consideration	1	meaning	1
capability	17	image	4	topic	1	consent	1	worry	1
proof	16	willingness	4	thinking	1	confidence	1	motive	1
opinion	16	urge	4	thing	1	concession	1	myth	1
hypothesis	16	unwillingness	4	theme	1	choice	1	weakness	1
determination	16	truth	4	suspicion	1	charge	1	probability	1
chance	16	temptation	4	supposition	1	characteristic	1		
tendency	13	task	4	subject	1	certainty	1		
will	12	stance	4	stereotype	1	benefit	1		
struggle	12	solution	4	scheme	1	beauty	1		
statement	12	sign	4	scepticism	1	audacity	1		
premise	12	requirement	4	rush	1	attribute	1		
objective	12	position	4	rumour	1	aspiration	1		
indication	12	perception	4	reply	1	anxiousness	1		
implication	12	mission	4	reminder	1	alternative	1		
case	12	insistence	4	readiness	1	admittance	1		
agreement	12	goal	4	reaction	1	acknowledgement	1		
way	11	feature	4	provision	1	accusation	1		
function	11	factor	4	prospect	1	reflection	1		
doubt	11	endeavour	4	propensity	1	punishment	1		

difference	11	definition	4	promise	1	priority	1
aim	11	convention	4	process	1	precondition	1
thought	9	consensus	4	principle	1	opposition	1
obligation	9	comment	4	principal	1	means	1
intention	9	basis	4	pretence	1	job	1
hope	9	awareness	4	powerlessness	1	interest	1
freedom	9	ambition	4	plan	1	instruction	1
feeling	9	acceptance	4	philosophy	1	frustration	1
fear	9	trouble	3	victory	1	event	1
explanation	9	time	3	understanding	1	enlightenment	1

Table 4.3: A word frequency list of cataphoric shell nouns found in SS (normalized per million words)

fact	364	role	13	necessity	5	understanding	3	answer	3
ability	364	result	13	justification	5	willing	3	analysis	3
idea	150	reality	13	indication	5	negation	3	advantage	3
attempt	132	premise	13	impression	5	message	3	acknowledgement	3
need	121	consensus	13	impetus	5	motivation	3	accusation	3
argument	93	concept	10	hypothesis	5	mission	3	acceptance	3
belief	85	thought	10	faith	5	misconception	3	advice	3
opportunity	76	suggestion	10	conception	5	line	3	affirmation	3
view	75	responsibility	10	effect	5	limitation	3	agreement	3
assumption	57	issue	10	drive	5	likelihood	3	alternative	3
right	57	ground	10	danger	5	judgement	3	capability	3
notion	57	freedom	10	context	3	irony	3	confidence	3
power	54	feature	10	example	3	interest	3		
failure	54	caveat	8	demand	3	interpretation	3		
duty	49	motive	8	provision	3	inference	3		
will	47	policy	8	scenario	3	incitement	3		
desire	47	pressure	8	stipulation	3	importance	3		
aim	44	requirement	8	doctrine	3	implication	3		
intention	44	unwillingness	8	wish	3	imperative	3		
point	36	struggle	8	wisdom	3	impact	3		
doubt	36	risk	8	perspective	3	habit	3		
inability	34	reluctance	8	prospection	3	inspiration	3		
effort	34	hope	8	urge	3	flaw	3		
potential	31	guarantee	8	truth	3	finding	3		
conclusion	31	goal	8	trust	3	fear	3		
willingness	28	function	8	trick	3	factor	3		
purpose	28	feeling	8	trend	3	explanation	3		
obligation	28	assertion	8	thing	3	illusion	3		
decision	28	criticism	8	tenet	3	dictum	3		
possibility	23	conviction	8	task	3	endeavour	3		
evidence	23	commitment	8	supposition	3	element	3		
problem	21	chance	8	myth	3	disadvantage	3		
incentive	21	choice	5	subject	3	difficulty	3		
concern	21	attitude	5	strategy	3	development	3		
basis	21	objective	5	stance	3	determination	3		
reason	18	contention	5	speculation	3	destiny	3		
promise	18	proposition	5	rule	3	definition	3		
presumption	18	exception	5	reminder	3	defence	3		
knowledge	18	weakness	5	remark	3	crux	3		
claim	18	thesis	5	movement	3	counterpoint	3		
difference	18	solution	5	question	3	contribution	3		
propensity	16	significance	5	quest	3	continuation	3		
theory	16	response	5	move	3	consequence	3		
probability	16	recognition	5	prospect	3	condition	3		
principle	16	reasoning	5	proposal	3	challenge	3		

opinion	16	realization	5	projection	3	campaign	3
capacity	16	proof	5	process	3	benefit	3
case	13	proclamation	5	position	3	awareness	3
way	13	perception	5	philosophy	3	aspiration	3
tendency	13	worry	5	permission	3	aspect	3
statement	13	object	5	option	3	approach	3

Table 4.4: A word frequency list of cataphoric shell nouns found in NS (normalized per million words)

ability	208	doubt	21	concern	10	likelihood	5	example	
fact	208	desire	21	consensus	10	key	5	job	5
need	115	claim	21	answer	10	justification	5	failure	5
attempt	104	chance	21	necessity	10	issue	5	initiative	5
tendency	73	incentive	21	suggestion	10	intention	5	desperation	5
evidence	68	premise	16	thought	5	indicator	5	realisation	5
view	57	possibility	16	understanding	5	importance	5	willingness	5
problem	52	point	16	thing	5	ideal	5	report	5
aim	52	method	16	supposition	5	hypothesis	5	preference	5
advantage	47	finding	16	step	5	hope	5	motivation	5
assumption	42	feeling	16	standard	5	ground	5	pressure	5
theory	36	discovery	16	stance	5	opinion	5	encouragement	5
effort	36	decision	16	solution	5	function	5	freewill	5
capacity	36	criticism	16	similarity	5	field	5	drive	5
belief	36	awareness	10	scenario	5	feature	5	myth	5
notion	31	disability	10	reluctance	5	fault	5	development	5
inability	31	duty	10	result	5	fallacy	5	endeavour	5
way	26	responsibility	10	requirement	5	explanation	5	contribution	5
purpose	26	role	10	proof	5	difference	5	perspective	5
implication	26	response	10	process	5	demerit	5		
disadvantage	26	readiness	10	plan	5	defence	5		
argument	26	question	10	prediction	5	consequence	5		
right	21	principle	10	perception	5	condition	5		
reason	21	power	10	trust	5	choice	5		
propensity	21	limitation	10	option	5	benefit	5		
potential	21	guarantee	10	observation	5	beauty	5		
opportunity	21	expectation	10	object	5	attraction	5		
knowledge	21	effect	10	mistake	5	acceptance	5		
idea	21	conclusion	10	misconception	5	susceptibility	5		
freedom	21	challenge	10	measure	5	concept	5		

Appendix 2

Table 4.5: Distinct cataphoric shell nouns combined from the three corpora

ability	conclusion	explanation	job	prediction	solution
acceptance	condition	fact	judgment	preference	speculation
accusation	confidence	factor	justification	premise	stance
acknowledgement	consequence	failure	key	pressure	standard
advantage	contention	faith	knowledge	presumption	statement
advice	context	fallacy	likelihood	principle	step
affirmation	continuation	fault	limitation	probability	stipulation
agreement	contribution	fear	line	problem	strategy
aim	conviction	feature	measure	process	struggle
alternative	counterpoint	feeling	message	proclamation	subject
analysis	criticism	finding	method	projection	suggestion
answer	crux	flaw	misconception	promise	supposition
approach	danger	freedom	mission	proof	susceptibility
argument	decision	function	mistake	propensity	task
aspect	defence	goal	motivation	proposal	tendency
aspiration	definition	ground	motive	proposition	tenet
assertion	demand	guarantee	move	provision	theory
assumption	demerit	habit	movement	purpose	thesis
attempt	desire	hope	myth	quest	thing
attitude	desperation	hypothesis	necessity	question	thought
attraction	destiny	idea	need	readiness	trend
awareness	determination	ideal	negation	realisation	trick
basis	development	illusion	notion	reality	trust
beauty	dictum	impact	object	reason	truth
belief	difference	imperative	objective	reasoning	understanding
benefit	difficulty	impetus	obligation	recognition	unwillingness
campaign	disability	implication	observation	reluctance	urge
capability	disadvantage	importance	opinion	remark	view
capacity	discovery	impression	opportunity	reminder	way
case	doctrine	inability	option	report	weakness
caveat	doubt	incentive	perception	requirement	will
challenge	drive	incitement	permission	response	willing
chance	duty	indication	perceptive	responsibility	willingness
choice	effect	inference	philosophy	result	wisdom
claim	effort	initiative	plan	right	worry
clarity	element	inspiration	point	risk	
commitment	encouragement	intention	policy	role	
consensus	endeavour	interest	position	rule	
concept	evidence	interpretation	possibility	scenario	
conception	example	irony	potential	significance	
concern	exception	issue	power	similarity	

Appendix 3

Table 4.7: A set of common shell nouns in three corpora

No.	SNs	Combined			AH			NS			SS		
		Raw freq.	Ind. cover %	Cum. cover %	Raw freq.	Ind. cover %	Cum. cover %	Raw freq.	Ind. cover %	Cum. cover %	Raw freq.	Ind. cover %	Cum. cover %
1	fact	484	12.34	12.34	303	13.52	13.52	40	8.55	8.55	141	11.63	11.63
2	ability	328	8.37	20.71	147	6.56	20.08	40	8.55	17.09	141	11.63	23.27
3	attempt	249	6.35	27.06	178	7.94	28.02	20	4.27	21.37	51	4.21	27.48
4	idea	175	4.46	31.52	113	5.04	33.07	4	0.85	22.22	58	4.79	32.26
5	need	129	3.29	34.81	60	2.68	35.74	22	4.70	26.92	47	3.88	36.14
6	belief	114	2.91	37.72	74	3.30	39.05	7	1.50	28.42	33	2.72	38.86
7	view	86	2.19	39.91	46	2.05	41.10	11	2.35	30.77	29	2.39	41.25
8	desire	81	2.07	41.98	59	2.63	43.73	4	0.85	31.62	18	1.49	42.74
9	notion	79	2.01	43.99	51	2.28	46.01	6	1.28	32.91	22	1.82	44.55
10	argument	78	1.99	45.98	37	1.65	47.66	5	1.07	33.97	36	2.97	47.52
11	assumption	64	1.63	47.62	42	1.87	49.53	9	1.92	35.90	21	1.73	49.26
12	failure	64	1.63	49.25	26	1.16	50.69	8	1.71	37.61	30	2.48	51.73
13	right	56	1.43	50.68	30	1.34	52.03	4	0.85	38.46	22	1.82	53.55
14	inability	55	1.40	52.08	36	1.61	53.64	6	1.28	39.74	13	1.07	54.62
15	effort	52	1.33	53.40	32	1.43	55.06	7	1.50	41.24	13	1.07	55.69
16	opportunity	51	1.30	54.71	17	0.76	55.82	4	0.85	42.09	30	2.48	58.17
17	evidence	51	1.30	56.01	29	1.29	57.12	13	2.78	44.87	9	0.74	58.91
18	claim	50	1.28	57.28	39	1.74	58.86	4	0.85	45.73	7	0.58	59.49
19	point	48	1.22	58.51	31	1.38	60.24	3	0.64	46.37	14	1.16	60.64
20	conclusion	44	1.12	59.63	30	1.34	61.58	2	0.43	46.79	12	0.99	61.63
21	decision	43	1.10	60.72	29	1.29	62.87	3	0.64	47.44	11	0.91	62.54
22	power	36	0.92	61.64	13	0.58	63.45	2	0.43	47.86	21	1.73	64.27
23	reason	35	0.89	62.54	24	1.07	64.52	4	0.85	48.72	7	0.58	64.85
24	chance	35	0.89	63.43	8	0.36	64.88	10	2.14	50.85	17	1.40	66.25
25	aim	35	0.89	64.32	8	0.36	65.24	10	2.14	52.99	17	1.40	67.66
26	problem	33	0.84	65.16	15	0.67	65.91	10	2.14	55.13	8	0.66	68.32
27	possibility	33	0.84	66.00	21	0.94	66.85	3	0.64	55.77	9	0.74	69.06
28	suggestion	31	0.79	66.79	24	1.07	67.92	3	0.64	56.41	4	0.33	69.39
29	theory	30	0.77	67.56	17	0.76	68.67	7	1.50	57.91	6	0.50	69.88
30	purpose	30	0.77	68.32	14	0.62	69.30	5	1.07	58.97	11	0.91	70.79
31	tendency	29	0.74	69.06	10	0.45	69.75	14	2.99	61.97	5	0.41	71.20
32	duty	27	0.69	69.75	6	0.27	70.01	2	0.43	62.39	19	1.57	72.77
33	doubt	26	0.66	70.42	8	0.36	70.37	4	0.85	63.25	14	1.16	73.93
34	capacity	26	0.66	71.08	13	0.58	70.95	7	1.50	64.74	6	0.50	74.42
35	intention	25	0.64	71.72	7	0.31	71.26	1	0.21	64.96	17	1.40	75.83
36	potential	22	0.56	72.28	6	0.27	71.53	4	0.85	65.81	12	0.99	76.82
37	opinion	19	0.48	72.76	12	0.54	72.07	1	0.21	66.03	6	0.50	77.31
38	way	18	0.46	73.22	8	0.36	72.42	5	1.07	67.09	5	0.41	77.72
39	premise	17	0.43	73.65	9	0.40	72.82	3	0.64	67.74	5	0.41	78.14
40	incentive	17	0.43	74.09	5	0.22	73.05	4	0.85	68.59	8	0.66	78.80
41	extent	16	0.41	74.50	1	0.04	73.09	5	1.07	69.66	10	0.83	79.62
42	difference	16	0.41	74.90	8	0.36	73.45	1	0.21	69.87	7	0.58	80.20
43	proof	15	0.38	75.29	12	0.54	73.98	1	0.21	70.09	2	0.17	80.36
44	knowledge	15	0.38	75.67	4	0.18	74.16	4	0.85	70.94	7	0.58	80.94
45	implication	15	0.38	76.05	9	0.40	74.56	5	1.07	72.01	1	0.08	81.02
46	freedom	15	0.38	76.43	7	0.31	74.88	4	0.85	72.86	4	0.33	81.35
47	hypothesis	14	0.36	76.79	12	0.54	75.41	1	0.21	73.08	1	0.08	81.44
48	concern	14	0.36	77.15	4	0.18	75.59	2	0.43	73.50	8	0.66	82.10

49	advantage	14	0.36	77.51	4	0.18	75.77	9	1.92	75.43	1	0.08	82.18
50	feeling	13	0.33	77.84	7	0.31	76.08	3	0.64	76.07	3	0.25	82.43
51	role	12	0.31	78.14	5	0.22	76.31	2	0.43	76.50	5	0.41	82.84
52	criticism	12	0.31	78.45	6	0.27	76.57	3	0.64	77.14	3	0.25	83.09
53	propensity	11	0.28	78.73	1	0.04	76.62	4	0.85	77.99	6	0.50	83.58
54	principle	9	0.23	78.96	1	0.04	76.66	2	0.43	78.42	6	0.50	84.08
55	responsibility	6	0.15	79.11	2	0.09	76.75	2	0.43	78.85	2	0.17	84.24
56	awareness	6	0.15	79.27	3	0.13	76.89	2	0.43	79.27	1	0.08	84.32
57	acceptance	5	0.13	79.39	3	0.13	77.02	1	0.21	79.49	1	0.08	84.41
58	choice	4	0.10	79.50	1	0.04	77.06	1	0.21	79.70	2	0.17	84.57
59	defence	3	0.08	79.57	1	0.04	77.11	1	0.21	79.91	1	0.08	84.65
60	benefit	3	0.08	79.65	1	0.04	77.15	1	0.21	80.13	1	0.08	84.74
61	response	9	0.23	79.88	5	0.22	77.38	2	0.43	80.56	2	0.17	84.90
62	question	5	0.13	80.01	2	0.09	77.47	2	0.43	80.98	1	0.08	84.98
63	guarantee	7	0.18	80.18	2	0.09	77.55	2	0.43	81.41	3	0.25	85.23
64	expectation	6	0.15	80.34	1	0.04	77.60	2	0.43	81.84	3	0.25	85.48
65	effect	8	0.20	80.54	4	0.18	77.78	2	0.43	82.26	2	0.17	85.64
66	necessity	6	0.15	80.69	2	0.09	77.87	2	0.43	82.69	2	0.17	85.81
67	challenge	5	0.13	80.82	2	0.09	77.96	2	0.43	83.12	1	0.08	85.89
68	understanding	3	0.08	80.90	1	0.04	78.00	1	0.21	83.33	1	0.08	85.97
69	thing	3	0.08	80.97	1	0.04	78.05	1	0.21	83.55	1	0.08	86.06
70	supposition	3	0.08	81.05	1	0.04	78.09	1	0.21	83.76	1	0.08	86.14
71	stance	5	0.13	81.18	3	0.13	78.22	1	0.21	83.97	1	0.08	86.22
72	solution	6	0.15	81.33	3	0.13	78.36	1	0.21	84.19	2	0.17	86.39
73	result	12	0.31	81.64	6	0.27	78.63	1	0.21	84.40	5	0.41	86.80
74	requirement	7	0.18	81.82	3	0.13	78.76	1	0.21	84.62	3	0.25	87.05
75	process	3	0.08	81.89	1	0.04	78.80	1	0.21	84.83	1	0.08	87.13
76	perception	6	0.15	82.05	3	0.13	78.94	1	0.21	85.04	2	0.17	87.29
77	option	4	0.10	82.15	2	0.09	79.03	1	0.21	85.26	1	0.08	87.38
78	misconception	3	0.08	82.22	1	0.04	79.07	1	0.21	85.47	1	0.08	87.46
79	hope	11	0.28	82.50	7	0.31	79.38	1	0.21	85.68	3	0.25	87.71
80	ground	10	0.26	82.76	5	0.22	79.61	1	0.21	85.90	4	0.33	88.04
81	function	12	0.31	83.07	8	0.36	79.96	1	0.21	86.11	3	0.25	88.28
82	explanation	9	0.23	83.30	7	0.31	80.28	1	0.21	86.32	1	0.08	88.37
83	consequence	6	0.15	83.45	4	0.18	80.46	1	0.21	86.54	1	0.08	88.45
84	condition	6	0.15	83.60	4	0.18	80.63	1	0.21	86.75	1	0.08	88.53
85	realisation	9	0.23	83.83	6	0.27	80.90	1	0.21	86.97	2	0.17	88.70
86	willingness	15	0.38	84.21	3	0.13	81.04	1	0.21	87.18	11	0.91	89.60
87	motivation	4	0.10	84.32	2	0.09	81.12	1	0.21	87.39	1	0.08	89.69
88	drive	10	0.26	84.57	7	0.31	81.44	1	0.21	87.61	2	0.17	89.85
89	myth	3	0.08	84.65	1	0.04	81.48	1	0.21	87.82	1	0.08	89.93
90	endeavour	5	0.13	84.77	3	0.13	81.62	1	0.21	88.03	1	0.08	90.02
91	contribution	3	0.08	84.85	1	0.04	81.66	1	0.21	88.25	1	0.08	90.10
92	concept	12	0.31	85.16	7	0.31	81.97	1	0.21	88.46	4	0.33	90.43
93	example	7	0.18	85.34	5	0.22	82.20	1	0.21	88.68	1	0.08	90.51
94	thought	12	0.31	85.64	7	0.31	82.51	1	0.21	88.89	4	0.33	90.84

Appendix 4

Table 4.10: A list of anaphoric shell nouns in AH (normalized per million words)

way	40	tradition	4	conclusion	3	provision	1	decision	1
argument	24	role	4	choice	3	proposal	1	deception	1
idea	21	purpose	4	area	3	propensity	1	danger	1
point	20	policy	4	action	3	pronouncement	1	convention	1
case	19	period	4	willingness	1	programme	1	contrast	1
view	17	link	4	understanding	1	procedure	1	contradiction	1
statement	16	line	4	truth	1	possibility	1	context	1
theory	13	knowledge	4	time	1	philosophy	1	consideration	1
question	13	image	4	thought	1	opinion	1	connection	1
change	12	hypothesis	4	technique	1	need	1	concern	1
reason	11	difference	4	task	1	myth	1	conception	1
method	11	threat	3	subject	1	mistake	1	concept	1
issue	11	theme	3	struggle	1	judgment	1	challenge	1
interpretation	11	suggestion	3	strength	1	interest	1	belief	1
approach	11	practice	3	strategy	1	information	1	awareness	1
claim	9	phenomenon	3	solution	1	inability	1	attempt	1
trend	8	perception	3	significance	1	hope	1	assumption	1
problem	7	order	3	sentiment	1	failure	1	assessment	1
notion	7	influence	3	sadness	1	factor	1	appreciation	1
aspect	7	illusion	3	revolution	1	evidence	1	application	1
stage	5	ideal	3	requirement	1	enterprise	1	appeal	1
situation	5	freedom	3	relief	1	dilemma	1	anticipation	1
principle	5	feeling	3	reality	1	difficulty	1	analysis	1
part	5	experience	3	realization	1	determination	1	agreement	1
effect	5	example	3	readiness	1	description	1	advice	1
distinction	5	desire	3	reaction	1	denial	1	acknowledgement	1
attitude	5	definition	3	punishment	1	defence	1	acceptance	1

Table 4.11 :A list of anaphoric shell nouns in NS (normalized per million words)

case	125	issue	21	example	10	resolve	5	distinction	5
method	120	factor	21	criticism	10	proposal	5	difficulty	5
way	99	concept	21	context	10	presumption	5	difference	5
idea	52	role	16	concern	10	practise	5	description	5
view	47	result	16	assumption	10	perception	5	decision	5
definition	47	purpose	16	aim	10	part	5	counterargument	5
reason	42	phenomenon	16	action	10	observation	5	conundrum	5
theory	36	information	16	understanding	5	objection	5	contradiction	5
problem	36	hypothesis	16	threat	5	notion	5	condition	5
evidence	36	act	16	thought	5	mystery	5	compromise	5
situation	31	viewpoint	10	theme	5	link	5	comment	5
knowledge	31	topic	10	tendency	5	judgment	5	claim	5
approach	31	suggestion	10	task	5	illusion	5	belief	5
statement	26	reaction	10	subject	5	finding	5	awareness	5
effect	26	predicament	10	stage	5	feeling	5	argument	5
trend	21	position	10	rule	5	fact	5	activity	5
technique	21	point	10	risk	5	explanation	5	ability	5
question	21	improvement	10	response	5	element	5		

Table 4.12: A list of anaphoric shell nouns in SS (normalized per million words)

view	47	difference	10	procedure	5	right	3	habit	3
case	43	criticism	10	presumption	5	result	3	foundation	3
way	41	concept	10	potential	5	response	3	feature	3

idea	39	act	10	phenomenon	5	reluctance	3	fear	3
theory	36	trend	8	opportunity	5	recognition	3	event	3
argument	36	technique	8	opinion	5	reading	3	doctrine	3
approach	36	reason	8	knowledge	5	rationale	3	disappointment	3
point	34	policy	8	justification	5	quest	3	desire	3
question	26	information	8	ideal	5	principle	3	conviction	3
issue	26	goal	8	failure	5	pretence	3	contradiction	3
statement	23	feeling	8	factor	5	predicament	3	context	3
problem	21	evidence	8	decision	5	practice	3	consideration	3
notion	21	condition	8	claim	5	possibility	3	connection	3
fact	21	concern	8	attitude	5	position	3	conclusion	3
situation	18	conception	8	advantage	5	period	3	choice	3
perspective	18	basis	8	uncertainty	3	outcome	3	challenge	3
aspect	18	application	8	tradition	3	occasion	3	appeal	3
change	16	understanding	5	time	3	necessity	3	alternative	3
reasoning	13	topic	5	threat	3	myth	3	aim	3
method	13	theme	5	task	3	move	3	agreement	3
distinction	13	tendency	5	target	3	measure	3	action	3
definition	13	subject	5	strategy	3	logic	3	account	3
assumption	13	stereotype	5	step	3	link	3	ability	3
role	10	sentiment	5	stand	3	limitation	3		
image	10	rule	5	stance	3	initiative	3		
example	10	risk	5	solution	3	incentive	3		
effect	10	requirement	5	scheme	3	implication	3		

Appendix 5

Table 4.13: Distinct anaphoric shell nouns combined from the three corpora

ability	compromise	evidence	limitation	pretence	sentiment
acceptance	concept	example	line	principle	significance
account	conception	experience	link	problem	situation
acknowledgement	concern	explanation	logic	procedure	solution
act	conclusion	fact	measure	programme	stage
action	condition	factor	method	pronouncement	stance
activity	connection	failure	mistake	propensity	stand
advantage	consideration	fear	move	provision	statement
advice	context	feature	mystery	punishment	step
agreement	contradiction	feeling	myth	purpose	stereotype
aim	contrast	finding	necessity	quest	strategy
alternative	conundrum	foundation	need	question	strength
analysis	conviction	freedom	notion	rationale	struggle
anticipation	counterargument	goal	objection	reaction	subject
appeal	criticism	habit	observation	readiness	suggestion
application	danger	hope	occasion	reading	target
appreciation	deception	hypothesis	opinion	realization	task
approach	decision	idea	opportunity	reality	technique
area	defence	ideal	order	reason	tendency
argument	definition	illusion	outcome	reasoning	theme
aspect	denial	image	part	recognition	theory
assessment	description	implication	perception	relief	thought
assumption	desire	improvement	period	reluctance	threat
attempt	determination	inability	perspective	requirement	time
attitude	difference	incentive	phenomenon	resolve	topic
awareness	difficulty	influence	philosophy	response	tradition
basis	dilemma	information	point	result	trend
belief	disappointment	initiative	policy	revolution	truth
case	distinction	interest	position	right	uncertainty
challenge	doctrine	interpretation	possibility	risk	understanding
change	effect	issue	potential	role	view
choice	element	judgment	practice	rule	viewpoint
claim	enterprise	justification	predicament	sadness	way
comment	event	knowledge	presumption	scheme	willingness

Appendix 6

Table 4.15: A set of common anaphoric shell nouns in three corpora

No.	SNs	combined			AH			NS			SS		
		Ra w freq	Ind. cover %	Cum. cover %	Raw freq	Ind. cover %	Cum. cover %	Raw freq	Ind. cover %	Cum. cover %	Raw freq	Ind. cover %	Cum. cover %
1	way	65	5.89	5.89	30	7.43	7.43	19	6.91	6.91	16	3.87	3.87
2	case	55	4.99	10.88	14	3.47	10.89	24	8.73	15.64	17	4.12	7.99
3	idea	41	3.72	14.60	16	3.96	14.85	10	3.64	19.27	15	3.63	11.62
4	view	40	3.63	18.22	13	3.22	18.07	9	3.27	22.55	18	4.36	15.98
5	method	36	3.26	21.49	8	1.98	20.05	23	8.36	30.91	5	1.21	17.19
6	argument	33	2.99	24.48	18	4.46	24.50	1	0.36	31.27	14	3.39	20.58
7	theory	31	2.81	27.29	10	2.48	26.98	7	2.55	33.82	14	3.39	23.97
8	point	30	2.72	30.01	15	3.71	30.69	2	0.73	34.55	13	3.15	27.12
9	approach	28	2.54	32.55	8	1.98	32.67	6	2.18	36.73	14	3.39	30.51
10	statement	26	2.36	34.90	12	2.97	35.64	5	1.82	38.55	9	2.18	32.69
11	question	24	2.18	37.08	10	2.48	38.12	4	1.45	40.00	10	2.42	35.11
12	issue	22	1.99	39.08	8	1.98	40.10	4	1.45	41.45	10	2.42	37.53
13	problem	20	1.81	40.89	5	1.24	41.34	7	2.55	44.00	8	1.94	39.47
14	reason	19	1.72	42.61	8	1.98	43.32	8	2.91	46.91	3	0.73	40.19
15	situation	17	1.54	44.15	4	0.99	44.31	6	2.18	49.09	7	1.69	41.89
16	definition	16	1.45	45.60	2	0.50	44.80	9	3.27	52.36	5	1.21	43.10
17	notion	14	1.27	46.87	5	1.24	46.04	1	0.36	52.73	8	1.94	45.04
18	trend	13	1.18	48.05	6	1.49	47.52	4	1.45	54.18	3	0.73	45.76
19	effect	13	1.18	49.23	4	0.99	48.51	4	1.45	55.64	4	0.97	46.73
20	knowledge	11	1.00	50.23	3	0.74	49.26	6	2.18	57.82	2	0.48	47.22
21	evidence	11	1.00	51.22	1	0.25	49.50	7	2.55	60.36	3	0.73	47.94
22	role	10	0.91	52.13	3	0.74	50.25	3	1.09	61.45	4	0.97	48.91
23	distinction	10	0.91	53.04	4	0.99	51.24	1	0.36	61.82	5	1.21	50.12
24	claim	10	0.91	53.94	7	1.73	52.97	1	0.36	62.18	2	0.48	50.61
25	concept	9	0.82	54.76	1	0.25	53.22	4	1.45	63.64	4	0.97	51.57
26	technique	8	0.73	55.49	1	0.25	53.47	4	1.45	65.09	3	0.73	52.30
27	example	8	0.73	56.21	2	0.50	53.96	2	0.73	65.82	4	0.97	53.27
28	difference	8	0.73	56.94	3	0.74	54.70	1	0.36	66.18	4	0.97	54.24
29	assumption	8	0.73	57.66	1	0.25	54.95	2	0.73	66.91	5	1.21	55.45
30	phenomenon	7	0.63	58.30	2	0.50	55.45	3	1.09	68.00	2	0.48	55.93
31	information	7	0.63	58.93	1	0.25	55.69	3	1.09	69.09	3	0.73	56.66
32	factor	7	0.63	59.56	1	0.25	55.94	4	1.45	70.55	2	0.48	57.14
33	feeling	6	0.54	60.11	2	0.50	56.44	1	0.36	70.91	3	0.73	57.87
34	concern	6	0.54	60.65	1	0.25	56.68	2	0.73	71.64	3	0.73	58.60
35	theme	5	0.45	61.11	2	0.50	57.18	1	0.36	72.00	2	0.48	59.08
36	link	5	0.45	61.56	3	0.74	57.92	1	0.36	72.36	1	0.24	59.32
37	action	5	0.45	62.01	2	0.50	58.42	2	0.73	73.09	1	0.24	59.56
38	understanding	4	0.36	62.38	1	0.25	58.66	1	0.36	73.45	2	0.48	60.05
39	threat	4	0.36	62.74	2	0.50	59.16	1	0.36	73.82	1	0.24	60.29
40	subject	4	0.36	63.10	1	0.25	59.41	1	0.36	74.18	2	0.48	60.77
41	practise	4	0.36	63.46	2	0.50	59.90	1	0.36	74.55	1	0.24	61.02
42	decision	4	0.36	63.83	1	0.25	60.15	1	0.36	74.91	2	0.48	61.50
43	context	4	0.36	64.19	1	0.25	60.40	2	0.73	75.64	1	0.24	61.74
44	task	3	0.27	64.46	1	0.25	60.64	1	0.36	76.00	1	0.24	61.99
45	contradiction	3	0.27	64.73	1	0.25	60.89	1	0.36	76.36	1	0.24	62.23

Appendix 7

Table 4.22: Types and sub-types of transitivity process construed by distinct shell nouns across three disciplinary domains (rounded figures normalized per million words)

Types of transitivity process	Sub-types of transitivity process	SN types and frequencies in AH	SN types and frequencies in NS	SN types and frequencies in SS
attributive relational process	Comparative	<i>difference(11),</i>	<i>similarity(5), difference(5),</i>	<i>difference(18),</i>
	Attitudinal	<i>problem(20), advantage(5), difficulty(3), weakness(1), inaptitude(1), flaw(1), danger(1), benefit(1), beauty(1), powerlessness(1)</i>	<i>problem(52), advantage(47), disadvantage(26), limitation(10), issue(5), demerit(5), danger(5), benefit(5), beauty(5), attraction(5)</i>	<i>problem(21), issue(10), weakness(5), danger(5), limitation(3), flaw(3), disadvantage(3), difficulty(3), benefit(3), advantage(3), crux(3)</i>
	Epistemic	<i>possibility(28), chance(16), certainty(1), probability(1)</i>	<i>chance(21), possibility(16), likelihood(5)</i>	<i>possibility(23), probability(16), chance(8), likelihood(3)</i>
	Deontic	<i>need(80), right(40), power(17), obligation(9), freedom(9), duty(8), role(7), task(4), requirement(4), mission(4), convention(4), necessity(4), challenge(3), pressure(3), authority(3), responsibility(3), precondition(1), job(1), force(1), custom(1)</i>	<i>need(115), freedom(21), right(21), responsibility(10), necessity(10), challenge(10), power(10), pressure(5), role(5), requirement(5), job(5), freewill(5)</i>	<i>right(57), duty(29), obligation(28), role(13), responsibility(10), requirement(10), freedom(10), necessity(5), task(3), mission(3), imperative(3), destiny(3), challenge(3), pressure(3)</i>
Identifying relational process	Dynamic	<i>ability(197), inability(48), opportunity(23), capability(17), tendency(13), way(11), potential(8), failure(8), temptation(4), approach(3), propensity(1), step(1), method(1), incapacity(1), means(1)</i>	<i>ability(208), tendency(73), capacity(36), inability(31), way(26), propensity(21), potential(21), opportunity(21), method(16), disability(10), step(5)</i>	<i>ability(364), opportunity(76), power(54), inability(34), potential(31), flexibility(18), capacity(16), failure(16), way(13), tendency(13), propensity(16), trend(3), approach(3), capability(3)</i>
	Neutral	<i>fact(406), case(12), truth(4), thing(1), reality(1),</i>	<i>fact(208), thing(5),</i>	<i>fact(364), reality(13), case(13), truth(3), thing(3),</i>

	Partitive	<i>function(11), aspect(), example(7), effect(5), condition(5), feature(4), factor(4), period(3), part(3), version(1), impact(1), element(1), characteristic(1), attribute(1), extent(1), competence(1), principal(1)</i>	<i>extent(26), effect(10), importance(5), condition(5), function(5), feature(5), example(5), perspective(5), scenario(5), field(5)</i>	<i>extent()26, feature(10), function(8), significance(5), effect(5), quality(3), importance(3), impact(3), form(3), factor(3), element(3), aspect(3), example(3), perspective(3), condition(3), context(3), scenario(3)</i>
Mental process	Evidential	<i>evidence(39), proof(16), indication(12), sign(4), indicator(1)</i>	<i>evidence(68), proof(5), indicator(5)</i>	<i>proof(5), indication(5), evidence(23)</i>
	Conceptual	<i>notion(68), theory(23), hypothesis(16), concept(9), knowledge(5), rule(3), philosophy(1), criterion(1), principle(1), dogma(1), myth(1), tenet(1), doctrine(1), provision(1)</i>	<i>theory(36), notion(31), knowledge(21), principle(10), hypothesis(5), concept(5), myth(5), standard(5)</i>	<i>notion(57), knowledge(18), theory(16), principle(16), concept(10), hypothesis(5), conception(5), philosophy(5), tenet(3), rule(3), doctrine(3), myth(3), provision(3), stipulation(3)</i>
	Creditive	<i>idea(152), belief(99), view(62), assumption(35), sense(23), impression(20), opinion(16), premise(12), agreement(12), thought(9), feeling(9), realization(8), ground(7), objection(7), conviction(7), concern(7), image(7), position(4), stance(4), perception(4), awareness(4), consensus(4), basis(4), definition(4), recognition(3), illusion(3), assurance(3), admission(3), discovery(3), ideal(1),</i>	<i>view(57), assumption(42), belief(36), idea(21), premise(16), feeling(16), finding(16), discovery(16), awareness(10), concern(10), consensus(10), understanding(5), supposition(5), stance(5), postulation(5), prediction(5), ideal(5), ground(5), perception(5), legitimacy(5), realization(5), thought(5), opinion(5), trust(5), acceptance(5), fallacy(5), misconception(5)</i>	<i>idea(150), belief(85), view(75), assumption(54), concern(21), basis(21), presumption(18), opinion(16), premise(13), sense(13), consensus(13), thought(10), ground(10), feeling(8), conviction(8), recognition(5), realization(5), perception(5), impression(5), faith(5), exception(5), attitude(5), validation(3), definition(3), supposition(3), stance(3),</i>

	<i>insight(1), fiction(1), faith(1), consideration(1), thinking(1), inference(1), stereotype(1), acknowledgement(1), admittance(1), ignorance(1), supposition(1), reflection(1), opposition(1), enlightenment(1), innovation(1), , denial(1), consciousness(1), fallacy(1), reference(1), understanding(1), concession(1), confidence(1), consent(1), misconception(1), meaning(1)</i>		<i>speculation(3), position(3), negation(3), wisdom(3), inference(3), finding(3), awareness(3), analysis(3), acknowledgement(3) , acceptance(3), trust(3), affirmation(3), agreement(3), illusion(3), understanding(3), confidence(3), permission(3), misconception(3)</i>
Dubitative	<i>doubt(11), scepticism(1), suspicion(1),</i>	<i>doubt(21), susceptibility(5),</i>	<i>doubt(36),</i>
Emotive	<i>fear(9), desperation(4), worry(1), hesitation(1), grace(1), eagerness(1), desperateness(1), anxiousness(1), frustration(1)</i>	<i>desperation(5)</i>	<i>worry(5), fear(3)</i>
Volitional	<i>desire(79), purpose(20), determination(16), will(12), objective(12), aim(11), intention(9), hope(9), drive(9), reluctance(8), incentive(7), wish(5), willingness(4), unwillingness(4), solution(4), goal(4), ambition(4), policy(3), motivation(3), impulse(3), impetus(3), scheme(1), readiness(1), plan(1), aspiration(1), priority(1), interest(1), inclination(1), demand(1), appeal(1), expectation(1),</i>	<i>aim(52), purpose(26), desire(21), incentive(21), readiness(10), expectation(10), willingness(5), solution(5), policy(5), object(5), key(5), intention(5), hope(5), object(5), preference(5), motivation(5), drive5(), initiative(5), plan(5), reluctance(5)</i>	<i>will(47), desire(47), intention(44), aim(44), purpose(28), willingness(28), incentive(21), unwillingness(8), reluctance(8), hope(8), goal(8), expectation(8), motive(8), policy(8), solution(5), objective(5), object(5), impetus(5), drive(5), wish(3), subject(3), strategy(3), trick(3), projection(3), motivation(3), interest(3), incitement(3), determination(3), aspiration(3), demand(3), inspiration(3), repress</i>

		<i>focus(1), initiative(), key(1), motive(1), prospect(1), audacity(1), subject(1) rumor(1), metaphor(1), irony(1)</i>		<i>ion(3), willing(3), prospect(3)</i>
Verbal process	Propositional			<i>dictum(3), irony(3)</i>
	Assertive	<i>claim(52), argument(50), point(42), conclusion(40), decision(39), suggestion(32), assertion(19), statement(12), implication(12), , explanation(9), interpretation(8), thesis(7), response(7), observation(7), account(5), comment(4), acceptance(4), demonstration(3), hint(3), topic(1), theme(1), reply(1), illustration(1), discussion(1), discourse(1), message(1), contention(1), allegation(1), information(1), line(1), proclamation(1), accusation(1), reminder(1), criticism(8)</i>	<i>argument(26), implication(26), claim(21), suggestion(10), point(16), decision(16), response(10), conclusion(10), answer(10), observation(5), justification(5), explanation(5), report(5), criticism(16)</i>	<i>argument(93), point(36), conclusion(31), decision(28), claim(18), statement(13), suggestion(10), assertion(8), thesis(5), response(5), proclamation(5), justification(5), contention(5), reminder(3), remark(3), judgement(3), line(3), implication(3), explanation(3), counterpoint(3), answer(3), accusation(3), advice(3), interpretation(3), message(3), criticism(8)</i>
	Rogative	<i>question(3).</i>	<i>question(10)</i>	<i>question(3), urge(3), Proposal(3), proposition(5)</i>
	Directive	<i>urge(4), proposal(3), proposition(7)</i>		<i>promise(18), guarantee(8), commitment(8), caveat(8),</i>
	Commissive	<i>promise(1), guarantee(3), commitment(1), warning(1)</i>	<i>guarantee(10)</i>	
Material process	Expressive			
	General	<i>event(1), campaign(1), movement(1), measure(1)</i>	<i>development(5), measure(5)</i>	<i>continuation(3), move(3), campaign(3), movement(3), development(3)</i>
	Specific	<i>attempt(239), effort(43), quest(7), insistence(4), endeavour(4), option(3), change(3),</i>	<i>attempt(104), effort(36), process(5), option(5), defence(5), choice(5), encouragement(5), endeavour(5),</i>	<i>attempt(132), effort(34), choice(5), work(3), support(3), quest(3), process(3), option(3), habit(3),</i>

	<i>fight(1), rush(1), process(1), defence(1), choice(1), charge(1), alternative(1), reaction(1), pretence(1), punishment(1), instruction(1), assessment(1), devotion(1), action(1), contribution(1)</i>	<i>contribution(5)</i>	<i>endeavour(3), defence(3), contribution(3), alternative(3)</i>
Attitudinal	<i>failure(48), struggle(12), mistake(7), trouble(3), success(1), victory(1), crisis(1), conspiracy(1)</i>	<i>mistake(5), fault(5), failure(5)</i>	<i>failure(38), struggle(8), risk(8)</i>

Appendix 8

Table 4.26: Types and sub-types of transitivity process construed by distinct anaphoric shell nouns across three disciplinary domains (rounded figures normalized per million words)

Types of transitivity process	Sub-types of transitivity process	SN types and frequencies in AH	SN types and frequencies in NS	SN types and frequencies in SS
Attributive relational process	Comparative	<i>distinction(5), difference(4), discrepancy(1), conflict(1), contrast(1),</i>	<i>difference(5), distinction(5),</i>	<i>distinction(10), difference(10),</i>
	Attitudinal	<i>problem(7), stigma(1), difficulty(1), significance(1)</i>	<i>problem(57), predicament(10), difficulty(5),</i>	<i>problem(23), advantage(5), burden(3),</i>
	Epistemic	<i>truth(1), possibility(1), danger(1), reality(1),</i>	<i>risk(5),</i>	<i>risk(5), uncertainty(3), possibility(3),</i>
	Deontic	<i>role(4), freedom(3), challenge(1), requirement(1), task(1), need(1),</i>	<i>role(16), task(5),</i>	<i>role(10), independence(8), requirement(5), right(3), necessity(3), challenge(3), task(3),</i>
	Dynamic	<i>Inability(1), Propensity(1),</i>	<i>trend(21), ability(5), tendency(5),</i>	<i>trend(8), tendency(5), potential(5), opportunity(5), ability(3),</i>
Identifying relational process	Neutral	<i>phenomenon(4)</i>	<i>phenomenon(16), fact(5),</i>	<i>fact(8), phenomenon(2),</i>
	Partitive	<i>part(1), factor(1), example(2), aspect(7),</i>	<i>part(5), factor(21), element(5), example(10),</i>	<i>example(10), aspect(18), basis(8), factor(5), feature(3), facet(3),</i>
Mental process	Evidential	<i>evidence(1),</i>	<i>evidence(57), finding(5), observation(5),</i>	<i>evidence(8), implication(3),</i>
	Conceptual	<i>idea(21), theory(13), concept(11), issue(11), image(5), hypothesis(5), principle(5), theme(3), thought(1), myth(1), wisdom(1), subject(1),</i>	<i>idea(52), theory(36), issue(21), concept(21), hypothesis(16), position(10), topic(10), subject(5), notion(5), theme(5), rule(5), mystery(5), conundrum(5),</i>	<i>idea(39), theory(36), issue(26), notion(21), image(10), concept(10), conception(8), topic(5), theme(5), stereotype(5), subject(5), rule(5), rationale(3), myth(3), logic(3), doctrine(3), principle(3),</i>
	Creditive	<i>point(20), view(17), interpretation(11), notion(7), attitude(5), knowledge(4),</i>	<i>view(47), knowledge(31), point(10), viewpoint(10),</i>	<i>view(46), point(36), perspective(18), reasoning(13), assumption(10),</i>

		<i>fascination(3), feeling(3), experience(3), illusion(3), understanding(1), sentiment(1), realization(1), perception(1), opinion(1), belief(1), assumption(1), anticipation(1), representation(1), philosophy(1), deception(1), contemplation(1), consideration(1), conception(1), awareness(1), analysis(1),</i>	<i>assumption(10), understanding(5) , thought(5), illusion(5), feeling(5), belief(5), awareness(5), presumption(5), perception(5)</i>	<i>note(10), feeling(8), presumption(5), opinion(5), knowledge(5), attitude(5), understanding(5), stand(3), stance(3), sentiment(3), recognition(3), reading(3), position(3), assumption(3), sentiment(3), direction(3), conviction(3), consideration(3), confusion(3),</i>
	Dubitative			
	Emotive	<i>Relief(1), concern(1), sadness(1),</i>	<i>concern(10),</i>	<i>concern(8), fear(3), disappointment(3),</i>
	Volitional	<i>policy(5), purpose(4), ideal(3), desire(3), strength(1), strategy(1), solution(1), programme(1), interest(1), hope(1), determination(1), decision(1), willingness(1), readiness(1)</i>	<i>purpose(16), aim(10), interest(5), compromise(5), resolve(5), decision(5),</i>	<i>policy(8), goal(8), ideal(5), decision(5), solution(3), rejection(3), desire(3), aim(3), target(3), strategy(3), scheme(3), reluctance(3),</i>
Verbal process	Propositional	<i>definition(3), information(1),</i>	<i>definition(47), information(16)</i>	<i>definition(13), Information(8),</i>
	Assertive	<i>argument(23), statement(15), claim(9), conclusion(3), judgement(1), interrogation(1), evaluation(1), description(1), denial(1), contradiction(1), acknowledgement(1), acceptance(1), pronouncement(1), defence(1), assessment(1),</i>	<i>statement(26), criticism(10), explanation(10), description(5), counterargumen t(5), claim(5), argument(5), judgment(5), comment(5), response(5),</i>	<i>argument(36), statement(23), criticism(10), conclusion(8), claim(5), justification(5), response(3),</i>
	Rogative	<i>question(13),</i>	<i>question(21),</i>	<i>question(26),</i>
	Directive	<i>suggestion(3), order(3), proposal(1), appeal(1), advice(1), application(1),</i>	<i>suggestion(10), proposal(5), directive(5),</i>	<i>application(8), tip(3), appeal(3),</i>
	Commissive	<i>threat(3),</i>		<i>threat(3),</i>

Material process	Expressive			
	General	<i>change(11), trend(8), tradition(4), action(3), choice(3),</i>	<i>act(16), action(10), reaction(10), move(5), activity(5),</i>	<i>change(16), act(10), tradition(3), preference(3), move(3), habit(3), event(3), alternative(3), action(3), choice(3), quest(3),</i>
	Specific	<i>enterprise(1), attempt(1), struggle(1), reaction(1), move(1)</i>		
	Attitudinal	<i>victory(1), mistake(1), failure(1),</i>		<i>failure(5),</i>

Appendix 9

Table 4.29: Changes of types of transitivity process across levels of study in AH (normalized per million words)

Types of transitivity process	Sub-types of transitivity process	SN types and frequencies in AH (L1)	SN types and frequencies in AH (L3)
Attribute relational process	comparative	<i>difference(5), distinction(2), discrepancy(2)</i>	<i>difference(18), distinction(9), contrast(3),</i>
	attitudinal	<i>problem(15), weakness(5), flaw(2), difficulty(2), beauty(2), irony(2)</i>	<i>problem(30), difficulty(9), weakness(3), advantage(3),</i>
	Epistemic	<i>possibility(36), chance(22), danger(5), truth(5), option(2), danger(2), certainty(2),</i>	<i>possibility(33), chance(15), reality(6), probability(3), truth(3),</i>
	deontic	<i>need(68), right(48), duty(17), freedom(15), role(12), obligation(10), responsibility(5), pressure(5), necessity(5), task(5), requirement(2), challenge(2)</i>	<i>need(81), freedom(18), necessity(12), mission(9), requirement(6), challenge(6), task(3), responsibility(3), duty(3), job(3), role(9),</i>
	Dynamic	<i>ability(169), opportunity(39), power(22), inability(22), tendency(17), capacity(15), propensity(5), step(2), potential(2), incapacity(2), approach(2),</i>	<i>ability(169), inability(75), power(27), opportunity(27), tendency(21), capacity(21), temptation(6), step(6), potential(6), propensity(6),</i>
Identifying relational process	neutral	<i>fact(467), phenomenon(5), thing(2),</i>	<i>fact(283), Phenomenon(3),</i>
	partitive	<i>example(19), function(7), aspect(7), factor(2), feature(2), attribute(2), part(2),</i>	<i>example(30), aspect(12), function(6), feature(6)</i>
Mental process	evidential	<i>evidence(22), proof(12), indication(10), reminder(7), sign(5), indicator(2), observation(5), implication(2)</i>	<i>evidence(36), implication(24), proof(15), indication(15), observation(9),</i>
	conceptual	<i>point(60), notion(44), theory(51), concept(22), thesis(12), hypothesis(15), issue(12), image(12), knowledge(7), philosophy(2), myth(2), rule(2), metaphor(2), doctrine(2), wisdom(2), subject(2), paradox(2)</i>	<i>notion(72), hypothesis(27), theory(21), idea(21), concept(15), issue(9), philosophy(6), theme(6), knowledge(3), image(3), theme(3), subject(3), thought(3), rationality(3), myth(3)</i>
	Creditive	<i>idea(147), belief(85), view(87), assumption(24), realization(12), opinion(12), impression(12), feeling(15), interpretation(15), concern(10), premise(10), illusion(10), thought(7), position(7), knowledge(7), recognition(5), principle(5),</i>	<i>idea(190), view(66), assumption(48), point(45), opinion(30), impression(27), principle(27), notion(15), attitude(12), thought(9), position(9), interpretation(9), objection(9), ground(9), feeling(9), agreement(9),</i>

		<p><i>perception(5), misconception(5), conviction(5), consensus(5), agreement(5), experience(5), consideration(5), understanding(2), tenet(2), stance(2), prospect(2), fiction(2), faith(2), dogma(2), discovery(2), prospect(2), insight(2), philosophy(2), deception(2), contemplation(2), conception(2), awareness(2), analysis(2),</i></p>	<p><i>realization(6), fascination(6), premise(6), inference(6), doctrine(6), conviction(6), basis(6), acceptance(6), thinking(3), stereotype(3), stance(3), reference(3), perception(3), illusion(3), fallacy(3), faith(3), discovery(3), consciousness(3), awareness(3), admittance(3), admission(3), reflection(3), interpretation(6), consensus(3), significance(3), understanding(3), sentiment(3), realization(3), perception(3), belief(3), anticipation(3), significance(3), suspicion(3), doubt(3), fear(9), frustration(6), worry(3), anxiousness(3),</i></p>
	dubitative	<i>doubt(17),</i>	
	emotive	<i>fear(10), desperateness(2), relief(2), concern(2), sadness(2),</i>	
	volitional	<i>desire(87), decision(43), purpose(22), refusal(22), determination(19), objective(15), aim(12), hope(12), willingness(10), policy(10), wish(7), intention(7), solution(7), reluctance(7), eagerness(5), determination(5), inclination(5), incentive(5), unwillingness(2), motivation(2), key(2), goal(2), ideal(2), strength(5), strategy(2), programme(2), interest(2),</i>	<i>desire(78), right(27), decision(36), hope(24), reluctance(21), purpose(18), willingness(15), intention(12), incentive(12), unwillingness(9), determination(9), ambition(9), aim(9), ideal(9), wish(6), motivation(6), solution(3), reluctance(3), readiness(3), plan(3), objective(3), motive(3), initiative(3), impetus(3), goal(3), expectation(3), drive(3), appeal(3), interest(3), purpose(6), readiness(3), policy(3), message(3),</i>
Verbal process	propositional	<i>information(5), definition(5), rumor(2),</i>	
	assertive	<i>argument(73), conclusion(58), claim(56), statement(27), assertion(10), comment(7), criticism(7), objection(5), explanation(2), demonstration(2), admission(2), denial(4), accusation(2), judgement(2), description(2), contradiction(2),</i>	<i>suggestion(75), claim(72), argument(72), conclusion(27), assertion(21), statement(21), response(9), criticism(9), insistence(9), explanation(6), demonstration(6), proclamation(3), contention(3), allegation(3),</i>

		<i>acknowledgement(2), acceptance(2),</i>	<i>assessment(6), acknowledgement(3), pronouncement(3), defence(3),</i>
	rogative directive	<i>question(36), suggestion(12), urge(5), proposal(5), demand(2), proposition(7), application(2)</i>	<i>question(9), urge(3), proposition(3), instruction(3), order(6), proposal(3), appeal(3), advice(3),</i>
	commissive	<i>guarantee(5), threat(5), assurance(2),</i>	<i>guarantee(3), commitment(3), assurance(3),</i>
Material process	expressive general	<i>change(19), trend(10), action(5), campaign(2),</i>	<i>tradition(9), custom(3), convention(3), change(6), trend(6), priority(3), action(3),</i>
	specific	<i>attempt(155), effort(34), quest(7), choice(7), insistence(5), endeavour(5), rush(2), option(2), alternative(2), enterprise(2), convention(2), measure(2),</i>	<i>attempt(259), effort(27), struggle(21), quest(6), reaction(3), option(3), endeavour(3), choice(6), pretence(3), conspiracy(3), reaction(3), move(3)</i>
	attitudinal	<i>failure(44), struggle(10), mistake(7), innovation(2), revolution(2), dilemma(2),</i>	<i>failure(75), mistake(6), success(3), struggle(3),</i>

Appendix 10

Table 4.30: Changes of types of transitivity process across levels of study in NS (normalized per million words)

Types of transitivity process	of	Sub-types of transitivity process	of	SN types and frequencies in NS(L1)	SN types and frequencies in NS(L3)
Attribute relational process	relational	comparative	attitudinal	<i>similarity(9), difference(9), problem(84), advantage(37), benefit(19), limitation(9), issue(9), disadvantage(9), demerit(9), beauty(9), attraction(9), difficulty(9), likelihood(9), chance(9), risk(9),</i>	<i>difference(12), distinction(12), problem(202), advantage(83), disadvantage(47), importance(24), limitation(12),</i>
				Epistemic	<i>possibility(36), chance(36),</i>
		deontic	<i>need(130), role(37), right(19), responsibility(19), duty(19), job(9), challenge(9),</i>	<i>need(107), necessity(24), task(12), requirement(12), pressure(12), role(12), right(12),</i>	
		Dynamic	<i>ability(241), tendency(93), opportunity(46), potential(37), capacity(28), disability(19), power(9), method(9), inability(9), freedom(9),</i>	<i>ability(356), tendency(95), inability(83), capacity(59), propensity(47), opportunity(36), freedom(36), trend(36), method(24), potential(12),</i>	
Identifying relational process	relational	neutral		<i>fact(149), thing(9), phenomenon(28),</i>	<i>fact(308),</i>
		partitive		<i>feature(9), example(9), part(9), factor(9),</i>	<i>factor(36), example(24),</i>
Mental process	relational	evidential		<i>evidence(84), indicator(9), finding(9), observation(9),</i>	<i>evidence(213), proof(12), implication(47), finding(71), observation(12),</i>
				conceptual	<i>theory(93), knowledge(65), notion(19), scenario(9), idea(84), point(19), concept(19), topic(9), theme(9), rule(9), mystery(9), issue(9), conundrum(9),</i>
		Creditive		<i>idea(186), view(92), belief(37), assumption(28), feeling(28), discovery(19), awareness(46), thought(19), stance(9), perspective(9), perception(9), opinion(9), hypothesis(9), , consensus(9), viewpoint(19), understanding(9), illusion(9),</i>	<i>idea(119), assumption(71), view(71), premise(36), belief(36), point(24), understanding(12), supposition(12), realization(12), principle(12), prediction(12), postulation(12), myth(12), misconception(12), ground(12), feeling(12), fallacy(12), discovery(12), consensus(12), presumption(12), perception(12),</i>
		dubitative emotive		<i>doubt(28), concern(28),</i>	<i>concern(24), desperation(12),</i>

	volitional	<i>aim(65), purpose(46), expectation(28), decision(19), desire(19), willingness(19, solution(9), reluctance(9), readiness(9), plan(9), object(9), key(9), intention(9), initiative(9), hope(9), drive(9), ideal(9), purpose(28), interest(9), compromise(9),</i>	<i>aim(59), incentive(47), decision(24), readiness(12), purpose(12), motivation(12), expectation(12), resolve(12), function(12),</i>
Verbal process	propositional	<i>definition(56), information(9),</i>	<i>definition(36), information(24),</i>
	assertive	<i>claim(28), response(19), point(9), explanation(9), conclusion(9), argument(19), answer(9), criticism(28), defence(9), statement(37), explanation(9), description(9), counterargument(9),</i>	<i>argument(47), claim(36), suggestion(47), acceptance(24), report(12), justification(12), conclusion(12), answer(12), statement(12), objection(12), judgment(12), explanation(12), criticism(12), comment(12),</i>
	rogative	<i>question(19),</i>	<i>question(36),</i>
	directive	<i>suggestion(9), proposal(9),</i>	<i>directive(12),</i>
	commissive	<i>guarantee(9), threat(9),</i>	
Material process	expressive		
	general	<i>trend(9), activity(9), action(9), act(9),</i>	<i>reaction(24), act(24), move(12), action(12), preference(12),</i>
	specific	<i>attempt(130), effort(19), option(9), endeavour(9), choice(9), measure(9)</i>	<i>attempt(119), effort(47), measure(12),</i>
	attitudinal	<i>failure(9),</i>	<i>mistake(12), fault(12),</i>

Appendix 11

Table 4.31: Changes of types of transitivity process across levels of study in SS (normalized per million words)

Types of transitivity process	Sub-types of transitivity process	SN types and frequencies in SS(L1)	SN types and frequencies in SS(L3)
Attribute relational process	comparative	<i>distinction(25), difference(13),</i>	<i>difference(44), distinction(4),</i>
	attitudinal	<i>problem(19), weakness(6), disadvantage(6), benefit(6), importance(6), irony(6),</i>	<i>problem(61), danger(17), advantage(13), weakness(4), flaw(4), difficulty(4), crux(4), burden(4)</i>
	Epistemic	<i>possibility(32), probability(25), chance(19), likelihood(6), risk(13),</i>	<i>possibility(26), reality(26), truth(9), chance(9), probability(4), Risk(9), uncertainty(4)</i>
	deontic	<i>need(134), right(83), obligation(51), duty(38), task(13), responsibility(6), necessity(6), mission(6), role(6),</i>	<i>need(91), right(91), duty(70), role(35), responsibility(13), requirement(13), obligation(13), task(9), challenge(9), requirement(9), necessity(4), challenge(4),</i>
Identifying relational process	neutral partitive	<i>ability(185), potential(64), opportunity(38), propensity(32), power(32), inability(25), trend(25), tendency(25), capacity(19), freedom(13), approach(6), fact(446), phenomenon(13), feature(25), factor(19), aspect(13), exception(13), basis(13), example(13),</i>	<i>ability(100), opportunity(96), power(48), potential(48), inability(30), freedom(17), capacity(17), tendency(17), pressure(13), trend(9), propensity(4), capability(4), fact(300), thing(9), basis(39), aspect(22), example(13), factor(4),</i>
		<i>evidence(25), implication(13), indication(6), finding(6), reminder(6),</i>	<i>evidence(95), poof(9),</i>
Mental process	conceptual	<i>notion(80), theory(83), issue(57), conception(25), knowledge(25), misconception(6), point(25), perspective(25), image(19), concept(19), doctrine(13), topic(6), subject(6), rule(6), rationale(6), myth(6), logic(6),</i>	<i>notion(30), theory(22), knowledge(17), issue(17), concept(13), conception(13), stereotype(9), theme(9), conception(4), subject(4), topic(4), subject(4), rule(4), principle(4), notion(4), image(4),</i>
	Creditive	<i>idea(273), assumption(102), belief(95), view(95), presumption(38), faith(38), awareness(38), feeling(38), attitude(38), premise(25), principle(19), motive(19), stance(13), ground(13), conviction(13),</i>	<i>idea(87), view(95), assumption(83), belief(52), point(39), thought(17), reasoning(17), note(17), perspective(17), principle(13), presumption(13), consensus(13), recognition(9), attitude(9), opinion(9), realization(9),</i>

		<p><i>consensus(13), understanding(13), thesis(13), tenet(6), supposition(6), rule(6), realization(6), prospect(6), position(6), perception(6), myth(6), hypothesis(6), interpretation(6), stand(6), stance(6), sentiment(6), recognition(6), reasoning(6), reading(6), position(6), point(6), assumption(6),</i></p>	<p><i>impression(9), ground(9), faith(9), conviction(9), attitude(9), wisdom(4), speculation(4), prospect(4), premise(4), inference(4), illusion(4), scenario(4), significance(4), confidence(4), projection(4), understanding(4), sentiment(4), conviction(4), consideration(4), confusion(4),</i></p>
	dubitative	<i>doubt(38),</i>	<i>doubt(26),</i>
	emotive	<i>concern(13),</i>	<i>concern(39), fear(13), worry(9),</i>
	volitional	<p><i>desire(83), intention(76), willingness(70), aim(57), purpose(45), decision(45), incentive(32), goal(32), interest(19), unwillingness(13), function(13), trick(6), strategy(6), reluctance(6), policy(6), objective(6), hope(6), drive(6), inspiration(6), agreement(6), solution(6), rejection(6),</i></p>	<p><i>aim(48), desire(35), decision(30), intention(22), incentive(22), purpose(17), expectation(17), hope(13), reluctance(13), policy(22), object(9), ideal(9), solution(9), wish(4), willing(4), unwillingness(4), strategy(9), objective(4), interest(4), incitement(4), goal(4), drive(4), target(4), scheme(4),</i></p>
Verbal process	propositional	<i>definition(19),</i>	<i>information(17), definition(9), message(4),</i>
	assertive	<p><i>argument(115), point(83), conclusion(45), criticism(38), statement(32), claim(25), assertion(13), response(6), remark(6), proclamation(6), explanation(6), contention(6), accusation(6), affirmation(6), acceptance(6), justification(6),</i></p>	<p><i>argument(148), conclusion(39), statement(39), point(30), claim(22), justification(17), suggestion(13), thesis(9), judgement(9), acknowledgement(9), response(13), explanation(4), dictum(4), counterpoint(4), contention(4), assertion(4), answer(4), criticism(4), objection(4), contribution(4),</i></p>
	rogative	<i>question(57),</i>	<i>question(9),</i>
	directive	<i>suggestion(6)</i>	<i>application(13), proposition(9), urge(4), advice(4), demand(4), proposal(4), tip(4), appeal(4),</i>
	commissive	<p><i>promise(51), commitment(13), caveat(13),</i></p>	<i>commitment(17), guarantee(13), caveat(4), threat(4),</i>
Material process	expressive		
	general	<p><i>change(25), act(25), tradition(6), preference(6), move(6), habit(6), event(6),</i></p>	<i>change(9), choice(4),</i>

specific	<i>alternative(6), action(6), attempt(146), struggle(25), effort(19), endeavour(13), initiative(13), move(6), habit(6), alternative(6), quest(6),</i>	<i>attempt(117), effort(44), option(17), quest(4), choice(4), initiative(13),</i>
attitudinal	<i>failure(45),</i>	<i>failure(78),</i>

