SIGNALLING NOUNS IN CORPORA OF LEARNER AND NATIVE SPEAKER ENGLISH

By

ABDULRAHMAN MOHAMMED ABDULRAHMAN ALHARBI

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 American and Canadian Studies

College of Arts and Law

University of Birmingham

August 2019

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This unpublished thesis/dissertation is copyright of the author and/or third parties.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 thesis investigates the use of signalling nouns (SNs) in writing by Arabic speakers who are learning English and by native speakers of English. The thesis has three aims: to compare the frequency of tokens and types of SNs in the two corpora; to compare the frequency of tokens and types of SNs in the language of learners at four levels of proficiency; to study in detail the usage of selected SNs in the two corpora. To address these aims, the thesis uses both quantitative and qualitative approaches.

The thesis examines the frequency of tokens and types of SNs in the learner and native speaker corpora. The key quantitative finding is that the learners were found to use fewer SN tokens and types compared to the native speakers. The thesis also investigates the frequency of SN tokens and types across the four levels of proficiency of the learners. The rationale for doing this investigation is to test the hypothesis that SNs are a developmental phenomenon. The quantitative findings show that the frequency of SN tokens and types increased as the level of proficiency of the learners increased. The thesis argues that the difference in proficiency of writers is responsible for the variation in the use of SNs. It also suggests that writing quality can be enhanced by the use of SNs.

The thesis then undertakes three detailed studies of specific SNs, with particular focus on: general SNs, stance SNs and causal SNs. First, three frequent and shared general SNs in the two corpora are thoroughly examined. The study shows the uses of these nouns in discourse and compares the learners with the native speakers of English in their uses of these nouns. Similarities were present in the uses of these nouns. Both groups of writers were found to frequently use these nouns within clause boundaries, to specify them in clauses and to use the meaning categories of these nouns with approximately equal frequency. Next, stance SNs and causal SNs used in the two corpora are investigated. These studies explore how stance and causality through SNs are constructed, identify the patterns associated with the construction of stance and causality through SNs, and compare the learners with the native speakers of English in their uses of stance SNs and causal SNs. The learners were found to use stance SNs and causal SNs significantly less than the native speakers. The learners were found to be less varied in their uses of stance SNs and causal SNs.

The thesis discusses the findings with relevant literature and offers pedagogical implications and suggestions for further research on SNs.

DEDICATION

This thesis is dedicated to my inspiring mother and father. It is also dedicated to my loving wife and little gorgeous son (Yamin).

ACKNOWLEDGEMENTS

Having completed my thesis, it is time to acknowledge the support that I have received throughout the different stages of my PhD journey.

First and foremost, I would like to express my profound gratitude to my supervisor, Professor Susan Hunston. She has been an exceptional supervisor and I have been privileged to have her as my PhD supervisor. She has been always there supporting me at all levels (personal and intellectual) throughout the various stages of my thesis writing. I am grateful for her support, patience, sound advice, and generous and constructive feedback on my work. The thesis would have been impossible to come to fruition without her professional supervision.

I would also like to immensely thank my parents for being supportive and patient. They have been an incredible source of encouragement.

I would also like to thank my wife (Shahad) for her endless support, love and encouragement. I am incredibly lucky to have her in my life. I also extend my thanks to my brothers and sisters for continuing to ask after me and for encouraging me to do my best.

Finally, my special thanks go to my friend and fellow doctoral colleague (Waleed Alsubhi). His company has been a pleasure. I will not forget the good many times that we have spent together at Costa, sipping coffee and discussing ideas related to our dissertations.

Table of Contents

CHAPTER 1: INTRODUCTION....1
1.1. Overview....1
1.2. The background to the current thesis....1
1.3. The rationale for the study....3
1.4. The aims and the research questions of my study....6
1.5. The corpora used and the methodology of the study....7
1.6. Outline of the thesis....7
CHAPTER 2: LITERATURE REVIEW....9
2.1. Overview....9
2.2. Sinclair's text organisation model and SNs....9
2.3. Learner corpus research: issues and contributions to understanding learner language....12
2.4. SNs in the literature....19
2.4.1. Early influential precursors to the phenomenon of SNs....20
2.4.2. Further development in the phenomenon of SNs....21
2.4.3. Flowerdew and Forest (2015) and SNs....29
2.5. Lexical cohesion and SNs....31
2.6. Learners and the use of SNs....33
2.7. Towards a definition of SNs....36
CHAPTER 3: METHODOLOGY....40
3.1. Overview....40
3.2. Description of the corpora used in the study....40
3.3. Steps of identification of SNs....45
3.4. General criteria that distinguish the class of SNs....47
3.4.1. SNs in formulaic phrases....49
3.4.2. Partitive SNs....51
3.4.3. Bidirectional instances of SNs....52
3.5. Repeated instances of SNs....53
3.6. Quantification and further investigation of SNs....56
3.7. Summary of the chapter....56

CHAPTER 4: THE FREQUENCY OF SNS IN THE LEARNER AND THE NATIVE SPEAKER CORPORA 58

4.1. Overview 58

4.2. Reporting of frequency of SNs in the two corpora 58

4.3. Frequency of SNs in the learner corpus 59

4.4. Frequency of SNs in the native speaker corpus 69

4.5. SNs in the learner and the native speaker corpora: comparison 72

4.5.1. Frequency of SNs 73

4.5.2. SN types 86

4.6. Summary of the findings and concluding remarks 91

CHAPTER 5: CLOSE ANALYSIS OF THREE MOST FREQUENT AND SHARED SNS IN THE LEARNER AND NATIVE SPEAKER CORPORA 93

5.1. Overview 93

5.2. Reference categories and textual meanings identified: a brief overview 93

5.3. *Time(s)* 95

5.3.1. Reference categories of *time(s)* 95

5.3.2. Meanings associated with the SN *time(s)* 99

5.4. *Thing(s)* 104

5.4.1. Reference categories of *thing(s)* 104

5.4.2. Uses of the SN *thing(s)* 108

5.5. *Way(s)* 117

5.5.1. Reference categories of *way(s)* 117

5.5.2. Meanings associated with the SN *way(s)* 120

5.6. Summary of the findings of the chapter 123

CHAPTER 6: SNS AND STANCE 124

6.1. Overview 124

6.2. SSNs: a literature review 124

6.3. Patterns and stance nouns 129

6.3.1. Construction of stance within clause boundary 130

6.3.2. Construction of stance across clause boundaries through prospection and encapsulation 137

6.4. Analysis and findings of SSNs in the two corpora 147

6.4.1. *Opinion(s)* 148

6.4.2. *Fact(s)* 157
6.4.3. *Idea(s)* 161
6.5. Summary of the chapter 168
CHAPTER 7: SNS AND CAUSALITY 170
7.1. Overview 170
7.2. The concept of causality 170
7.3. Patterns and CSNs 172
7.3.1. Signalling of causality within clause boundary 174
7.3.2. Signalling of causality across clause boundary 180
7.4. Analysis and findings of CSNs in the two corpora 191
7.4.1. *Reason(s)* 192
7.5. Summary of the findings and concluding remarks 206
CHAPTER 8: DISCUSSION AND CONCLUDING REMARKS 208
8.1. Overview 208
8.2. Summary and discussion of findings 208
8.2.1. What is the frequency of SNs in the learner and native speaker corpora, and how the two corpora compare? 208
8.2.2. What is the frequency of SNs in the learner corpus across the proficiency levels of the learners? 210
8.2.3. How are the three most frequent and shared SNs in the learner and native speaker corpora used in relation to the reference categories these nouns may fall into, the grammatical forms of specification and what meanings of these nouns can be identified? How do the two corpora compare with each other with these regards? 211
8.2.4. What are the patterns that can be associated with SNs signalling causality and those signalling stance? How comparable are the two corpora in their uses of SNs to signal those two discourse functions? 214
8.3. Pedagogical implications 218
8.3.1. Refining current advice: the SN *thing(s)* 223
8.3.2. Expanding the repertoire: SNs and causality 225
8.3.3. Introducing the personal: SNs and stance 228
8.4. Assessment of the methodology 229
8.5. Concluding remarks and suggestions for further research 230
APPENDIX A: CEPA test specifications 232
APPENDIX B: Candidate SNs in descending order according to their overall raw frequency (signalling and non-signalling) in the learner corpus 234

APPENDIX C: Non-lemmatised SNs in descending order according to raw frequency in the learner corpus……..236

APPENDIX D: Candidate SNs in descending order according to their overall raw frequency (signalling and non-signalling) in the native speaker corpus…….238

APPENDIX E: Non-lemmatised SNs in descending order according to raw frequency in the native speaker corpus…….241

APPENDIX F: Signalling and non-signalling instances of the SN *reason(s)* in the learner corpus…….244

APPENDIX G: Signalling and non-signalling instances of the SN *reason(s)* in the native speaker corpus……249

APPENDIX H: Most frequent patterns of the SN *thing* in the learner corpus…….257

APPENDIX I: Most frequent patterns of the SN *thing* in the native speaker corpus……...262

APPENDIX J: The frequency of all the adjectival types occurring with the CSN *reason(s)* across the two corpora……....265

APPENDIX K: The frequency of all the enumerative types occurring with the CSN *reason(s)* across the two corpora……....266

REFERENCES…….267

List of Figures

Figure 1-1 An extract taken from a writing textbook (Pavlik and Segal, 2007, p.9).....5

Figure 4-1 Normalised frequency of the top-5 SNs across the four levels per 100,000 words 65

Figure 4-2 Normalised frequency and total number of SN types occurring in each level.....68

Figure 5-1 Meanings associated with the specification of the SN *time(s)* in the learner and native speaker corpora 99

Figure 5-3 Meanings associated with the SN *way(s)*.....120

Figure 8-1 An extract taken from a writing textbook (Pavlik and Segal, 2007, p.9).....219

List of Tables

Table 2-1 Previous studies on the phenomenon of signalling nouns 21
Table 2-2 Schmid's syntactic patterns used in the identification of shell nouns 26
Table 3-1 Comparison of the two corpora used in the present study 44
Table 3-2 Formulaic phrases of SNs in the learner corpus 49
Table 4-1 Normalised frequency of SNs in the learner corpus per 100,000 words 59
Table 4-2 Frequencies of the four most frequent SNs in the learner corpus 62
Table 4-3 Total normalised frequency of both grammatical forms of SNs in the learner corpus 64
Table 4-4 Normalised frequency of SNs per 100,000 words across levels along with the number of SN types occurring in each level 66
Table 4-5 Total normalised frequency of SNs in the native speaker corpus 69
Table 4-6 Essay prompt for native speaker corpus 71
Table 4-7 Total normalised frequency of SNs across the two corpora 73
Table 4-8 Log-likelihood results for the total number of SNs in the two corpora 74
Table 4-9 List of SNs in the two corpora shared and unshared 86
Table 4-10 SNs as used in the highest-level learners (CEPA 5 and CEPA 6) and the native speaker corpus 90
Table 5-1 Reference categories of the SN *time(s)* across the two corpora 95
Table 5-2 Specification of the SN *time(s)* in the learner and native speaker corpora 98
Table 5-3 Raw frequency of the three meanings associated with the SN *time(s)* across the two corpora 103
Table 5-4 Frequency of *thing(s)* across the two corpora 104
Table 5-5 Reference categories of *thing(s)* across the learner and native speaker corpora 105
Table 5-6 Specification of the SN *thing(s)* in the learner and native speaker corpora 107
Table 5-7 The most frequent patterns of the SN thing across the two corpora (* indicates some instances of divergent patterns; *** refers to redundant use) 111
Table 5-8 Specification categories of the SN *way(s)* 117
Table 5-9 Specification of the SN *way(s)* in the learner and native speaker corpora 119
Table 5-10 Meanings of the SN *way(s)* across the two corpora 122
Table 6-1 Patterns associated with SSNs in the learner and native speaker corpora 146
Table 6-2 Stance signalling nouns appearing in the top-25 SNs in the two corpora 147

Table 6-3 Patterns associated with the SSN *opinion(s)* as identified in the two corpora (* refers to an erroneous use of preposition and possessive pronouns found to occur once in the learner corpus) 149
Table 6-4 Raw frequencies and percentages of reference uses of *opinion* across the two corpora 153
Table 6-5 Construction of stance with the SSN *opinion* in the two corpora 156
Table 6-6 Patterns associated with the SSN *fact* in the two corpora 158
Table 6-7 The construction of stance through the SSN *fact* in the two corpora.....161
Table 6-8 Patterns identified to occur with the SSN *idea(s)* in the learner and native speaker corpora 162
Table 6-9 The construction of stance through idea across the two corpora.....165
Table 7-1 The frequency of quantifier + CSN in the BNC.....187
Table 7-2 List of all the patterns associated with CSNs along with their frequencies in the two corpora 190
Table 7-3 Causal signalling nouns in the two corpora along with their raw and normalised frequencies 192
Table 7-4 Raw and normalised frequencies of *reason(s)* and because across the two corpora 193
Table 7-5 Total raw frequency of adjectival modifiers along with the number of types.....198
Table 7-6 Frequency of the adjectives preceding the CSN *reason* along with the TTR 200
Table 7-7 Frequency of enumerative tokens preceding the CSN *reason* along with the TTR 202
Table 7-8 Total number of vague and precise modifiers preceding the CSN *reason* along with percentages.....203

CHAPTER 1: INTRODUCTION

1.1. Overview

The aim of this chapter is to introduce the present thesis on the phenomenon of signalling nouns. Section 1.2 addresses the background to the present thesis. The rationale for focusing on signalling nouns is presented in Section 1.3. Section 1.4 presents the aims and the research questions of my study. The data used in the current thesis are briefly addressed in Section 1.5. The chapter ends with Section 1.6, which provides the outline of the thesis.

1.2. The background to the current thesis

The topic of the present study is signalling lexis in the writing of Arabic-speaking learners of English. More particularly, it concerns the use and function of one particular type of signalling lexis, that is *signalling nouns*, as used in a learner corpus. The term *signalling nouns* (henceforth SNs) is taken from Flowerdew and Forest (2015) who coined the term. In fact, their study has been influential in my thinking on SNs. SNs in the present study mean abstract/borderline nouns, which rely on the co-text for their meaning to be specified. Examples of this type of nouns are *reason*, *idea*, *thing*, *way* and *problem*. These SNs, as other SNs, can refer to previously mentioned propositions/entities (e.g. *this idea*) or propositions/entities to come (e.g. *the thing is that…*), either within or across clause boundaries. Example (1) below shows an instance of SN use taken from the learner corpus used in the present study.

(1): … and we faced another **problem**, they said that there are no rooms available for us (**Learner corpus)**

The SN *problem* (in **bold**), in example (1), refers cataphorically to a stretch of discourse (underlined), which is found across a clause boundary. This underlined stretch of language provides the specifics of the SN in question and the SN itself signals how these specifics are meant to be interpreted.

Over the past four decades, nouns such as *thing*, *idea*, *fact* and *problem* have been the subject of much research. Researchers have investigated such type of nouns from a wide range of aspects, resulting in a wide range of different nomenclatures that highlight these aspects (*general nouns* as in Halliday and Hasan (1976) and Mahlberg (2005); *vocabulary 3* as in Winter (1977); *anaphoric nouns* as in Francis (1986); *carrier nouns* as in Ivanic (1991); *advance* and *retrospective labels* as in Francis (1994); *shell nouns* as in Hunston and Francis (1999) and Schmid (2000) and *signalling nouns* as in Flowerdew (2003) and Flowerdew and Forest (2015))[1]. This on-going interest in SNs and related phenomena highlights the importance of this type of nouns in discourse. It also draws attention to a wide range of functions that these nouns might serve in discourse by employing different methodological approaches. Therefore, the aforementioned studies show similarities as well as differences in the description of what I refer to in the current thesis as SNs. In my study, I follow Flowerdew and Forest's (2015) illuminating approach towards the investigation of SNs; their approach looks at SNs from a discourse perspective, a point that, I believe, gives the broadest picture of SNs. In fact, the discourse approach has led me to extend Flowerdew and Forest's definition of SNs. My definition of SNs is discussed and presented in Section 2.7 in the literature review chapter.

[1] A more detailed review of the relevant studies on the phenomenon of SNs is provided in Chapter 2.

The previous body of research that has provided descriptions of SNs and related concepts has broadly investigated SNs and related concepts in academic genres or journalistic prose (see Benitez-Castro, 2015 for a similar argument). Similarly, studies that looked at SNs and related concepts in learner language have chiefly focused on the use of this type of nouns in academic genres. These two facts raise the question of whether SNs are only genre specific or whether they are used in a wide range of genres. Since the present study examines SNs in non-academic genres, it is hoped that the study will broaden our understanding of SNs in such genres, which is largely underexplored.

The majority of investigations into phenomena related to SNs have limited the identification of this type of nouns to automated pre-specified patterns. This methodological approach has its own limitation. It may restrict the identification of instances of SNs as well as patterns that may be associated with the phenomenon of SNs. Therefore, in the present study I have built on Flowerdew and Forest's (2015) methodological approach towards the identification of SNs which does not pre-specify patterns in the identification process. I have also made some modifications to better understand SNs in learner language. This point brings us to the rationale for my thesis, which is presented in the next section.

1.3. The rationale for the study

There are many reasons why I have chosen the phenomenon of SNs as a target linguistic area for investigation. One of these is based on my personal experience as an EFL teacher at a Saudi University. Having taught writing modules to Arabic speakers who are learning English, I have noticed that these learners often have difficulty with using signalling nouns as a resource to link stretches of text. Example (2) is an excerpt taken from a paragraph written

by an Arabic speaker who is learning English. The excerpt is taken from the learner corpus used in the present study.

(2): It combines love, respect, dignity and honour towards people and their country. **This** gives the film a well balance, so the viewers don't get bored. Troy has brought the past and history of Greece to us, so we could appreciate what they have been through.

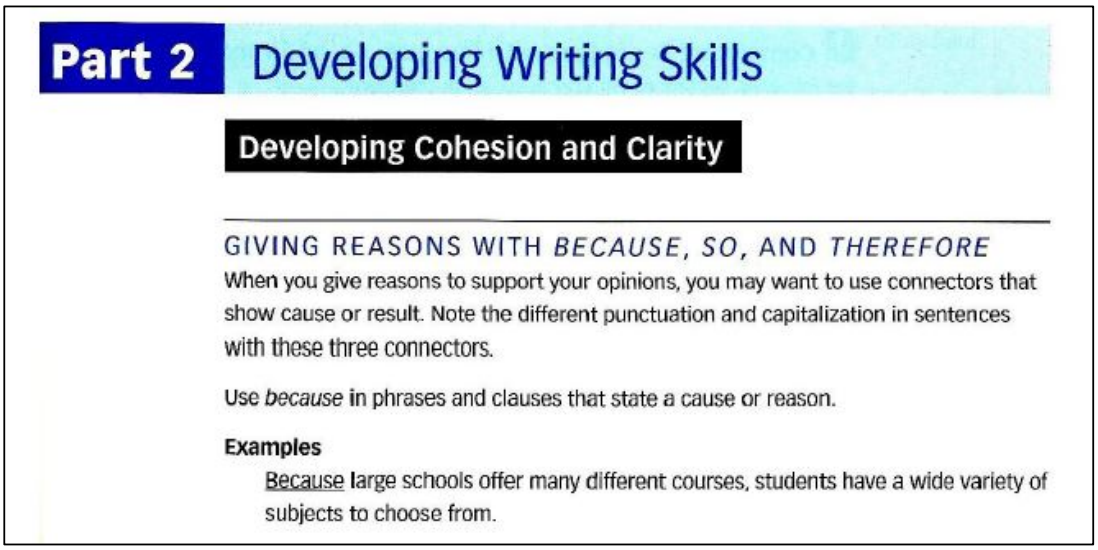
In the above excerpt, although the Arabic speaker who is learning English makes use of an anaphoric pronoun "this" in the second sentence, the meaning and the exact connection between the first and second sentences could have been made clear if the learner had used an instance of SN such as "combination".

I have also noticed the importance of signalling nouns not only in linking pieces of text together, but also in adding additional meanings to the signalled propositions/entities, meanings such as writer's stance, writer's intended organisation and so on. The following paragraph is taken from a published academic book (Hoey, 2005, p.117). The paragraph illustrates an instance of the SN *claims*.

> There are two **claims** here. The first is that one of the things we know about the words we use is that we expect them to participate in cohesion, or not. Put more formally, words (or nested combinations) may be primed to participate in cohesive chains or links, or to avoid them. The second claim is that we know what kind of cohesion to expect.

In the above quote, the use of the SN "claims" serves to help the author to signal his own stance on the following stretches of discourse which provide the specifics of the noun. The specifics of the noun are to be interpreted as *claims* that are not yet proven. The author could have used other alternative SNs either to cast his stance differently by using nouns such as "arguments" and "facts" or even to avoid expressing his stance by using nouns such as "points" and "aspects".

Furthermore, what caught my attention in writing textbooks is that there is less focus on the resource of SNs as a way to create cohesive links between segments of discourse as opposed to the thorough treatment of other cohesive resources, most notably conjunctions. Figure 1-1 illustrates this practice.



Part 2 Developing Writing Skills

Developing Cohesion and Clarity

GIVING REASONS WITH *BECAUSE, SO,* AND *THEREFORE*

When you give reasons to support your opinions, you may want to use connectors that show cause or result. Note the different punctuation and capitalization in sentences with these three connectors.

Use *because* in phrases and clauses that state a cause or reason.

Examples

Because large schools offer many different courses, students have a wide variety of subjects to choose from.

Figure 1-1 *An extract taken from a writing textbook (Pavlik and Segal, 2007, p.9)*

In Figure 1-1, the targeted learners of English (L1 Arabic) are taught how to give reasons. They are taught how to do so by using the resource of conjunctions. What is interesting here is the absence of SNs as another possible resource by which causal meanings can be signalled in texts.

The other reason is that SNs play a pivotal role in the organisation of discourse (e.g. Hoey, 1983; Flowerdew, 2003, Flowerdew and Forest, 2015). SNs are one of the lexical signals by which writers/ speakers are able to organise a discourse and readers/listeners are able to construe that discourse (Hoey, 1983, p.63). In addition to the primary importance of SNs as noted above, SNs, as pointed out by Flowerdew (2003, p.329), are "prevalent in academic discourse" (see also Flowerdew, 2006; Flowerdew and Forest, 2015). Thus, the control over SNs can be seen as "an important dimension in the development of academic literacy" (Flowerdew, 2006, p.346).

Another reason is that although there has been a fairly substantial amount of work in the literature on Arabic speaking learners of English and on cohesion (e.g. Khalil, 1989; Al-Jarf, 2001; Mohamed-Sayidina, 2010; Alharbi, 2011), no research up to now has investigated the use of SNs in this group of learners.

Based on the aforementioned reasons, focusing on SNs in the writing of Arabic-speaking learners of English as a topic for investigation is strongly justified. This brings us to the aims and research questions of the current thesis.

1.4. The aims and the research questions of my study

The overarching aims of the present study are to explore how the resource of SNs are used in written texts produced by speakers of Arabic who are learning English and in what ways the uses of SNs by this group of learners either differ from or resemble those found to occur in comparable written texts by native speakers of English.

In accordance with these aims, the main research questions are as follows:

1. What is the frequency of SNs in the learner and native speaker corpora, and how the two corpora compare?
2. What is the frequency of SNs in the learner corpus across the proficiency levels of the learners?
3. How are the three most frequent and shared SNs in the learner and native speaker corpora used in relation to the reference categories these nouns may fall into, the grammatical forms of specification and what meanings of these nouns can be identified? How do the two corpora compare with each other with these regards?
4. What are the patterns that can be associated with SNs signalling causality and those signalling stance? How comparable are the two corpora in their uses of SNs to signal those two discourse functions?

1.5. The corpora used and the methodology of the study

The present study is based on two corpora. The first is a learner English corpus comprising texts produced by L1 Arabic speakers. The corpus includes writing at several levels of proficiency. The second is a native speaker of English reference corpus which is a sub-corpus of the ICNALE corpus. More detailed descriptions of the two corpora are found in Section 3.2 of the methodology chapter.

The methodology of the present study is largely comparative and corpus-based. The study employs a combination of both quantitative and qualitative methods to gain a wider understanding of how SNs are used in the learner and native speaker corpora.

1.6. Outline of the thesis

This introductory chapter is followed by seven chapters. The following chapter (Chapter 2) locates the present thesis in its context. The aims of the chapter are (1) to address the broader context that my thesis fits into, (2) to review the relevant studies that investigated the

phenomenon of SNs and related concepts and (3) to provide my own definition of SNs. Chapter 3 gives account of the methodology of the present thesis, describing the corpora used, addressing steps of identification of SNs and the criteria used in the identification process. The following four chapters present the findings of the thesis. Chapter 4 gives account for the results relating to overall frequencies of SNs in the learner and native speaker corpora. Chapter 5 focuses on three most frequent and shared SNs in the two corpora with the aim to examine the uses and the functions of these nouns in context. Chapter 6 presents the findings of SNs with an inherent stance function in the two corpora, illustrating patterns associated with these nouns as well as findings relating to how these nouns are used in the two corpora. Chapter 7 provides the results concerning SNs signalling causality in discourse, identifying patterns associated with these nouns and also illustrating how the two corpora compare in the use of these nouns. The thesis ends with Chapter 8. The aims of this chapter are to summarise and discuss the main findings of the thesis, to suggest some pedagogic implications and to present concluding remarks.

Finally, some practical points should be noted before proceeding to the following chapters of the current thesis. Examples from the learner and native speaker corpora are numbered consecutively starting with (1) in each chapter. All examples from the two corpora are presented as they occur with no modification apart from marking the SN in question in **boldface** and underlining its specification. All the original spelling and grammar are retained. Tables and Figures are also numbered consecutively starting with the number of the chapter in which they occur: 2-1, 2-2, and so on.

CHAPTER 2: LITERATURE REVIEW

2.1. Overview

This chapter aims to contextualise the present study within its broader and specific context in the literature. This is done in two parts. The first part of the chapter begins with delineating the broader picture that the thesis fits into. This is presented in Sections 2.2 and 2.3. The second part provides the specific context of the present study by reviewing the relevant studies that have investigated the phenomenon of SNs and related concepts. This is found in Sections 2.4, 2.5, and 2.6. The chapter concludes with my definition of SNs, which is presented in Section 2.7.

2.2. Sinclair's text organisation model and SNs

As mentioned in Chapter 1, one of the functions of SNs is to organise written texts. This can be done by means of encapsulation and prospection. When SNs serve these two mechanisms, the realisation of these nouns is either in the preceding or the following sentence(s). In this sense, the specification of SNs in text is tied in with the notion of prospection and encapsulation proposed in Sinclair's (1993) model of written text organisation. In his model, Sinclair (1993) provides a useful perspective on organisation or what can be called discourse cohesion of written texts, which takes into consideration the relations that exist between sentences in text.

He argues that cohesion can be created by acts of reference, which he calls encapsulation and prospection. Encapsulation is best defined as "the means of recovering what has been previously said in order to add what is new" (Rego, 2001, p.81). Prospection, on the other hand, "… occurs where the phrasing of a sentence leads the addressees to expect something specific in the next sentence" (Sinclair, 1993, p.12). Examples (1) and (2) below, found in the

native speaker corpus used in the current study, illustrate the encapsulation and prospection function with the SNs *argument* and *reasons*. The SN *argument* in example (1) is specified in the underlined stretch of text in the previous sentence that it encapsulates whereas the specification of the SN *reasons* in example (2) is found in the following sentences that the SN prospects.

(1): Some people argue that having jobs would take up too much of the students' time and therefore, impact negatively on the students' studies. As far as I can see, this **argument** is not really valid.

(2): I have three **reasons**. First, all non smokers are paying customers and have the right to enjoy their meal in a smoke free environment.

Unlike Halliday and Hasan's (1976) view on lexical cohesion which deals with relations established between two lexical items in text where one item refers to another (see below), Sinclair's approach extends this notion by investigating relations that exist between sentences by means of encapsulation and prospection. Example (3) is presented below and is followed by a brief comment to illustrate the difference between the two approaches to cohesion.

(3): I believe that smoking should not be banned in restaurants in Japan for the following three **reasons**. Firstly, smoking is not illegal in society, so what right does an individual establishment have to rewrite the rules that we all live by and accept?

One of the instances in which one item establishes a relation with another in example (3) is found in the use of the lexical item *smoking*. *Smoking* is repeated in the second sentence, thus

creating cohesion in Halliday's and Hasan cohesion model. Sinclair (1993) points out that this type of cohesion is point-to-point cohesion because the noun phrase *smoking* refers to the noun phrase *smoking*. Sinclair's cohesion model is rather different. Sinclair would refer to the single noun phrase *reasons*, referring to a group of clauses or stretch of language, as an instance of point to a stretch of language cohesion. Sinclair's model thus identifies relations between sentences that are not only point to point but also point to a stretch of language through means of encapsulation and prospection.

With regard to SNs, they often fit into Sinclair's cohesion model of prospection and encapsulation since this category of nouns can prospect what comes next in text and encapsulate what has gone before in text. Therefore, the proper use of SNs in a text is fundamental in linking sentences in a way that makes that text interpretable as a coherent stretch of discourse. Thus, it can be said that SNs add to cohesion and continuity in text (Flowerdew and Forest, 2015).

Having shown how the concepts of encapsulation and prospection are related to the concept of SNs, it might also be worth pointing out, however, that SNs and prospection/encapsulation are not the same thing but in a sense are the obverse of one another. Sinclair (1993) identifies the role of the noun in encapsulating (or prospecting) a stretch of language. Flowerdew and Forest (2015) identify the role of a stretch of language in specifying a noun. However, it is also important to note that Sinclair's concepts of encapsulation and prospection include language features other than nouns as well. For example, according to Sinclair (1993), words like 'however', 'therefore' and 'moreover' encapsulate the preceding sentence.

2.3. Learner corpus research: issues and contributions to understanding learner language

Learner corpora are defined as "electronic collections of spoken and written texts produced by foreign or second language learners" (Granger, 2004, p.124). Hunston (2002) observes that this type of corpus provides a wealth of information about how learners differ from each other and from language produced by native speakers drawn from a comparable corpus. It is this practice of comparison which, according to Granger (1998, p.12, 2007, p.175), "lies at the heart of CLC-based studies"[2]. Granger (1998a, 2003, 2007) calls this comparative approach: *Contrastive Interlanguage Analysis* (CIA). This approach is to be differentiated from the traditional *Contrastive Analysis* (CA) (e.g. Lado, 1957) which is an analytical approach based on the premise that L2 productions (errors) are mainly predicted by L1 transfer and thus focuses primarily on the comparison of different language systems (e.g. Arabic and English) in an effort to predict areas of difficulties that learners of these languages would have. If Arabic and English, for example, were looked at from the perspective of traditional CA, it would be said that English has an indefinite article system while Arabic does not. So, speakers of Arabic who are learning English would have difficulties acquiring the indefinite article system due to the absence of the indefinite article system in Arabic. The traditional CA approach is widely criticised for its sole emphasis on L1 transfer in L2 acquisition and also for its negligence of learner interlanguage (Péry-Woodley, 1990).

Drawing on Péry-Woodley's (1990) concern with regard to learner interlanguage, Granger's (1998a, 2003, 2007) Contrastive Interlanguage Analysis is not concerned with comparing and contrasting different language systems but rather it is primarily about the comparison of learner interlanguage with a comparable native speaker language or with other learner

[2] CLC is an abbreviation for computer learner corpus.

interlanguages. To illustrate this further, Granger's approach would look at the output from learners and would compare that with the output from native speakers or indeed with the output from other learners and then would draw conclusions from the output rather than what is known about learners' L1s structures. In other words, the output from the learners is treated as being the interlanguage of the learners, which is independent from the learners' L1s and L2s. As far as the investigation of learner interlanguage is concerned, Granger's approach has its advantages. One of the advantages is that the focus is on how learners use the target language rather than the difference between the learners' target language and their L1. Hence, learners are viewed as users of a version of the target language rather than users of an (other) L1. For example, speakers of English who are learning French are viewed as users of a version of French rather than users of English. Another useful point about this approach is that it allows for two types of comparison. As noted earlier, one is concerned with the comparison of learner interlanguage with native speaker language while the other is concerned with the comparison of learner interlanguage with other learner interlanguage. The present study fits into the former approach of comparison. The aim of this approach, according to Granger (1998a, p.13; 2007, p.175), is to uncover non-native like features of language produced by learners by identifying 'overuse' and 'underuse' of lexical items. There are two controversial elements in the approach, which are worth mentioning. One is related to the terminology used in CIA studies. The other is concerned with the notion of nativeness. These two points are discussed below.

'Overuse' and 'underuse' are widely used in these studies to highlight the difference in frequencies between learner interlanguage and native speaker language. These terms evoke negative connotations. When a particular lexical item is said to have been 'overused' by learners, the implication is that the learners in some instances should not have used that

lexical item. The same reasoning applies when a particular lexical item is said to have been ‘underused’ by learners. It has been warned in the CIA literature, though, that these terms are neutral (e.g. Gilquin and Paquot, 2008; Lee and Chen, 2009; Granger, 2015). Gilquin and Paquot, (2008, p.38), for example, warn that “the terms ‘overuse’ and ‘underuse’ are descriptive, not prescriptive, terms”. According to the authors, these terms should be understood as a quantitative indication of difference but not as a qualitative indication of the quality of learner language under study. Likewise, Lee and Chen (2009, p.153) note that these terms provide a statistical description and emphasise that “there are no prior assumptions that all such overused items necessarily represent bad writing practices. They merely point to differences which merit further, qualitative investigation”. Although Gilquin and Paquot, (2008), Lee and Chen (2009) and Granger (2015) go to great lengths to disconnect themselves from the notion of judging purely on the basis of overuse and underuse, the fact remains that it is quite difficult to get away from that negative interpretation when using these terms. In view of what actually happens in practice, one might argue that the matter goes beyond the issue of terminology. In most of CIA studies, researchers talk about ‘overuse’ and ‘underuse’ in the learner corpus and they very rarely talk about ‘overuse’ and ‘underuse’ in the native speaker corpus. This suggests that, in practice, the ‘overuse’ and ‘underuse’ terms are used as judgmental terms. Aston (2011, p.11) acknowledges this common practice and states that “many analyses of learner corpora have treated all quantitative differences from reference group behaviour as undesirable—as things that learners should learn not to do, rather than as things that may help them communicate and learn more effectively”. To avoid sounding too judgmental in reporting quantitative differences, the terms ‘overuse’ and ‘underuse’ will be avoided in the present study and replaced by terms such as ‘over-representation’ and ‘under-representation’. Although it might be argued that these alternatives still sound judgmental and that they only detach judgment from the user and

make it a fact about the corpus, these alternatives are less negative and are thus used in the present study to report quantitative disparities between the learner corpus and the native speaker corpus. This brings us to the second area of controversy: the reference to native speaker norms in CIA studies.

The use of a native speaker norm in the first place as the benchmark against which a learner language is judged is not unproblematic. In fact, there is a wide range of objections to the use of such a norm. For example, many studies have pointed out that the labelling of speakers into native or non-native is often based on non-linguistic factors (e.g. Rampton, 1990; Holliday, 2009). Rampton (1990) argues for the concept of expert speaker to replace that of native speaker. Rampton's argument is a valid one as it brings the attention to the speaker's linguistic ability rather than their life-history. In other words, the attention is shifted "from 'who you are' to 'what you know'" (Rampton, 1990, p.99). In addition, many other researchers in the field of English as a Lingua Franca (ELF) have questioned the relevance of the use of native speaker norms (e.g. Jenkins, 2006 a, 2006 b, 2011; Seidlhofer, 2001, 2010; Mauranen, 2012). These ELF researchers and many others argue that English has become a global language that is predominantly used in communication between non-native speakers of English of different L1 backgrounds as opposed to communication between native speakers of English and non-native speakers of English. Thus, ELF scholars argue that it is irrelevant to use native speakers' norms as a benchmark against which the acceptability of a learner language is determined. It is suggested that ELF should be treated as "a mode of communication in its own right" (Seidlhofer, 2010, p.158). Therefore, instead of comparing language produced by non-native learners of English with that of native speakers, it makes better sense to compare that language with "performances by successful users of English as a Lingua Franca" (Aston, 2011, p.11). This argument implies that there are proficiency levels

of ELF speakers, a point which in turn raises two interrelated questions. The first is that how ELF competence in English assessed in the first place is assessed. The second question is concerned with the documentation of the range of competence levels of speakers of ELF: are these proficiency levels well delineated? The answer to the first question is that competence in any given language is usually judged based on a well-defined variety of that language. In CIA studies, for example, "all the studies that compare learners of different proficiency levels are in fact based on an underlying L1 norm [English as a native language], as proficiency is usually assessed with an L1 target in mind" (Granger, 2009, pp.18-19). In the case of ELF, however, ELF as a well-defined variety of English does not exist. Therefore, it would be difficult to assess competency levels of ELF users. The answer to the second question is rightly observed in Ferguson (2009, p.127): not enough information is given to delineate the proficiency levels of ELF "to define more precisely the language of the expert ELF user, or, indeed, to provide a means of distinguishing between ELF acceptable features and others not so acceptable". Given the responses to the two questions raised above, it can be argued here that comparing language produced by non-native speakers of English with a native speaker norm, in fact, is needed and is thus followed in the present study.

There are, of course, other reasons why this position is taken in the present study. One of these reasons is pragmatic. Mauranen's (2012) distinction between a learner of English and a user of ELF is highly relevant. Mauranen (2012) points out that the difference between the two roles is a matter of context as much as a matter of identity. She goes on to say the same person can be an ELF user in one context (e.g. outside the classroom) and a learner of English in another (inside the classroom). The learner corpus used in the present study represents people who are in the context of classroom and learning English. In that context, native speaker English is their target. Another reason is that it is widely observed in the

literature that learners from a wide range of contexts show a preference towards conforming to a native speaker norm (e.g. Timms, 2002, Moore and Bounchan, 2010). This attitude towards a native speaker norm is also shared by English language teachers, both native and non native (e.g. Jenkin, 2007; Young and Walsh (2010); Tajeddin, Alemi, and Pashmforoosh, 2018).

I have explained the rationale behind the use of a native speaker corpus as a reference corpus in the present study. Yet, what remains to be made mention of is the obvious issue that may arise from such use: how comparable is the native speaker corpus used in the present study to the learner corpus? This issue will be discussed in the methodology chapter in Section 3.2.

Corpus-based comparisons between learners and native speakers have been going on since 1998 (e.g. Granger's (1998c) edited book) and continue to this day (e.g. Wang, 2017). The interest was initially on L2 learners of English, however, the focus has been expanded to include L2 learners of other languages. The "Learner corpora around the world" website[3] lists 166 corpus projects, the majority of these corpora are targeted at L2 English while the rest focus on different L2s. What the website shows is that the work on learner corpora has grown over the past twenty years in terms of the availability of learner corpora.

In Granger's edited book (1998c), many of the studies show how comparing learner language with native speaker language may yield interesting generalisations to be made about learner language. For instance, Altenberg and Tapper (1998) find that Swedish learners of English make less use of conjuncts such as *however*, *though* and *yet* while these learners make

[3] https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpora-around-the-world.html#Learner_corpus-based_datasets (accessed 4 March, 2018)

extensive use of other conjuncts such as *furthermore, still* and *of course*. Comparing seven groups of European learners from the International Corpus of Learner English (ICLE) with the LOCNESS native speaker corpus, Ringbom (1998) finds that learners are different from native speakers in terms of the style of the discourse in that they overuse lexical items that are vague such as *thing* and *people*.

As mentioned previously, the comparison of learner language with native speaker is still used in recent scholarship. Wang (2017), for example, compares the Chinese Learner English Corpus (CLEC) with the LOCNESS native speaker corpus. The aim of her study is to compare the two corpora in terms of the use of intensifiers (e.g. *very*, *absolutely* and *so*). The findings of the study show that the Chinese learners use more tokens of intensifiers than native speakers do but they use far fewer types. The study also shows that the learners are limited in their repertoire of intensifiers' collocational patterns that are acceptable in English, a point that is found to characterise the Chinese learners' writing as "awkward learner English" (Wang, 2017, p.124).

The findings in Altenberg and Tapper (1998) and Ringbom (1998) and Wang (2017) have uncovered some aspects of learner language and thus may be of pedagogic use if taken into consideration when teaching. Hunston (2002,p.208) notes that findings obtained from corpus-based comparison do not often find their way into pedagogy and suggests a number of ways in which results obtained from L1 and L2 comparison may be translated into pedagogy. Although Granger (2009) concurs with Hunston's point that there is often weak link between the results obtained from the learner corpus research and language teaching, she goes further to say that this link should not be always necessary since what determines the translatability of learner language features uncovered by corpus research into classroom action are factors

such as “learner needs, teaching objectives and teachability” (p.22). The usefulness of learner corpus research to inform pedagogy, nonetheless, cannot be denied. Granger (2009, p.27) herself acknowledges this:

> This research [referring to learner corpus research] has resulted in learner-corpus informed writing aids that show typical examples of native and learner writing and provide a rich semantic, syntactic, phraseological and stylistic description of the ways particular functions and concepts are lexicalized in academic English.

The present study is set within the field of learner corpus research. The study investigates the use of SNs in Arab learners who are learning English, a population of learners that is extremely underexplored in the field of corpus research, and thus aims to contribute to the field by uncovering features of language produced by this group of learners. The findings of the study may be of pedagogic value to teachers of Arabic-speaking learners of English.

Having put the present study within its broader context, I will now move on to the second part of this chapter, which reviews previous work on signalling nouns and related concepts. This will be presented below in the following three sections (Sections 2.4, 2.5, and 2.6).

2.4. SNs in the literature

A number of studies have addressed SNs and related concepts. In this section, I will discuss first the early influential precursors to SNs in Section 2.4.1, then the studies that have developed the concept of SNs further in Section 2.4.2 and finally the seminal work of Flowerdew and Forest (2015) in Section 2.4.3.

2.4.1. Early influential precursors to the phenomenon of SNs

Halliday and Hasan (1976) and Winter (1977) were the first to draw attention to the potential role played by lexical items in discourse. Halliday and Hasan's (1976) study is addressed in Section 2.5. Winter's contribution to the phenomenon of SNs is reviewed below.

The concept of signalling in discourse can be traced to the pioneering work by Winter (1977). In this work, Winter classifies the vocabulary of English whose main function is to signal the semantic relationships that might exist between clauses into three categories: Vocabulary 1: subordinating items (e.g. *whereas, while*); Vocabulary 2: those grammatical set of words whose function is to link independent sentences (*however*, *so*); Vocabulary 3: those lexical items such as nouns (e.g. *contrast*), verbs (e.g. *compare*) and adjectives (e.g. *different*) serving to signal relations between clauses in discourse. Vocabulary 1 and 2 are dealt with in any standard grammar under the headings of 'subordinators', 'coordinators' and 'sentence adverbs'. Winter's innovation was to note that words playing other roles in the clause (e.g. subject, verb etc.) could perform the same clause connection function.

The third set of vocabulary (Vocabulary 3) items proposed by Winter is interesting in that although they are open class (lexical) words, they resemble close class (grammatical) words in their linking function. As postulated by Winter (1977), they represent a closed system and as a result there are a limited number of Vocabulary 3 items (only 108 lexical items). Although Winter's Vocabulary 3 includes nouns, verbs and adjectives that are capable of signalling relations between clauses, nouns in particular have received a considerable amount of attention in the literature. It could be said here that Winter's concept of vocabulary 3 nouns has inspired many researchers to investigate the phenomenon of SNs and to develop Winter's ideas on the subject. In fact, the research projects conducted on SNs and related phenomena

have greatly expanded the number of nouns with linking function in discourse and also have shown that this class of noun is open-ended, a point that goes against Winter's argument that vocabulary 3 are a closed set (e.g. Francis, 1986, 1994; Hunston and Francis, 1999; Schmid, 2000; Flowerdew and Forest, 2015). These studies will be addressed below.

2.4.2. Further development in the phenomenon of SNs

Following the early work by Halliday and Hasan (1976) and Winter (1977), the phenomenon of SNs has been studied in various ways, which has resulted in different definitions and distinct terminologies assigned to these nouns (e.g. Francis, 1986, 1994; Hunston and Francis, 1999; Schmid, 2000; Flowerdew, 2003; Flowerdew and Forest, 2015). Table 2-1 shows the labels attached to nouns in these studies along with the description assigned to these terminologies.

Table 2-1 *Previous studies on the phenomenon of signalling nouns*

Study	Label	Brief description
Francis (1986)	A- nouns	A- nouns are nouns which functions to signpost "to a stretch of discourse preceding it in terms of how the writer chooses to label or interpret for the purposes of his/her argument…[and these nouns] must be presented as the given information" (p.4). This type of noun must refer to a clause or clauses, not a noun phrase.
Francis (1994)	Advance/ retrospective nouns	Retrospective nouns are those functioning to anaphorically recover a previously mentioned stretch of language whereas Advance nouns are those serving to cataphorically refer to a stretch of language. Similiarly to A- nouns, the referent of retrospective and advance nouns must be a stretch of language larger than a noun phrase.

Table 2-1 *(Continued)*

Hunston and Francis (1999)	Shell nouns	Shell nouns refer to those nouns identifiable (1): "on the basis of their behaviour" and (2): on the their need for lexical specification in the surrounding environment (p.185).
Schmid (2000)	Shell nouns	Shell nouns are abstract nouns that have the capability of serving as "a conceptual shells for complex, proposition-like pieces of information" (p.4). This type of noun is identified based on four syntactic patterns which provide "a clearly defined slot where the shell noun can be expected" (p.40).
Flowerdew (2003)	Signalling nouns	"Any abstract noun, the meaning of which can only be made specific by reference to its context" (p.2).
Flowerdew and Forest (2015)	Signalling nouns	Abstract nouns whose meaning lacks specificity when treated in isolation but by reference to their linguistic context, the meaning of these nouns is specified.

Table 2-1 shows the attempts that have been made to address the phenomenon of signalling nouns. No matter how these nouns are labelled, the underlying idea behind the use of this type of abstract nouns is broadly the same. The idea can be summed here in two main aspects. First, the meaning of these nouns is to be sought from the context where these nouns occur. Second, these nouns link clauses. The different writers stress either one or the other of these aspects. However, looking at the description of the labels assigned to these nouns indicates that there certainly exists an overlap and sometimes a difference between the labels and the concepts that these authors assign to the phenomenon of SNs.

A-nouns, for instance, as identified by Francis (1986) are limited to those nouns serving to refer back to a previously mentioned clause or several of clauses in text. Francis' later study (1994) referred to the same phenomenon using the label 'retrospective nouns'. Signalling nouns whose specification is anaphoric and across clauses as identified in Flowerdew (2003) and Flowerdew and Forest (2015) can also be said to overlap with Francis' A-nouns and retrospective nouns.

In her 1994 publication, Francis extends the identification of nouns that have a discourse metadiscursive function to include nouns that refer to a stretch of language in a cataphoric direction. She uses the label of 'Advance nouns' to refer to these nouns. What labelled as Advance noun in Francis (1994) overlaps with signalling nouns whose specification is cataphoric across clause as found in Flowerdew (2003) and Flowerdew and Forest (2015).

Notwithstanding the overlap between the labels and concepts assigned to the phenomenon of SNs as found in Francis (1994) and Flowerdew (2003) and Flowerdew and Forest (2015), there are some differences. For example, Francis' defining feature of retrospective and advance nouns is that these nouns must refer to a stretch of language larger than a noun phrase. Although Flowerdew (2003) and Flowerdew and Forest (2015) follow Francis' rule, they are more flexible in that "[they] *do not exclude nominal groups as the referent of SNs if such a nominal group exhibits clause-like semantics, in having a deverbal noun, deadjectival noun or gerund as its head* " (Flowerdew and Forest, 2015, p.28, emphasis in original). Another area where Francis (1994) greatly differs from that of Flowerdew (2003) and Flowerdew and Forest (2015) is that Francis (1994) limits its focus on nouns that are specified across clause boundary. However, this focus is broadened in Flowerdew (2003) and Flowerdew and Forest (2015) to include nouns that are realised within clause boundary.

Given the overlap and the difference between Francis' two publications (1986, 1994) and the studies by Flowerdew (2003) and Flowerdew and Forest (2015), as shown above, it is expected that the lists of nouns that these authors arrive at will also show a degree of overlap and difference, and this is in fact the case. For example, the latest list of SNs (850 SN types) that appear in Flowerdew and Forest's (2015) study includes many nouns that are found in Francis's (1994) study and many that are not. Examples of nouns found in both studies, to list a few, are *argument*, *definition*, *example* and *illustration*. Examples of nouns that are only found in Flowerdew and Forest's (2015) list are *dichotomy*, *drawback*, *expansion*, *generalisation*, *epistemology* and *axiom* while nouns such as *drivel*, *affair*, *irony*, *blunder* and *affront* are only identified in Francis' (1994) study. The difference in the SN types identified by the two authors as shown in the aforementioned examples may be due to the nature of the data from which nouns are drawn or possibly due to the difference in definition; Francis' nouns are identified to occur in a corpus drawn from journalistic prose whereas Flowerdew and Forest's nouns are identified to occur in a corpus drawn from academic texts, spoken and written.

Hunston and Francis (1999) identify a class of nouns on the basis of the complementation pattern this class is associated with. They label this class as *shell nouns*[4]. According to these authors, this type of noun requires lexical specification in the surrounding text and this specification can be found in the that-appositive clause that follows *shell nouns*. Unlike almost all other studies mentioned in the table, Hunston and Francis (1999) identify shell nouns on formal grounds rather than functional ones. Schmid (2000) takes a similar stance in his identification of shell nouns. However, he identifies four typical patterns that can be

[4] It is useful to note that the label 'shell nouns' is borrowed from Schmid (Hunston, 2018, personal communication).

associated with *shell nouns*: (1) shell noun + that-clause, (2) shell noun + be + that-clause, (3) this + shell noun and (4) that + be + shell noun (p.22). Thus, the expansion of the patterns of shell nouns by Schmid (2000) makes it possible to find nouns in Schmid's list that would not appear in Hunston and Francis's (1999). Examples of such nouns, to mention a few, are *aim*, *answer* and *achievement*.

Having explained how the studies that dealt with the phenomenon of SNs either differ or overlap with one another, it is useful to provide examples that illustrate this point. Nouns like *argument, allegation, assertion, assumption* and *analysis*, for instance, can be examples of all the labels assigned to the phenomenon of SNs exemplified in Table 2-1 above while a noun like *example* cannot appear as an example of the label of *shell noun* in Hunston and Francis's (1999) sense, as this noun cannot be followed by an appositive that-clause.

The studies mentioned in Table 2-1 can all be said to be of relevance to my current project as they all address SNs from various angles, of course with varying degrees of relevance and also with varying perspectives attached to this signalling phenomenon of abstract nouns. In terms of the perspectives of the aforementioned studies dealt with abstract nouns, it can be said here that there are two areas of emphasis placed on these nouns: one of which is structural (Hunston and Francis, 1999; Schmid, 2000) while the other is discourse oriented (Francis, 1986, 1994; Flowerdew, 2003, Flowerdew and Forest, 2015). In the structural approach to abstract nouns, the focus has been largely on the identification of nouns that function as shells of the lexical specification given in a content clause on grounds of form, whereas in the discourse oriented perspective, the emphasis has been on the functional features of abstract nouns at a discourse level, i.e. how these nouns function metadiscursively to refer to a stretch of language not only at a sentence level but also across sentences.

Each of the two aforementioned perspectives associated with the phenomenon of SNs in the literature has its own advantages and disadvantages. The advantages and disadvantages will be addressed below in the discussion of Schmid's (2000) study and Flowerdew and Forest's (2015) study (see Section 2.4.3 below). The selection of these studies in particular for further discussion is because these studies are the two most comprehensive studies to date that investigated the phenomenon of SN with the first advocating the structural approach towards the identification of SNs and the latter prioritising the discourse approach.

One of the most comprehensive studies that investigated the phenomenon of SNs (under the label of *shell nouns*) is that of Schmid (2000). The corpus used in Schmid's (2000) study is the British element of the Bank of English. At the time of the study, the corpus consisted of 225 million words drawn from a wide range of texts, written and spoken collected from a variety of sources, with the written texts skewing the corpus data. As mentioned previously, Schmid's (2000) definition of the noun class referred to as shell nouns is structural par excellence. In order to be defined as a shell noun, the noun has to occur in four syntactic patterns. These patterns along with their illustrative examples as they appear in Schmid's (2000, p.22) study are illustrated in Table 2-2 below.

Table 2-2 *Schmid's syntactic patterns used in the identification of shell nouns*

Pattern number (abbreviation)	**The pattern used to identify shell nouns**	**Examples illustrating the pattern**
One (N-cl)	Shell noun + postnominal clause	Mr Bush said Iraq's leaders had to face **the fact** that the rest of the world was against them. (BBC)
Two (N-be-cl)	Shell NP + *be* + complementing clause	**The advantage** is that there is a huge audience that can hear other things you may have to say, (PAPERS)

***Table 2-2** (Continued)*

Three (th-N)	Referring item + (premod) + shell noun	(Mr Ash was in the clearest possible terms labelling my clients as anti-semitic.) I hope it is unnecessary to say that **this accusation** is also completely unjustified, (PAPERS)
Four (th-be-N)	Referring item as subject + *be* + shell noun (phrase)	(I won the freshmen's cross-country. - Mm.) That was **a great achievement** wasn't it? (SPOKEN)

To arrive at the list of shell nouns (670 shell nouns), Schmid (2000) only uses the first two pre-specified patterns as shown in Table 2-2 above. He then uses the list of nouns identified to occur in these two patterns to search for patterns three and four. His study largely focuses on the first two patterns, as acknowledged by Schmid (2000, p.42), where he points out that nouns found to occur only in pattern three and four as shown in Table 2-2 are excluded as these patterns (three and four), according to Schmid (2000, p.42), do not feature a noun as a shell noun.

Such a structural approach towards the phenomenon of SNs is useful in two ways. First, the structural approach to identify the SN category is undoubtedly preferred when examining large corpora; studies by Hunston and Francis (1999) and Schmid's (2000) are cases in point. When dealing with large corpora, it is difficult to look at concordance lines of all individual nouns to decide whether they belong to a certain word class. Thus, it is often the case that the researchers use pre-specified patterns to retrieve all the nouns that are used in such patterns. Second, such an approach shows a degree of objectivity in the retrieval of the nouns under investigation through automated queries.

In spite of the two obvious advantages of the structural approach towards the phenomenon of SNs, there are many disadvantages. The main disadvantage is that this approach is restrictive in terms of the identification of the phenomenon of SNs. This is the case in the studies by Hunston and Francis (1999) and Schmid (2000) where the identification of SNs in the former is too restricted to nouns occurring in one syntactic pattern (N-that appositive clause) and in the latter is limited to two syntactic patterns (N-cl and N-be-cl) as illustrated in Table 2-2. Such a narrow approach towards the identification SNs does not provide a wider picture of the patterns associated with the phenomenon of SNs and also the range of SNs identified (see the advantages of the discourse approach towards SNs addressed below for comparison). In fact, the reliance on pre-specified automated patterns to investigate the phenomenon of SNs is bound to lead to the omission of other potential patterns that may be associated with this class of nouns. This drawback is acknowledged by Schmid (2000, pp.51-53). One of the problems caused by the reliance on the automated syntactic patterns, as pointed out by Schmid (2000), is that linguistic components may occur between the SN and its realisation and thus the noun is not identified by the automated method as a shell noun as in the following example: "**The next part** of the project is to go back and to identify where these products come from" (p.52) where the prepositional phrase (*of the project*) following the shell noun is held responsible for not making the computer identify the noun *part* as a shell noun. This issue of linguistic items found to occur between the noun and its realisation as shown in the aforementioned example is estimated by Schmid (2000) to be responsible for the missing of 30-40% of uses of shell nouns. Another problem associated with the structural approach is the fact that plural SNs are not included in the automated queries used to arrive at the list of shell nouns (see Hunston and Francis, 1999 and Schmid, 2000). Given all these disadvantages, the structural approach towards the identification of the phenomenon of SNs is not reliable and does not provide a wider scope of the nouns that may function as SNs.

2.4.3. Flowerdew and Forest (2015) and SNs

The most comprehensive and recent study that investigated the phenomenon of SNs is that of Flowerdew and Forest (2015). The corpus used in the investigation of SNs in this study (600,000 words) is drawn from academic texts, spoken (lectures) and written (academic journals and textbooks) collected from a wide range of sources. In dealing with the phenomenon of SNs, Flowerdew and Forest (2015, p.6) emphasise that the phenomenon is best accounted for "not only as a lexical and structural phenomenon, but also from the perspective of discourse". It is this discourse approach towards SNs that has underpinned Flowerdew and Forest's (2015) study. In fact, this approach is emphasised in the definition of SNs proposed by the authors as shown in Table 2-1 presented earlier; the definition does not associate SNs with certain structural patterns. Similar to Flowerdew and Forest (2015), Francis' (1986; 1994) studies also stress that this category of nouns is best understood as a discourse phenomenon and thus is best identified on the basis of that perspective rather than on formal grounds.

Unlike Hunston and Francis (1999) and Schmid (2000), Flowerdew and Forest (2015) do not use pre-specified automated syntactic patterns as a way to identify each instance of an SN but rather they look at all abstract nouns found in the list of all words occurring in their corpus. These nouns are then examined to determine whether they are used as SNs. This procedure is facilitated by the use of corpus software which helps in providing the concordance lines of any selected noun and the on-screen function which makes it possible to examine any selected noun within its wider textual context.

The discourse approach towards the phenomenon of SNs demands a human involvement in the identification of an instance of an SN: its function and its pattern. This point does have its

own disadvantages. One is that the analysis may involve a degree of subjectivity on the part of the analyst. The second is that the approach, as acknowledged by Flowerdew and Forest (2015), is time consuming and a laborious task. Despite these two inevitable problems, the approach lends itself to be extremely useful as far as the phenomenon of SNs is concerned as will be discussed below.

One of the benefits of the discourse approach is that it provides a broader approach towards the identification of the phenomenon of SNs. The identification of any instance of noun as an SN is not done on the basis of form but rather on functional grounds. Thus, the range of SNs identified and the functions of these nouns in discourse are not narrowed down.

Flowerdew and Forest's (2015) discourse approach, for example, has helped the authors to identify each instance of an SN in their corpus. Despite the fact that the corpus used in their study is extremely small in size compared to that of Schmid (2000), Flowerdew and Forest (2015) have been able to identify 845 SN types thanks to the discourse and methodological approach followed in the investigation of the phenomenon as discussed earlier.

My perspective on SNs is based on that of Flowerdew (2003) and Flowerdew and Forest (2015), i.e. the discourse oriented perspective (see Section 2.8 for further details). This is not to say that the structural perspective is not going to be useful in my project. Indeed, this perspective is useful because it helps in identifying the structural patterns associated with SNs under investigation. The identification of these patterns might account for the possible discourse functions or behaviour of these nouns in discourse.

Having discussed Flowerdew and Forest's (2015) study with regard to its approach to the phenomenon of SNs, I now turn to address its findings. The study shows a number of interesting results. The frequency of SNs in the social sciences is higher in comparison to the natural sciences, a difference which is postulated to be due to the frequent occurrence of technical terms in the natural sciences (Flowerdew and Forest, 2015, p.183). The use of SNs is also found to be influenced by particular genres and disciplines. For example, it is suggested that the "higher frequency of more general nouns like *thing*, *point* and *way*" in lectures is influenced by the nature of spoken mode (Flowerdew and Forest, 2015, p.184). With respect to the syntactic patterns associated with SNs, it is found that across clause SNs are used more than within clause SNs.

2.5. Lexical cohesion and SNs

Since the current study is concerned with SNs as an available lexical resource to create cohesion between stretches of discourse within text, it is useful to briefly touch on lexical cohesion and its relationship with the phenomenon of SNs. Halliday and Hasan (1976) state that lexical cohesion is created by means of vocabulary where one lexical item in a text enters into a semantic relationship with another item in the same text. Among the lexical resources they identified is what they call *general nouns*.

When discussing lexical cohesion, the two authors draw attention to an aspect of lexical cohesion (*general nouns*), which lies in a continuum between lexical cohesion and grammatical cohesion as shown in the widely quoted example below (Halliday and Hasan, 1976, p. 279):

I turned to the ascent of the peak. | The ascent / The climb / The task / The thing / It | is perfectly easy.

The above sentence shows a number of lexical cohesive resources (Halliday and Hasan, 1976, p.279): the underlined phrase (‘the thing’) is an example of the phenomenon named *general noun*. It is clear that this noun occurs on the paradigm between the grammatical pronoun *it* and the other obvious lexical items. The location of *thing* in the paradigm in the above example is a feature of SNs. Therefore, the general noun *thing* in the example can be regarded as an instance of SNs in the present study. However, not all general nouns discussed by Halliday and Hasan (1976), indeed most of them, cannot be regarded as SNs. The reason is that these authors include many concrete general nouns such as *people, man, woman, creature* (Halliday and Hasan, 1976, p.274).

SNs can be distinguished from other resources of cohesion by reference to their abstractness and position between lexical and grammatical set of words. The phenomenon of SNs can be best seen as a lexico-grammatical cohesive resource, which is capable of linking stretches of language in text.

So far, I have reviewed previous studies on SNs and related concepts. I now move on to review research on SNs in learner writing. This is addressed in the following section.

2.6. Learners and the use of SNs

It is vitally important that learners of English demonstrate their ability to produce written texts that are cohesive and coherent. One way of doing so can be facilitated through the proper use of SNs. It has been found that this type of noun is prevalent in academic discourse (Flowerdew, 2003; Flowerdew and Forest, 2015). The pivotal role of SNs in discourse makes it significant to investigate the use of SNs by learners of English and to compare it to that of native speakers of English.

The phenomenon of SNs in student writing has received less attention in the literature. Among the few studies that dealt with this type of noun are those by Flowerdew (2006, 2009, 2010) and Tåqvist (2016). These studies will be reviewed below.

Flowerdew (2006, 2009) investigates SNs in the English writing of L1 Chinese students. The corpus used in the study comprises 210 essays written by this group of learners at tertiary level. His focus is mainly on the inappropriate use of SNs. The identification of the misuses of SNs is done manually by the researcher himself with the help of a research assistant. They then provide a categorisation of SN errors: *errors in colligation*, *collocation inaccuracies*, *wrong choice of SNs*, and *deletion of SNs*.

This study, however, has suffered several limitations that the writer himself has acknowledged. One of these shortcomings is that the study has not compared the uses of signalling nouns by the learners to that of native speakers of English which could have helped

to pinpoint any discrepancy in terms of under or over representation of SNs and "to what extent their [SNs] range of (i.e. different types) might vary with that of native speakers" (Flowerdew, 2006, p.359). Another drawback is related to the methodology, which favours error analysis. This analytical approach, as pointed out by Tåqvist (2016, p.45), has been widely disfavoured in the recent "linguistic scholarship". Error analysis ignores correct use of linguistic features and places more emphasis on misuse of those features, a point that does not provide a complete picture of the learner language under investigation. A similar point has been made above in Section 2.3 in relation to Contrastive Analysis.

To rectify the limitation of his earlier study, Flowerdew (2010) incorporates a native speaker corpus to his study while he uses the same learner corpus used in the 2006 study. The aim was to compare the writing of Chinese learners with that of native speakers. The findings show that SNs are more common in the native speaker corpus. This in turn suggests that learners of English use SNs less than English native speakers do. The study also concludes that native speakers employ wider range of functions in terms of the specification of SNs. However, in my view the comparability of the two corpora is not exact, because the native speakers were not writing under timed conditions, as the learners were, and could revise their essays. This may have influenced the functions and frequencies of SNs found to occur in the native speaker corpus. It has been suggested in the literature that by allowing writers to revise their texts the overall organisational quality of their texts is improved (John, 2005).

Tåqvist (2016) resembles Flowerdew (2010) in that it compares the writing of non-native speakers of English with native speakers at academic contexts (Tåqvist refers throughout to these nouns as discourse organising nouns). However, the group of learners investigated in Tåqvist (2016) is Swedish. The academic essays written by advanced Swedish students of L2

English (200,033 words) are drawn from the Swedish sub-corpus of the International Corpus of Learner English (ICLE-SW). The aim of the study is to compare the essays produced by Swedish learners of English in terms of the use of discourse-organising nouns (DONs) with other two groups: L1 student essays (324.304 words) drawn from the Louvain Corpus of Native English Essays (LOCNESS) and essays of expert writers of L1 English (253.733 words) drawn from the learned sub-corpora of the International corpus of English (ICE).

To identify potential DONs, Tåqvist (2016) uses two methodological steps that are worth commenting on. First, she follows the structural approach to the identification of DONs. Following Schmid (2000), two pre-specified patterns: (1) "*the* + [noun] + *that*" and (2) "*the* + [noun] + [be] + *that*" are used (Tåqvist, 2016, p.65). The academic sub-corpus of the corpus of contemporary American English (COCA) is chosen to search for the patterns with a cut-off point of 500 most frequent instances of each pattern. Second, 20 essays from the ICLE-SW and 20 from the LOCNESS are manually examined. Interestingly, Tåqvist is able to identify 74 DONs types in the second step and 47 in the first step (the two lists of nouns identified in the two steps overlap and differ at the same time). In light of my previous discussion of the limitations of the structural approach towards the identification of the phenomenon of SNs and related concepts (see Section 2.4.2), identifying fewer DONs types in the first step is not surprising. As noted earlier, the structural approach is bound to retrieve hits in which the word tagged as noun does not belong to the class being looked at. This happens to be the case in the identification process of DONs in Tåqvist's study. For example, out of the 500 nouns identified to occur in the pattern "*the* + [noun] + *that*", as pointed out by Tåqvist, more than half (276 instances) is found to feature nouns with no discourse organising function (e.g. *the cab that*) (Tåqvist, 2016, p. 65).

Drawing DONs from the three corpora by using two different methods is considered a strength in Tåqvist's study. However, her method does prioritise L1 rather than L2 usage in identifying DONs and therefore may not capture the full range of DONs used by L2 writers.

The study shows that the L2 writers use more DONs tokens than the other two groups of writers. The study also shows that the L2 students stand out with regard to the fewer number of DONs types used. Furthermore, the L2 writers are found to use more vague DONs (e.g. *thing*) and more DONs marking involvement (e.g. *question*). The study concludes that the use of DONs in the writing by the L2 and the L1 students (though to a lesser degree in L1 students), compared with the expert writing, could in part be perceived as falling short of the formal features of academic written discourse.

2.7. Towards a definition of SNs

Flowerdew and Forest (2015, p.1) define SNs as "abstract nouns which are non-specific in their meaning when considered in isolation and which are made specific in their meaning by reference to their linguistic context" both across clause and within clause. There are two important points in this definition. Firstly, the noun must be abstract. Secondly, when observed in discourse, its meaning cannot be determined from the noun alone but the reader needs to look elsewhere in the linguistic context to find out what the noun means. This definition is used in this paper for two main reasons. The first reason is that the definition adopts a broader perspective on the phenomenon of SNs than others in the literature (e.g. Francis, 1986, 994; Hunston and Francis, 1999; Schmid, 2000). The second reason is that the definition makes it clear that the qualifying criterion for an instance of abstract noun to be regarded as a signalling noun is functional. This means that SNs must be lexically specified in their surrounding environment. The lexical specification of any signalling noun (SN) refers

to the stretch of language that gives the specific meaning of that signalling noun either within the same clause or across clause boundary (Flowerdew and Forest, 2015).

Flowerdew and Forest's definition of SNs makes it clear that what constitutes an instance of SN is that the noun has to be abstract and has to be specified in text. However, what constitutes an abstract noun in the first place might be problematic. On the one hand, there are nouns that can be distinguished with ease as abstract nouns. Examples of these nouns include *problem*, *reason*, and *attitude*. On the other hand, there also exists a number of nouns which are in the borderline between abstract and concrete nouns. Examples of these nouns are *holiday*, *vacation* and *trip*. Nouns such as these are normally excluded from the literature on SNs. The recent list of SNs in Flowerdew and Forest's (2015) study does not include these nouns. These nouns may not have occurred in their corpus and thus are not listed accordingly. Their corpus is drawn from academic texts and these nouns are less likely to occur in such contexts. It is difficult to know as to whether they would include these nouns if they had actually occurred in their corpus.

To be more inclusive of all nouns that function to signal a stretch of discourse in text, the definition of SNs should be slightly modified. Thus, SNs are best defined as nouns that are abstract or on the borderline between abstract and concrete nouns that are non-specific in their meaning when considered in isolation, but when treated in their linguistic context, their meaning is supplied. This definition makes it possible to divide SNs into two categories: core SNs (abstract nouns) and marginal SNs (ordinary nouns that are not widely recognised as abstract nouns).

As shown in the discussion of SNs and related concepts, there is no scholarly agreement on what constitutes the stretch of discourse that supplies the meaning of an SN. Francis (1994), for instance, suggests that the *lexicalisation* (her terminology) must be a clause or clauses but not noun phrases. Schmid (2000) also suggests that the *shell content* (his terminology) is realised by clauses and not by noun phrases. However, other scholars do not restrict the realisation of the SN to clauses only but they also allow for noun phrases. This broader view is found in Flowerdew (2003) and Flowerdew and Forest (2015), however with the caveat, as noted earlier, that the noun phrase providing the specification for an SN has to be nominalisation (a noun derived from verb, adjective or gerund). The current study takes even a broader view than that of Flowerdew (2003) and Flowerdew and Forest (2015) by allowing for noun phrases even if the nouns do not exhibit clause like semantics achieved by nominalisation. Taking such a broader view is explained in the following reason. Allowing for a broader definition of what constitutes the specification of an SN provides a more inclusive examination of the phenomenon of SN usage in student and native speaker writing (see Chapter 5 for further discussion and examples). In this study, I follow Flowerdew and Forest (2015) in using the term 'specification' (or 'specifics') to refer to the stretch of language that provide the specific meaning of an SN. It should be noted that Flowerdew and Forest (2015) also use the term 'realisation' interchangeably with 'specification'. The reason why I use 'specification' instead of 'realisation' is that the latter is widely used in the literature (e.g. Systemic Functional Linguistics (SFL)) as a concept to explain the interrelation between form and meaning in text.

SNs are important in writing as they have the potential not only to create cohesion by means of encapsulation but also to serve as a discoursal resource capable of organizing written discourse by means of prospection (see Section 2.2 in this chapter for further discussion of

these two mechanisms: encapsulation and prospection). Therefore, it is useful to study this type of noun to see what and how SNs are used in essays by both native speakers of English and Arabic speakers who are learning English. This is one of the main objectives of the current study. As noted in Section 1.3 presented earlier, there has been no study so far that has investigated the use of SNs in writing by Arabic speakers who are learning English. Therefore, investigating SNs among this group of learners will partially bridge this gap in knowledge. The study will also be particularly useful for many reasons. Firstly, it will tell us how the phenomenon of SNs is used by this type of learners and whether the use of SNs by these learners is different from that of native speakers of English. Secondly, the study will provide useful contrast with the very limited studies that exist in the literature (see Section 2.6). These studies indicated that Chinese learners of English and also Swedish learners of English had issues in using SNs in their writing. Therefore, the current study will show whether the issues associated with the use of SNs found in the essays written by these two groups of learners can be found in another group of English learners who come from a distinct language background, i.e. Arabic speakers who are learning English. This will make it possible to determine whether the problems experienced by these groups of learners are universal. Another aim of the project arises from the fact that the work on the phenomenon of SNs is still evolving and has not reached a stage where one can say it is a well established field of study. This fact is truly explained by the on-going interest placed on this type of nouns in the literature as shown in Table 2-1. Therefore, this study will also attempt to identify new points of interest with regard to this class of nouns such as new patterns associated with SNs when serving a particular discourse function.

The following chapter provides a description of the learner and native speaker corpora used in the thesis. It also outlines the steps as well as the criteria followed in the identification and quantification of SNs.

CHAPTER 3: METHODOLOGY

3.1. Overview

This chapter is concerned with the methodology of the current study. The chapter is divided into two main parts. The first part provides a description of the two corpora used in the study and a discussion of how comparable these corpora are. This is presented in Section 3.2. The second part addresses the procedures followed in dealing with the phenomenon of SNs in the present study. This part is divided into three sections. The first section (3.3) presents the steps undertaken in the identification process of SNs in the two corpora. The second section (3.4) addresses the criteria that are used to determine what constitutes an SN in this study. The section also discusses the methodological issues that arise from these criteria and shows how these issues are dealt with in the study. The third section (3.5) provides a discussion of the repeated instances of SNs. The chapter ends with a comment on the SNs to be investigated in the study presented in Section 3.6, and a brief summary of the chapter in Section 3.7.

3.2. Description of the corpora used in the study

The learner corpus used in the study is drawn from the BUiD Arab learner corpus (BALC). This corpus has been compiled by the British University in Dubai (BUiD). The corpus is available online among the list of learner corpora around the world presented by Sylviane Granger and her colleagues at the University of Louvain, Belgium. However, access to the BALC data was kindly given by Nicholas Groom from the University of Birmingham who directed the BALC project in a joint collaboration with Mick Randall from the British University in Dubai.

The BALC corpus contains written essays collected from three different sources: secondary schools (146 essays), first year university students (36 essays) and essays by students sitting the Common Educational Proficiency Assessment test (CEPA)[5] (1676 essays). Only CEPA essays are used in the study for three main reasons. Firstly, CEPA essays constitute the largest number of essays (1676 essays) found in the BALC corpus, comprising 253,211 word tokens. Secondly and more importantly, CEPA essays are assigned bands from one to six in ascending order of language proficiency. This will make it possible to compare the writing of students at different levels of proficiency in terms of the use of SNs, which is one of the aims of the present study. Finally, these essays were written under exam conditions and were timed. By having these two conditions met in the CEPA essays, it can be guaranteed that the essays being investigated are written by students alone. Hence, I can be confident that a picture of the students' actual writing ability can be obtained.

Although there are, as I have mentioned earlier, six proficiency levels of the CEPA essays (CEPA1 to CEPA6), essays of CEPA1 and CEPA2 are excluded from the present study. The reason why these essays are not used is due to the fact that these essays are poorly written and lack structure. To illustrate this, the following are examples of CEPA1 and CEPA2 along with criteria based on which these essays are assessed as CEPA1 and CEPA2 respectively.

(1): `I go holiday in my family.in the see she was swimming and go ham grand father and grand mother.my brother go swimm and took my Help, Help. my father ran speed to Help brother.` **(A CEPA1 student's whole text)**

[5] CEPA is a placement test for admission to higher education institutions in the United Arab Emirates. Emirati students in their final year of secondary schooling are required to sit the test. More detailed information about the CEPA test is given in Appendix A.

Criteria for CEPA1:

- Message is frequently unclear.
- Essentially unable to make sentences or multi-word messages.
- Vocabulary is extremely limited.
- Spelling errors are frequent, even in common words, except when copied.
- May confuse upper and lower case letters.
- Punctuation is mainly missing or inaccurate.
- Text may be so short that it is difficult to assess meaningfully. (UAE Ministry of Higher Education and Scientific Research, 2014)

(2): I liked holiday so much. A one day in the holiday I go to the baech with my family. I see people, canten and tree. I doing swimming, run, play volyball and Footbal. This holiday is a very wonderful becouse I see anew friend and bired. My holiday is a wonderfull. **(A CEPA2 student's whole text)**

Criteria for CEPA2:

- Can convey only the simplest ideas.
- Attempts to produce short sentences and phrases independently, but with little control of sentence structure.
- Vocabulary is limited to common words.
- Can spell a few common words accurately.
- Some evidence of punctuation, but usually inaccurate. (UAE Ministry of Higher Education and Scientific Research, 2014)

It can be seen from the criteria for CEPA1 and CEPA2 that the essays comprising these two proficiency levels are very low. Another point is that in the CEPA1 and CEPA 2 criteria there is no mention of anything beyond the level of sentence. However, this is not the case in the criteria for CEPA3, CEPA4, CEPA5 and CEPA6 (see Appendix A) where concepts such as cohesion and text organisation are included in the assessment criteria. Thus, it is decided that CEPA1 and CEPA2 are not suitable for the investigation of the resource of SN. The learner corpus used in the present study, therefore, consists of CEPA3, CEPA4, CEPA5 and CEPA6, comprising a total of 211,178 words (1099 essays).

The native speaker corpus used in the study consists of written essays by native speakers of English consisting of 90,613 word tokens. This corpus (400 essays) is a subset of a larger corpus of the International Corpus Network of Asian Learners of English (ICNALE), developed by Ishikawa from the University of Kobe, Japan (see Ishikawa, 2013, for more details about the ICNALE project). The ICNALE corpus has two access versions: the online version[6] and the download version. The online version is accessible for everyone and is meant, as noted by Ishikawa (2013), to be accessed by users who are less familiar with corpus queries while the download version is targeted at corpus linguistic researchers who can freely use any corpus tools to process the data (Ishikawa, 2013). The download version of the ICNALE is made available to researchers upon request. I contacted Ishikawa and was kindly given access to the ICNALE corpus data.

The question that arises here is that how comparable the ICNALE native speaker corpus is to the BALC corpus. In Section 2.3 in the literature review chapter, I have raised this issue and

[6] The website of the ICNALE corpus can be accessed at http://language.sakura.ne.jp/icnale/.

it is now the time for a discussion of the matter. Table 3-1 demonstrates how comparable the two corpora are.

Table 3-1 *Comparison of the two corpora used in the present study*

Parameter	The BALC	The ICNALE	Comparable
Condition of writing	Timed and under exam conditions	Timed and under exam conditions	Yes
The corpus size	211,178 tokens	90,613 tokens	No
Number of topic prompts asked	2 topics	2 topics	Yes
Topics addressed	(1): Best holiday/vacation/ weekend (2): Favourite film/hobby	(1): Banning of Smoking (2): Part-time jobs	No

Table 3-1 shows that there are similarities and differences in the two corpora used in the study. This in turn implies that the two corpora are comparable, though only to a certain extent. On the one hand, the essays in the two corpora are timed and written under exam conditions, increasing the degree of comparability. The two corpora are also similar in the number of topic prompts that the writers were asked to respond to. On the other hand, however, the size of the two corpora is not the same. As mentioned earlier, the BALC corpus used in the present study is 211,178 tokens whereas the ICNALE native speaker corpus is 90,613 tokens. This may impact on the reliability of the comparison to some extent. However, this disparity with regard to size can be overcome by normalisation, a procedure that is done in the present study to increase the level of comparability in the two corpora (see Section 4.2 in Chapter 4).

The final parameter presented in Table 3-1 is the exact topics addressed in the two corpora. It can be said that the two corpora show similarities and differences with this regard. Although the specific topics addressed in the two corpora are different, they are the same kind of topics in the sense that they are opinion topics, making the two corpora comparable. In both corpora, writers are expected not to report facts but rather to give their opinions. However, the kind of opinion they express is rather different in each corpus. In the BALC corpus, for example, writers provide personal opinions to personal issues. In contrast, in the ICNALE corpus, writers express personal responses to broader social issues. Based on the discussion of the comparison parameters above, it can be concluded that the two corpora show a fairly large degree of comparability and thus the ICNALE native speaker corpus is suitable to be used as a reference corpus in the present study. In the end, it is quite impossible to find a reference corpus that shares similarities with the learner corpus in every respect. Guo (2006, p.55) acknowledges this issue and points out that "[s]ome times, in the real world, however, we have to reach a compromise between what is desirable and what is available".

3.3. Steps of identification of SNs

The aim of the first stage of the study was to identify all the nouns that at any time were used as SNs. This was expected to include 'central and typical' SNs and also nouns which have a range of functions, and which are rarely used as SNs. To identify potential nouns with a signalling function in the two corpora of the current study, three initial steps were taken. First, a word frequency list derived from each corpus was first compiled by means of AntConc software (Anthony, 2014). All words appearing in these frequency lists with a cut-off point of 15 occurrences overall were closely examined. The aim was to identify potential abstract/ borderline nouns. Then, the concordance lines of all abstract/ borderline nouns that appeared in the generated list with the cut-off point in mind were generated in order to

establish whether a particular noun ever had a signalling function as defined in Section 2.7. Next, those nouns identified as SNs were examined in more detail by looking at the concordance lines. However, concordance lines are often insufficient to determine whether a particular noun can be regarded as SN. Therefore, it is necessary to look beyond the immediate context presented by concordance lines to determine whether a noun is being used as a SN. In fact, I often had to expand the concordance line to make an accurate identification of SNs. Thus, an integrated approach of corpus linguistics and discourse analysis has proved beneficial when investigating the discourse phenomenon of SNs.

The careful examination of identified SNs in the two corpora used in the current study is done in two stages. First, the concordance lines of each SN are generated with the help of AntConc corpus tool and then printed out for manual examination. Here, the obvious problem as mentioned previously is that SNs are not often specified within the same concordance line where they are located. The following examples taken from the learner corpus illustrate this issue:

(3): Holidays is the best **time** of my life. In holidays you can do what you want and when you want. **CEPA 5**

(4): Really it was the worst holiday I have ever spent. But the good **thing** is that we are safe, togather and ready to travel again. **CEPA 5**

(5): middle east, it faces the Arabian Gulf. I adore the UAE for many **reasons**, which I will try to explain in this essay. **CEPA 6**

(6): Finaly I liked it so much because it is a true **story** and shows that nothing is impossiple in our lives. **CEPA 5**

In the first two concordance lines (3) and (4), it is easy to identify as whether the noun can be labelled as an SN since the specification of the noun is spelled out in the same concordance line under examination. However, this is not the case in the second two examples (5) and (6), where the concordance lines fail to present the specifics of the SN. It is in this case that I had to go back and consult the AntConc software by using the text function that the software provides to see the expanded environment of the noun in order to establish whether the noun is specified and thus is regarded as an SN.

Secondly, the identified SNs are counted to provide the total number of each identified SN in the two corpora. The problem that arises in the counting stage is whether to include the repeated instances of SNs in the overall frequency. This issue is thoroughly explained in Section 3.5 below.

3.4. General criteria that distinguish the class of SNs

The broad criteria used for the identification of SNs in the current study are those used by Flowerdew and Forest (2015). To repeat the definition, SNs are abstract nouns that demand specification within text for their meaning to be established. However, the idea of abstraction in nouns is problematic since what constitutes an abstract noun cannot be taken for granted. It is a matter of individual interpretation. There certainly are nouns that can be distinguished with ease as abstract nouns. Examples of these nouns include *problem*, *reason*, and *idea*. On the other hand, there also exist a number of nouns which lie on the borderline between

abstract nouns and concrete nouns, not clearly identified as abstract nor as concrete in that they have no physical property. Examples of these nouns are *holiday*, *vacation*, and *trip*.

In the current study, nouns in the borderline between abstract and concrete nouns such as *holiday*, *trip* and *vacation* are counted as SNs since they function to signal a stretch of language in discourse. Thus, my definition of SNs in the study is extended to include such type of nouns. The inclusion of these nouns can be illustrated in the discussion that follows the two examples of SNs *trip* and *idea* (SNs in **bold** and their specifics underlined).

(7): Last summer, my mother and I went to Italy. The **trip** was her idea as a gift for me. **CEPA 6**

(8): They decided to watch a film, I suppoted the **idea** because they will stop talking and watch the film…

The nouns in bold, in examples (7) and (8), are used to encapsulate the underlined stretch of language that has been mentioned in the previous statement. The two nouns are performing a discourse function here despite the fact that the SN in the second example is clearly identified as an abstract noun while the SN in the first example is less clearly identified so. Thus, I see no reason not to include those nouns belonging to the borderline between abstract and concrete nouns when they are functioning to signal stretches of discourse either within clauses or across clauses as shown in the discussed example above.

In light of the preceding discussion, an overarching criterion for what constitutes a SN in the current study is a discourse oriented one. To be assigned the signalling membership, an abstract or borderline abstract noun has to be semantically unspecific where there is a demand for lexical specification in discourse. This specification is to be found either across clauses,

understood via the concept of “encapsulation” and “prospection” (see Section 2.2 for the definitions of these concepts) or within clauses.

My discourse perspective towards the identification of SNs is inspired by the illuminating study by Flowerdew and Forest (2015). In fact, their study is influential in my thinking on the phenomenon of SNs and thus is much cited in my current study. The next section will address some aspects of the specific criteria that Flowerdew and Forest (2015) set for SNs identification. These criteria are useful in my identification of SNs in the two corpora of the study.

3.4.1. SNs in formulaic phrases

Flowerdew and Forest (2015) exclude from their study SNs that occur in formulaic phrases that are used as adjuncts (see Table 3-2 below for examples). For my study, however, I have decided to include these SNs in the overall counts of SNs. In fact, SNs occurring in formulaic phrases are worthy of investigation. The importance lies in the fact that the SNs in these patterns signal a stretch of discourse like other SNs not found in these patterns. Examples of SNs in formulaic phrases as found in the learner corpus are shown in Table 3-2 below:

Table 3-2 *Formulaic phrases of SNs in the learner corpus*

SNs in formulaic adjunct phrases:
In my opinion – in addition – in fact – in the end – in other words – in conclusion – for example – at the same time

The SNs occurring in formulaic phrases as illustrated in Table 3-2 are frequently used in the learner corpus. In fact, many of these nouns, if not exclusively occurring in their formulaic phrases, are found to frequently occur in such phrases. The SNs: *opinion* and *fact*, as will be shown in Chapter 6, are cases in point. The former noun is exclusively used in its fixed phrase (*in my opinion*) while the latter noun is found to predominantly occur in its fixed phrase (*in fact*). So, my concern here is that omitting SNs that occur in formulaic phrases would not give us a full picture of SN uses in the learner corpus. Thus, a decision is taken to include these SNs in the study. The following are examples of some of the SNs listed in Table 3-2 as found to occur in the learner corpus:

(9): In **conclusion**, I learned a few historic facts about the history of Rome and I even learned a few Italian words. **CEPA6**

(10): In **conclusion**, I would like to say that I'm glad that I have this amazing system. **CEPA5**

(11): in **addition** my gradfather was very tiered so I was afreid. **CEPA4**

(12): In **addition**, I use my labtop to do my homework. **CEPA5**

(13): For **example** I travel with my family to the (thiland) **CEPA4**

(14): , for **example**. you can see the amazing birds of the seagulls with their white wings soaring up in the sky. **CEPA6**

3.4.2. Partitive SNs

Partitive nouns are nouns like *part*, *kind*, and *type*. This type of nouns is included in the list of SNs in the current study. Interestingly, *type* and *kind* are not found in Flowerdew and Forest's list of SNs, although they provide *kind* as a typical example of SNs identified in the literature (p.10) and they also assert that partitives are abundant in academic texts (p.62). This seems paradoxical since the corpus that they analyse in their study was mainly academic. However, Flowerdew and Forest's (2015) methodological approach towards partitives offers an explanation why partitives are not included in their list of SNs. According to them, partitives serve to pre-modify other nouns and also other SNs and thus it is "the canonical SN" that should be assigned the signalling function (p.62).

The methodological approach that Flowerdew and Forest's (2015) study adopts when treating instances of partitives is problematic for the following reasons. First, it is not always the case that partitives act as pre-modifier of other SNs. An illustrative example of this issue, found in the native speaker corpus used in the current study, is presented below:

(15): The bad **part** of this is that it will place unnecessary stress on students who are already buried with schoolwork.

(16): There is no way to please this **kind** of individual.

Examples (15) and (16) show that it is not always the case that partitives pre-modify SNs. Thus, it would be more accurate to treat instances of partitives as SNs in their own right. My approach is that even if the partitives pre-modify canonical SNs, it is the partitives that act as the SNs.

Thus, my approach towards noun phrases that include partitives is that the partitive only should be regarded as SNs. When one says “part of the problem”, for instance, it is not the whole problem that is being signalled here but rather the “part” thing of it. Examples (17) and (18) are illustrative of this. In these examples, I count *part* as an SN but not *trip* or *experience*.

(17): There, I learnt how to walk, trot and canter with a horse. Actually, that was the most interesting **part** of the whole trip. **(CEPA6 Student)**

(18): One enormous **part** of any gourmet experience is a variety of smells one can enjoy in addition to the taste of the food, and cigarette smoke absolutely destroys this **part** of the experience. **(Native speaker)**

3.4.3. Bidirectional instances of SNs

This specific criterion is interesting. In some cases, an SN may establish two complement clauses that specify the lexicalisation of the noun, but only one clause can be said to be “equative” with the SN under investigation (Flowerdew and Forest, 2015, p.55). In the identification of SNs in the current study, a number of instances of bivalent SNs have been identified. The following example taken from the learner corpus illustrates this phenomenon:

(19): Our **aim** was to have fun, but fishing is not the greatest **way** to have fun while a one year old is crying. **(CEPA 6 student)**

The student in the above example makes use of two SNs: *aim* and *way*. The SN *aim* is unidirectional in that it establishes one complement clause (to-infinitive clause) as its specification. The second SN *way*, however, is different in that it establishes two complement clauses as part of the lexical specification for *way* as seen in the underlined (ing-clause) and in the shaded (to-infinitive-clause). It is only the underlined (ing-clause) that spells out the content of *way* and hence is identified as the specification of the SN *way*.

3.5. Repeated instances of SNs

As a starting point here, it is useful to reiterate that SNs are to be counted as SNs only when they are specified within text. It is widely assumed in the literature on SNs that any given SN has a single specification (Flowerdew and Forest, 2015, p.50). They acknowledge that although this assumption is effective as a starting point in getting a grip on the essential characteristics of SNs, it could be misleading since SNs may have one specification in common and may also have more than one specification in the evolving discourse.

I am in agreement with Flowerdew and Forest's observation with regard to the complication that arises from repeated instances of SNs. Example (20), taken from the learner corpus, shows a repeated instance of an SN that illustrates the complex nature of repeated instances of an SN when one is to identify and quantify SNs.

(20): The last summer holiday was the worst holiday. I went to India with my parents, sisters, brothers and ofcours our maid. This was the really **idea** of our maid. She told us that India was the most perfect place to visit. She also talled us that their are many tourisum. I wasn't really satisfied about going to India but my father said it

was a really a good **idea** so that he could make money in India. We travelled there and when we just intered the airport all the poor people rushed to us, because they knowed that we are from the Gulf, and they have an **idea** that all the gulf people are rich. **(CEPA 5 student)**

In this example, the SN *idea* is repeated three times. The first instance is anaphoric SN functioning to encapsulate the previous statement. The second use of *idea* is anaphoric encapsulating the gerund clause (going to India), which is the same specification of *idea* in the first instance but spelled out differently. The third instance, however, is understood to be a new *idea* not related to the previous instances of *idea*. It is unproblematic to regard this as a separate instance of SN since the specification is new and unrelated to the previous specifics of the same noun. It is when a repeated instance of an SN shares the same specification in discourse that quantifying SNs gets complicated and raises methodological questions. The main raised questions are listed below and answered afterwards:

- Do we only treat the first instance of a repeated SN sharing the same lexical specification as SN and the rest as merely repetitions of the same SN?
- Do we only treat the first instance of a repeated SN when the first instance acts as a 'hyper theme' and the repeated instances of the SN are merely multiple lexical specification of the 'hyper theme' SN?

The answer to these questions depends on the focus of the study. My discourse focus determines that the repeated instances of a given SN when sharing a single lexical specification are counted as SNs since they add to lexical cohesion through repetition. The repeated SN can function to signal a stretch of language as well as to evaluate discourse.

Looking back at example (20) above, the first instance of *idea* is signalling a previously mentioned text while the second one is evaluating the *idea* as good. Not including a repeated SN, in this case, does not render the full picture of how the resource of SNs operates at a discourse level.

With regard to the second question, the 'hyper theme' SN and the repeated SNs providing specifics for it are all counted as SNs. The following example is illustrative of this complexity:

(21): I liked this place for many **reasons**. The first **reason** is that it is the house of allah where we can be ask allah what we want and forgivenes.The second **reson** is that all the rich, poor, black or white are the same in this place. **(CEPA 5 Student)**

Example (21) shows three instances of the SN *reason*. The first occurrence of the SN functions as a hyper theme prospecting what is to come in the forward discourse (the first *reason* and the second *reason*). The second and third instances are quite problematic as to whether they should be included as SNs or only considered as specifics of the main SN in hyper theme position and not counted separately. The decision in the current study is to include all the three instances as SNs. The reason for this choice is that the study prioritises how SNs function in discourse not only across clauses but also within clauses. Another reason is that by not including them as SNs the style of the students in repeating the SN will be missed.

3.6. Quantification and further investigation of SNs

So far, I have discussed the steps and the criteria used in the present study to identify SNs. Once the SNs in the learner and native speaker corpora had been identified, the next step was to quantify them. (A list of all candidate SNs, with their raw frequencies, as signalling and non-signalling, is given in Appendix B (learner corpus) and Appendix D (native speaker corpus). A list of SN uses only, with raw frequencies, is given in Appendices C and E. For illustrative purposes, examples of signalling and non-signalling instances of the SN *reason(s)* are provided in Appendix F (learner corpus) and Appendix G (native speaker corpus).) The outcome of this process is a list of all the SNs and the frequency with which they are used as SNs in each of the two corpora. (The list of SNs, with their normalised frequencies in the learner corpus, is presented in Table 4-1 and the list of SNs, with their normalised frequencies in the native speaker corpus, is shown in Table 4-5.) The results of the frequency of all the SNs identified to occur in the learner and native speaker corpora are provided in Chapter 4. Then, particular SNs were selected for further detailed investigation. First, three frequent and shared SNs in the two corpora were chosen. The results of these three nouns are given in Chapter 5. Next, SNs with particular discourse functions were examined. The results of the selected SNs are presented in Chapter 6 and Chapter 7. The rationale for the selection of the SNs is explained in each chapter. Further details of the methodology are also given in each chapter.

3.7. Summary of the chapter

This chapter has given a description of the corpora used in the present study, as well as a detailed account of how comparable the two corpora are. The chapter has also provided an account of the steps followed to identify and quantify the SNs in the present study. I have argued that the set of SNs identified by Flowerdew and Forest (2015) needs to be expanded to

account for the writing in the learner corpus particularly. Thus, although I have followed Flowerdew and Forest's (2015) methodology for the most part, I have interpreted it in light of my data. The following chapter presents the findings of the study with regard to the overall frequency of SNs in the learner and native speaker corpora.

CHAPTER 4: THE FREQUENCY OF SNs IN THE LEARNER AND THE NATIVE SPEAKER CORPORA

4.1. Overview

This chapter provides the results of the study of frequency of SNs in the learner and native speaker corpora. The aim is to give an account of the frequency of SNs in these two corpora. The first section (4.2) explains how the frequency of SNs in the two corpora is reported. The other two sections (4.3 and 4.4) present the list of SNs in the two corpora of the study. Section 4.5 compares the use of SNs in the two corpora in terms of frequency and type. The findings of the chapter are summarised in Section 4.6.

4.2. Reporting of frequency of SNs in the two corpora

The first step in the process is to establish the relative frequency of the SNs in the learner and native speaker corpora. This is carried out by normalising the frequency of the individual SNs and the overall frequency of identified SNs in the two corpora. The frequency of a given SN is calculated per 100,000 words (formula: a normalised frequency = raw frequency of an SN * 100,000 ÷ overall corpus size).

As mentioned in the methodology chapter, the two corpora used in the current study are of unequal sizes. Thus, using raw frequency to explain the differences between the two corpora in terms of the use of SNs may distort the results obtained since raw frequency does not tell us much about the effect of the size of the two corpora being compared. It is for this reason that normalised frequency is appealed to instead of raw frequency as normalised frequency takes into consideration the size-difference of the two corpora used. However, the crucial point when normalising data is the appropriateness of the choice of a common base based on

which data is normalised. It has been noted that "results obtained on an irrationally enlarged or reduced common base are distorted" (McEnery, Xiao, and Tono, 2006, p.53). As shown in the formula presented above, the frequency of SNs is normalised to 100,000 words. The reason for choosing 100,000 as a common base in this study is that this base can be comparable to the size of the two corpora used in the current study; the size of the learner corpus is 211,178 word tokens while the native speaker corpus is 90,613. Thus, when using 100,000 as a common base, the normalised frequency is often less than the raw frequency. This is important in rendering a reliable representation of the actual frequency of SNs used in the corpora or sub-corpora being compared.

4.3. Frequency of SNs in the learner corpus

Table 4-1 presents all the instances of SNs in the learner corpus in descending order according to normalised frequency (As noted in Chapter 3, Table 4-1 includes these nouns only when they are used as SNs). Table 4-1 also provides the normalised frequency of each SN in the singular and in the plural.

Table 4-1 *Normalised frequency of SNs in the learner corpus per 100,000 words*

Rank	Non-lemmatised SN	Normalised singular frequency per 100,000 words	Normalised plural frequency per 100,000 words	Normalised frequency per 100,000 words
1	Holiday	235	4	239
2	Place	126	22	148
3	Time	139	5	144
4	Thing	75	52	127
5	Trip	53	0	53
6	End	43	0	43
7	Story	41	0	41
8	Reason	17	18	35
9	Problem	20	7	27
10	Vacation	26	0	26
11	Kind	11	12	23

Table 4-1 *(Continued)*

Rank	Non-lemmatised SN	Normalised singular frequency per 100,000 words	Normalised plural frequency per 100,000 words	Normalised frequency per 100,000 words
12	Addition	20	0	20
13	Way	16	3	19
14	Idea	18	1	19
15	Opinion	19	0	19
16	Moment	14	3	17
17	Example	14	2	16
18	Part	15	1	16
19	Activity	12	2	14
20	Group	12	2	14
21	Experience	13	1	14
22	Chance	13	0	13
23	Fact	13	0	13
24	Journey	12	0	12
25	Conclusion	11	0	11
26	View	9	1	10
27	Dream	8	1	9
28	Feeling	8	1	9
29	News	9	0	9
30	Rest	8	0	8
31	Event	3	3	6
32	Plan	5	1	6
33	Death	6	0	6
34	Word	2	3	5
35	Accident	5	0	5
36	Job	4	1	5
37	Point	5	0	5
38	Information	4	0	4
39	Minute	3	1	4
40	Favourite	3	1	4
41	Change	2	1	3
42	Hope	3	0	3
43	Matter	3	0	3
44	Nature	2	0	2
45	Use	1	1	2
46	Means	1	1	2
47	Advice	1	0	1
48	Thought	1	0	1
		Total**= 1084**	Total= **151**	Total= **1235**

As shown in Table 4-1, the learner corpus contains 48 SN types. The overall normalised frequency of these types accounts for 1235 tokens per 100,000 words. A closer look at the table clearly shows that the top-10 SNs represent most of the occurrences of all SNs that are used in the corpus, comprising nearly 70% of the overall total of SNs. The SNs that are most frequently used in the learner corpus relate to the topics that the students are asked to write about. The essays comprising the learner corpus can be categorised as: narrative, in which learners are asked to write about the worst holiday ever or a perfect vacation; and as descriptive, in which they are asked to describe a favourite possession or movie. These essays can also be said to be opinion-oriented in nature in that students are encouraged to express their personal opinions on the topics they choose to address.

Table 4-1 shows that the choices of SNs in the learner corpus are often sensitive to the topic being talked or written about. As will be argued below, this is true of the native speaker corpus as well. This is in line with Flowerdew and Forest's (2015, p.91) argument that the choice of particular SNs "can be linked to the subject area and vocabulary of the field". In their study, the SN *right*, for example, is found to be predominantly used in the law corpus. My learner corpus, however, suggests that one can go further and argue that the specific topic and not just the discipline and genre has an influence on the choice and the frequency of particular SNs. There are a number of SNs that can be said to have been influenced by the topics covered in the learner corpus and would not be expected to occur in other learner corpora unless the essay topics were the same. Examples of these topic related SNs are *holiday*, *trip*, and *vacation*. These examples are located in the borderline of abstract and concrete nouns as discussed in the methodology chapter. These nouns are not included in lists of SNs in the existing literature that has dealt with the phenomenon of SNs and similar phenomena (Hunston and Francis, 2000; Schmid, 2000 Mahlberg, 2005; Flowerdew and

Forest, 2015). The reason for the exclusion of such SNs in the aforementioned studies can be attributed to three factors: (1) the focus of these studies as to which aspect of SNs is investigated, (2) the context from which the data of these studies were drawn and (3) the interpretation of the notion of 'abstract'.

Table 4-1 also shows that the students rely on four SNs: *holiday(s)*, *place(s)*, *time(s) and thing(s)*. These nouns come at the top of list of all instances of SNs used in the learner corpus. More importantly, these four nouns together account for over half (53.27%) of all the SNs in the corpus. Table 4-2 below gives more detail of these very frequent nouns. It presents the total frequency of these nouns in the corpus in general and the total instances where these nouns are used as SNs. Nouns are counted as SNs only when they are lexically specified within the text, as discussed in the methodology chapter.

Table 4-2 *Frequencies of the four most frequent SNs in the learner corpus*

Noun (singular and plural)	**Total frequency of the noun in the corpus**	**Total raw frequency of the noun in the corpus when used as SN (%)**	**Normalised frequency per 100,000 words for each grammatical forms of the SN**
Holiday	1625	497 (31%)	235
Holidays	97	8 (8%)	4
Place	506	267 (53%)	126
Places	189	47 (25%)	22
Time	867	294 (33.33%)	139
Times	75	11 (14%)	5
Thing	352	159 (39%)	75
Things	365	109 (33%)	52

Table 4-2 shows that the most frequent SN in the learner corpus is *holiday(s)*. This SN comes at the top of the table, with the total of 497 singular instances and 8 plural instances. The noun *place(s)* is the second most frequent SN. *Place(s)* occurs 267 times in its singular form and 47 times in its plural case. The third common SN is *time(s)*, accounting for 305

occurrences; 294 instances of them are singular and plural form is realised in 11 occasions. *Thing(s)* ranks fourth in the list, comprising 159 and 109 instances in the singular and plural form respectively.

The frequency findings of these four nouns indicate that the grammatical forms of these SNs are unequally frequent. This is apparent in the case SNs *holiday*, *place* and *time* where the singular form is by far more frequent than the plural form of these nouns. However, this is less clear in the case of the SN *thing* where both grammatical forms of the noun are frequently used in the corpus. There are a number of reasons for this apparent discrepancy. One reason is that some SNs are only used in singular form because they are non-countable nouns in the first place, at least in some contexts. Examples, as found in the learner corpus, include *rest*, and *end* as shown in the following lines:

(1): I became sick so I spent the **rest** of my holiday in the hospital and my best friend didn't come... **(CEPA 5)**

(2):...that it was time to go back to Al Ain. This was the **end** of my miserable weekend. **(CEPA 6)**

The second reason is that the distributional behaviour of SNs may be affected by the genre or the discipline where these nouns are used (Flowerdew and Forest, 2015, p.87). This is a credible observation, as I believe that the distribution of grammatical forms of SNs, as the choice of SN itself, is often influenced not only by genres and disciplines, but also by the topic that is being talked or written about. This is demonstrated in the extensive use of, for example, the SNs: *holiday* and *vacation* in their singular form. The topic that students are asked to respond to can be held responsible for the extensive use of the singular form over the

plural form. As mentioned previously, the learners are asked to write about **"**the**"** worst holiday ever or **"**the**"** perfect vacation.

A closer look at the list of all SNs as shown in Table 4-1 supports the observation that SNs tend to show preference to one grammatical form over the other. SNs in the learner corpus are predominately used in their singular form, not only in the case of the first four most frequently SNs in the list as mentioned above but also in the overall uses of all SNs in the corpus. The following table shows this:

Table 4-3 *Total normalised frequency of both grammatical forms of SNs in the learner corpus*

Total normalised singular frequency of SNs per 100,000 words	**Total normalised plural frequency of SNs per 100,000 words**
Total= 1084	Total= 151

Table 4-3 clearly shows that SNs in the learner corpus are more likely to be used in their singular form. As will be observed below, this is also true of the native speaker corpus.

I now turn to the question of learner progression. It is hypothesised in this thesis that the learners at the various levels vary with regard to the use of the resource of SNs. To initially test this hypothesis, the first top-5 SNs as found in the list of SNs in the learner corpus will be examined. Figure 4-1 below shows the normalised number of instances of the most commonly used SNs in the learner data across the proficiency levels:

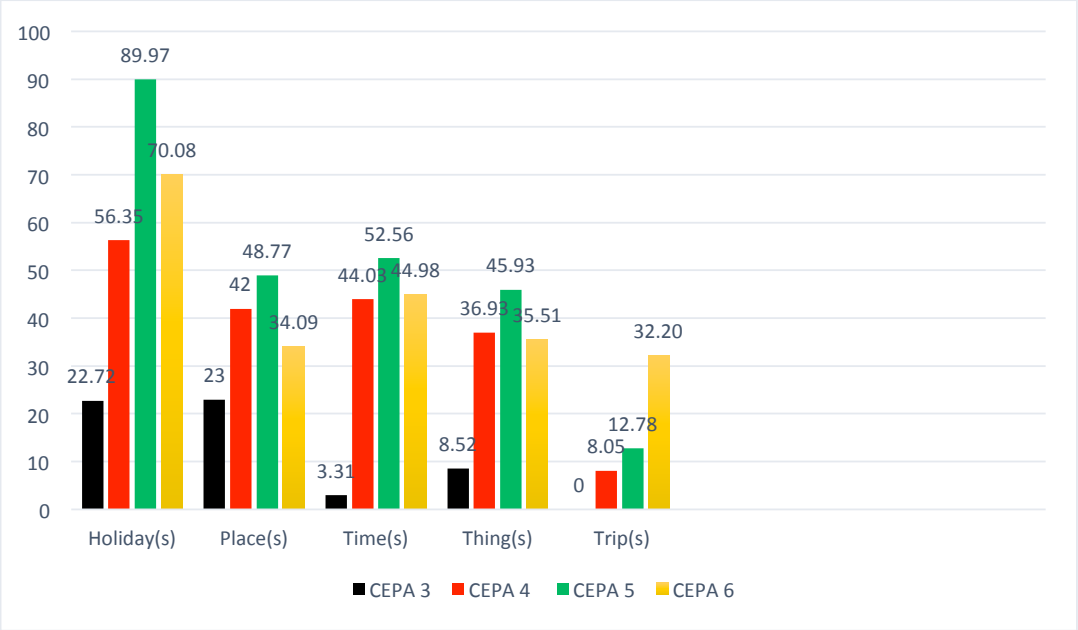


Figure 4-1 *Normalised frequency of the top-5 SNs across the four levels per 100,000 words*

As demonstrated by Figure 4-1 above, the learners across the four proficiency levels make use of all the SNs in question with the exception of the SN *trip*, which is not used as an SN in CEPA 3 essays (the lowest proficiency essays). As mentioned earlier, *holiday* and *trip* are topic related SNs and thus an appropriate use of these nouns can be said to create cohesion in the learners' writing. The highly rated essays in terms of cohesion (CEPA 5 and CEPA 6) contain more instances of these two topic-related SNs than the other essays of the lowest proficiency levels (CEPA 3 and CEPA 4).

It is also observed from the normalised frequency of the four SNs, as shown in the figure, that the essays comprising CEPA 3 make less use of these nouns. The increase of the use of SNs as proficiency level increased is noticeable (the apparent fall between CEPA 5 and CEPA 6 will be discussed below). This can be an indicator that the learners gain a greater understanding of how SNs are used as proficiency level increases. To better assess whether

such understanding is developed in learners' writing across different levels, it would be useful to (1) look at the total normalised frequency of all SNs in each level and (2) to investigate the types of SNs used across proficiency level.

Table 4-4 presents the same list of SNs as shown in Table 4-1. However, Table 4-4 provides the normalised frequency of the SNs used by the learners in each level of proficiency. It also shows the total number of SN types across the levels.

Table 4-4 *Normalised frequency of SNs per 100,000 words across levels along with the number of SN types occurring in each level*

Non-lemmatised SN	Normalised frequency per 100,000 in CEPA 3	Normalised frequency per 100,000 in CEPA 4	Normalised frequency per 100,000 in CEPA 5	Normalised frequency per 100,000 in CEPA 6
Holiday	22.72	56.35	89.97	70.08
Place	23	42	48.77	34.09
Time	3.31	44.03	52.56	44.98
Thing	8.52	36.93	45.93	35.51
Trip	0	8.05	12.78	32.20
End	3.78	12.31	14.67	12.31
Story	1.89	11.36	17.04	10.41
Reason	1.89	4.26	10.89	18.46
Problem	0.94	7.57	12.78	5.68
Vacation	0	1.42	7.10	17.04
Kind	0.94	6.62	9.94	6.15
Addition	0.47	4.26	8.52	6.62
Way	0.47	2.36	6.62	9.94
Idea	0.94	5.20	6.62	6.15
Opinion	2.36	7.10	5.20	4.26
Moment	0	0.47	5.68	7.57
Example	4.26	3.78	3.31	4.73
Part	0.94	3.31	4.73	7.10
Activity	5.68	2.84	3.78	2.36
Group	0	0.94	4.73	8.52
Experience	0	0	2.36	11.36
Chance	0	0	0.94	11.83
Fact	1.89	0.47	1.42	9.47
Journey	0.47	1.89	5.20	4.73
Conclusion	0.47	2.36	2.84	5.20
View	0	1.89	2.84	5.20

Table 4-4 *(Continued)*

Non-lemmatised SN	Normalised frequency per 100,000 in CEPA 3	Normalised frequency per 100,000 in CEPA 4	Normalised frequency per 100,000 in CEPA 5	Normalised frequency per 100,000 in CEPA 6
Dream	0	0.94	2.84	4.73
Feeling	0	2.84	0.94	4.73
News	0	1.42	4.73	2.36
Rest	0	1.42	2.84	3.78
Event	0	0	1.89	4.26
Plan	0	1.42	1.42	3.31
Death	0	1.89	1.42	2.36
Word	0	0.47	0.94	4.26
Accident	0	1.42	2.36	1.42
Job	0	0.47	1.42	3.31
Point	0	1.42	0.47	3.31
Information	0	1.42	1.89	0.47
Minute	0	0	0.47	3.78
Favourite	0	0	1.89	2.36
Change	0	0	0.94	1.89
Hope	0	1.42	.47	.94
Matter	0	.94	.47	1.42
Nature	0	0.47	0.94	0.94
Use	0	0	0.94	1.42
Means	0	0	0	1.89
Advice	0	0	0.94	0.47
Thought	0	0	0.47	0.47
Total of normalised frequency of SNs in each level	**84.36**	**285.73**	**417.91**	**445.83**
The number of SN types identified in each level	**19**	**38**	**47**	**48**

Table 4-4 demonstrates that there is a quite substantial difference between the levels. The normalised frequency of SNs has witnessed a sharp increase as the proficiency level of the learners increases. Figure 4-2 below presents a line graph that illustrates the difference in terms of the frequency and type of SNs across the proficiency levels:

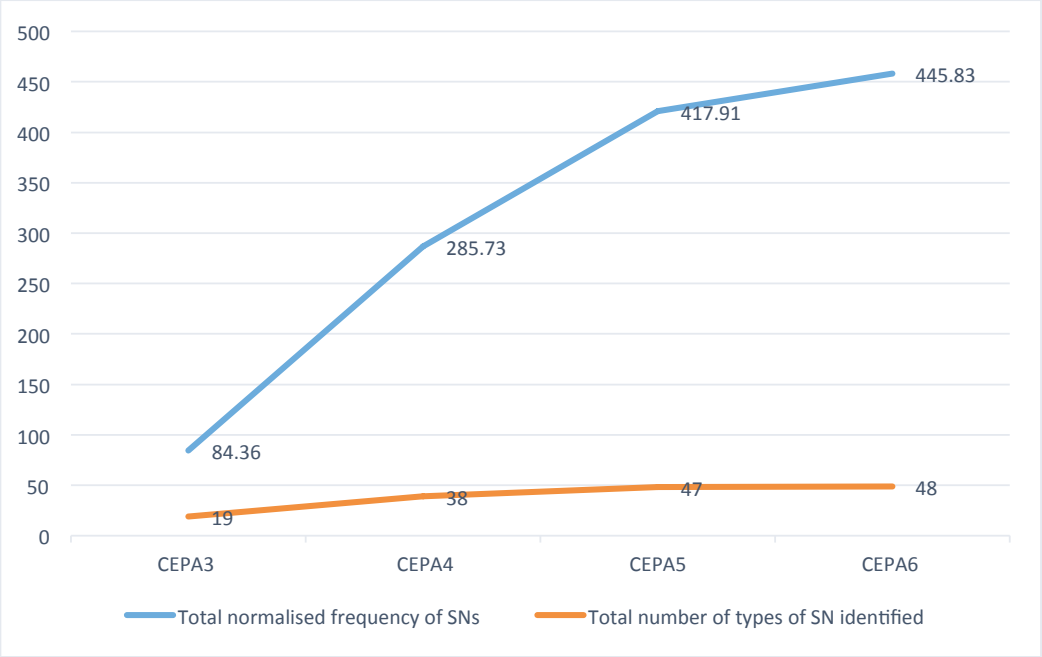


Figure 4-2 *Normalised frequency and total number of SN types occurring in each level*

Overall, the figure illustrates that there is an increase in the number and type of SNs used across the four proficiency levels. As displayed in Figure 4-2, the normalised frequency of SNs per 100,000 words stands at 84.36 in CEPA 3. That number increases dramatically to 285.73 in CEPA 4 and then to 417.91 in CEPA 5, which is followed by a steady increase in CEPA 6 with 445.83 SNs per 100,000 words. Figure 4-2 also shows that there is a slight increase of SN types across the proficiency levels; 19 in CEPA 3, 38 in CEPA 4, 47 in CEPA 5 and 48 in CEPA 6. In conclusion, both tokens and types of SNs have increased as the level of the learners' proficiency increases. However, this increase has levelled out in the higher proficiency levels: CEPA 5 and CEPA 6.

A closer look at Table 4-4 presented earlier shows that there are more instances of the more abstract SNs in the higher proficiency levels: CEPA 5 and CEPA 6 as opposed to the lower proficiency levels: CEPA 3 and CEPA 4. For example, the instances of the classical abstract

SN *reason(s)* in CEPA 5 and CEPA 6 are 29.35 per 100,000 words. The corresponding figure in CEPA 3 and CEPA 4 is 6.15 per 100,000. Similarly, the abstract SNs *problem(s)* and *idea(s)* occur more in CEPA 5 and CEPA 6 as opposed to CEPA 3 and CEPA 4.

4.4. Frequency of SNs in the native speaker corpus

Table 4-5 presents all the instances of SNs in the native speaker corpus in descending order of normalised frequency. Table 4-5 also gives the normalised frequency of each SN in the singular and in the plural.

Table 4-5 *Total normalised frequency of SNs in the native speaker corpus*

Rank	Non-lemmatised SN	Normalised singular frequency per 100,000 words	Normalised plural frequency per 100,000 words	Normalised frequency per 100,000 words
1	Time	240	11	251
2	Way	108	10	118
3	Reason	31	86	117
4	Ban	66	29	95
5	Thing	64	31	95
6	Right	68	3	71
7	Idea	64	2	66
8	Opinion	53	6	59
9	Fact	52	2	54
10	Example	53	0	53
11	Issue	49	3	52
12	Case	40	11	51
13	Experience	42	4	46
14	Part	42	3	45
15	Statement	44	0	44
16	Choice	40	3	43
17	Skill	8	34	42
18	Job	18	18	36
19	Place	16	20	36
20	Problem	30	6	36
21	Opportunity	28	6	34
22	Habit	30	1	31
23	Kind	24	5	29
24	Result	28	1	29
25	Chance	26	2	28
26	Freedom	24	0	24

Table 4-5 *(Continued)*

Rank	Non-lemmatised SN	Normalised singular frequency per 100,000 words	Normalised plural frequency per 100,000 words	Normalised frequency per 100,000 words
27	Point	21	3	24
28	Addition	22	1	23
29	Solution	20	3	23
30	Variety	22	1	23
31	Benefit	3	17	20
32	Sense	20	0	20
33	Majority	19	0	19
34	Matter	18	1	19
35	Need	16	3	19
36	Situation	16	3	19
37	End	18	0	18
38	Position	15	1	16
39	Activity	4	10	14
40	Responsibility	12	2	14
41	Smell	12	2	14
42	Cost	8	5	13
43	Risk	11	2	13
44	Effect	4	6	10
45	Business	9	1	10
46	Cause	10	0	10
47	Pressure	5	1	6
48	Use	6	0	6
49	Thought	3	2	5
50	Deal	3	1	4
51	Stress	3	0	3
52	Culture	2	0	2
53	Hope	2	0	2
54	Mean	0	1	1
		Total= 1592	**Total= 363**	**Total= 1955**

As presented in Table 4-5, the native speaker corpus contains 54 SN types. The overall normalised frequency of the SNs in the corpus is 1955 SNs per 100,000 words. This overall frequency of SNs will be compared with the overall number of SNs in the learner corpus in Section 4.5 below. The top-five SNs in the native speaker corpus are *time(s)*, *way(s)*, *reason(s)*, *ban(s)* and *thing(s)*. The normalised frequency per 100,000 words of these SNs is

251 instances for *time(s)*, followed by *way(s)* (118), *reason(s)* (117), *ban(s)* (95) and *thing(s)* (95).

Among the top-5 SNs used in the native speaker corpus, two SNs can be said to have been influenced by the topic on which native speaker of English participants are asked to respond to. The most clear instances of these topic-related SNs that occur among the five most frequent SNs used by native speakers is *ban(s)* while the most surprising one is the SN *reason(s)*. Including *reason(s)* as among topic-dependant SNs needs some justification. The respondents are asked to respond to the following prompt:

Table 4-6 *Essay prompt for native speaker corpus*

Do you agree or disagree with this statement? Use reasons and specific details to support your claim (A) It is important for college students to have a part time job. (B) Smoking should be completely banned at all the restaurants in the country.

The prompt presented in Table 4-6 includes the SN *reason(s)* and therefore this SN is said to be topic-dependant. Another topic-related SN that occurs among the top-15 SNs in the native speaker corpus is *statement*. This SN also appears in the prompt and thus is regarded as a topic-dependant SN. The following are examples of how the SNs *reason(s)* and *statement* are used in the native speaker corpus:

(3): I agree with the **statement** that it is important for college students to have a part time job.

(4): Smoking should be banned in all restaurants in the country, is a strong but true **statement**. This **statement** shows that the government is taking an active role in the health and welfare of the community.

(5): I disagree with this **statement**. One of the main **reasons** is because it creates billions of revenue for the government that can be then recycled and used to restore old buildings or some other areas of need in our cities.

(6): I agree with this **statement** for various **reasons**. The first **reason** is because studying costs a lot of money and it should not just be the job of the parents to pay for all of our education.

As argued before in this paper and as shown in the above examples (3) to (6), the choice of SNs is often topic sensitive. This is not say that all the SN types found in a corpus would be directly related to the topic being addressed but surely the topic can influence the choice and the frequency of particular SNs. This observation lends support to what has already been expressed in the literature that has dealt with the phenomenon of SN with regard to the influence of discipline and genre on the choices of SNs (e.g. Flowerdew and Forest, 2015). However, it goes beyond that to add that the topic **itself** that one is asked to talk or write about is likely to influence the use of certain SNs.

4.5. SNs in the learner and the native speaker corpora: comparison

The aim of this section is to compare the learner and native speaker corpora, in order to determine to what extent the two corpora either resemble or differ in the use of SNs. The

section is organised into two main sections. Section 4.5.1 is concerned with the comparison of the two corpora in terms of SN tokens and Section 4.5.2 compares SN types used in the two corpora.

4.5.1. Frequency of SNs

Table 4-7 presents the overall frequency of SNs in both corpora, in order to determine to what extent the two corpora resemble or differ in the use of SNs.

Table 4-7 *Total normalised frequency of SNs across the two corpora*

Total Normalised frequency of SNs in the learner corpus per 100,000 words	**Total Normalised frequency of SNs in the native speaker corpus per 100,000 words**
Total= 1235	**Total= 1955**

As shown in Table 4-7, the learner corpus contains fewer SNs per 100,000 words than the native speaker corpus. Since the total frequency of SNs in the two corpora is normalised, it can be concluded that there is under-representation in the utilisation of the resource of SNs in the learner corpus compared with the native speaker corpus. However, to be confident as to whether this difference is significant, I have used the log-likelihood tool created by Rayson (2017), available online at http://ucrel.lancs.ac.uk/llwizard.html (accessed February 10, 2017). This calculator is useful in that it takes into consideration the different sizes of corpora (see Rayson and Garside, 2000 for more detailed on how this tool works). According to Rayson (2017) and Rayson and Garside (2000), significance is interpreted depending on the obtained log-likelihood result as follows: (1) if the result is more than 6.63, the probability of the difference between the two corpora occurring by chance is less than 1% with the

probability value of $p<0.01$; (2) if the outcome is 3.84 or greater, the likelihood of the difference between the compared corpora happening by chance is less than 5% with the probability value of $p<0.05$. The $p<0.05$ value is widely accepted as the minimum statistical measure of significance (Dörnyei, 2007).

Table 4-8 below shows the log-likelihood results of the total frequency of SNs in both corpora obtained from likelihood tool created by Rayson (2017):

Table 4-8 *Log-likelihood results for the total number of SNs in the two corpora*

Total Normalised frequency of SNs in the learner corpus per 100,000 words	**Total Normalised frequency of SNs in the native speaker corpus per 100,000 words**	**Log-likelihood result (LL)**
Total= 1244	**Total= 1890**	**-217.97**

As shown in Table 4-8, the difference between the two corpora is highly significant. There is much under-representation of SNs in the learner corpus compared to that of native speakers. This is indicated by the minus symbol attached to the log-likelihood score. As stated by Rayson (2017), the log-likelihood value itself is always a positive number. However, the addition of a minus or plus symbol is to indicate either under-representation or over-representation in corpus 1 compared to corpus 2.

This significant difference in terms of the use of SNs between the learner and the native speaker corpora at least in the light of the quantitative result obtained can be linked to a difference in proficiency. Arabic speakers who are learning English — even at advanced

proficiency levels— do not use SNs to organise and to evaluate a stretch of discourse as much as the native speakers do.

The quantitative analysis as shown in Figure 4-2 and Table 4-8 suggests the use of SNs is linked to proficiency. However, it would be more reliable to adopt a more qualitative approach to assess as to what extent overall proficiency can be enhanced by encouraging learners to use more SNs. The next section will shed more light on the learners' use of SNs.

A qualitative analysis of the use of SNs in the learners' writing

The aim of this section is to qualitatively answer the following question: would the learners' writing be better if they use SNs? In order to answer this question and to discuss the effect of the use of SNs in the students' writing, essays from the learner corpus have been randomly selected for deeper scrutiny. Having closely looked at these essays, instances have been identified where if the learners had used SNs their writing could have been better. Some examples of these instances are provided below. In these examples, I discuss the influence of the use of SNs on the quality of the selected texts in terms of the cohesion of the texts (linking stretches of discourse together), the expression of stance and causality in the texts. I also make suggestions of where SNs could have been added to make the students' writing better. It might be worth noting that SNs cannot just be added to existing sentences. Using an SN would in most cases involve rephrasing of existing discourse, and it is that rephrasing along with the use of an SN that constitutes the improvement in the learner's writing.

Before I present the qualitative findings, two points regarding the task undertaken in this section are worth mentioning. The first is that the task involves looking for something that is

not there in the learner's writing. Hence, it is a difficult task to pursue. The second is that the task involves a degree of subjectivity. This brings us to the findings.

Two whole essays taken from the randomly selected essays are provided below in examples (7) and (8). The reason why the whole essay is presented in each of these examples is to show how the use of SNs within the context of the whole text could have made the text better. However, any subsequent examples following these two examples will be limited to extracts taken from the learner corpus. All the examples are supplied with a suggested revision of the text in question, which points out the instances where SN(s) could have been added. Any addition to the text is presented between brackets.

In Example (7), there are a number of instances where the addition of SNs to the text could have made the text better. The possible SNs that could have been used by the learner are *reasons*, *opinion*, *fact*, and *conclusion*.

Example (7):

A learner's text:
The best time of my life is the holiday. It let me have a rest. It helps me to discover many new things such as:new clutures by travelling and new hobbies by practicing in the club. In my last summer holiday, I travelled to India. The crowdest country. I went there with my family. I planned for what I was going to do and the places I am going to visit. First I visited Mumbai city. It was very big city with many buildings and shopping centers. I mostly liked the beach it was refreshing. Then, I visited the (M.g city) which was located on the mountain, it was the most beautiful city I have ever seen. And so on I moved with my family from city to another. We saw many attractive things like the zoo and the Taj Mahel.

I think it was my best holiday. It helped me to learn the Indiain cluture. You should really visit India.

A suggested revision of the learner's text with the addition of SNs:

The best time of my life is the holiday [for two main **reasons**]. **[**The first is that**]** it let me have a rest. [The second is that] it helps me to discover many new things such as:new clutures by travelling and new hobbies by practicing in the club.

In my last summer holiday, I travelled to India. I went there with my family. [In my **opinion**,] [India is] the crowdest country. [Inspite of this **drawback**, it is possible to have good **experiences** there.] I planned for what I was going to do and the places I am going to visit. First I visited Mumbai city. It was very big city with many buildings and shopping centers. I mostly liked the beach. It was refreshing. Then, I visited the(M.g city) which was located on the mountain. It was the most beautiful city I have ever seen. And then I moved with my family from one city to another. We saw many attractive things like the zoo and the Taj Mahel.

[In **conclusion,**] I think it was my best holiday. It helped me to learn the Indiain cluture. You should really visit India.

As shown in example (7), the learner begins his/her text by presenting his/her opinion about a topic: "The best time of my life is the holiday.". This is followed by two sentences: "It let me have a rest." and "It helps me to discover many new things such as:new clutures by travelling and new hobbies by practicing in the club.". The textual link between these two statements in relation to the first sentence of the text that presents the learner's opinion may be unclear. The use of an SN at the end of the first sentence, however, could have made this link more explicit. As shown in the suggested version of the learner's text in example 7, the learner could have used, for example, the SN *reasons* at the end of the first sentence to signal that the coming discourse is to present the reasons for his/her opinion and the relationship

between these statements is to be understood as that of cause and effect. Thus, the two sentences following the first sentence are to be seen as the specifics of the SN and thus these two sentences link up with the SN *reasons* found in the hyper thematic position of the first sentence. Figure 4-3 below demonstrates this cohesive link created by the use of an SN between the three sentences. Thus, the use of SN may be said to have improved the learner's text in terms of cohesion and the organisation of the text, more particularly the first paragraph of the text.

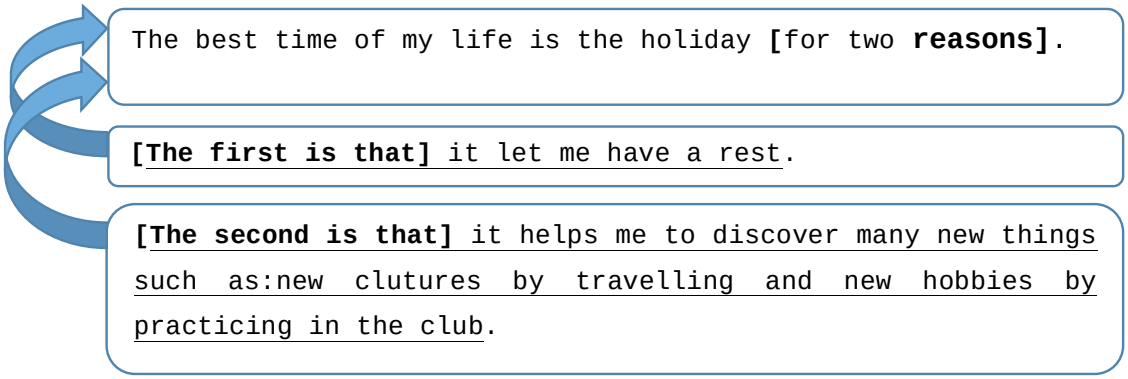


Figure 4-3 *Cohesive links between the first three sentences taken from the suggested version of the learner's text as shown in example (7)*

A number of instances where the use of SNs could have improved the level of the learner's essay in example (7) can be found in the second paragraph. One of these instances can be seen in the following: "In my last summer holiday, I travelled to India. The crowdest country.". Although the second sentence here is incomplete, it is more likely to be understood as an evaluation of some part of the previous sentence. However, this evaluation could have been signaled better, besides the structural changes that should have been made to the sentence, by the use of, for example, the SN *opinion* as shown in the following: "[In my **opinion**,] [India is] the crowdest country.". By using the SN *opinion*, the learner could have incorporated his stance towards the information he/she is providing and thus eliminating the questions that the readers of his/her

text might have in minds: is the information provided a fact?, is it just the opinion of the author?, or is it the opinion of someone else that the writer of the text might have failed to reference? Thus, the use of the SN in such a sentence can be said to have effectively improved the quality of the learner's writing.

The learner's second paragraph could be also made better by adding a sentence with two instances of SNs: "[Inspite of this ***drawback***, it is possible to have good ***experiences*** there.]". By using the SN *drawback*, the learner could have assessed India being crowded as a bad thing, a point that the learner had made in his/her previous sentence. Thus, a link between the two sentences could have been established by the use of the noun. The use of the SN *experiences* could have helped in signaling what follows in the learner's paragraph.

The final SN that could have been added in the learner's essay is *conclusion*. The instance where this SN appears to be necessary can be found in the beginning of the learner's final paragraph. The learner begins this paragraph with a sentence assessing his/her holiday. This sentence could be seen as an abrupt sentence, considering the fact that the learner in the previous paragraph recounts his/her holiday experiences. Thus, it would have been better if the learner had signalled the sentence properly. One way of doing so can be achieved by the use of the SN *conclusion* which signals that the writer is no longer recounting his/her experiences but is rather giving his conclusion based on what has been mentioned previously as shown in the example below:

[In **conclusion,**] I think it was my best holiday. It helped me to learn the Indiain cluture. You should really visit India.

The SN *conclusion* in the above example is used in its fixed phrase pattern. Such use may seem less difficult for the learners as the SN is not integrated into the sentence. Therefore, the learner could have made his writing more sophisticated by integrating the SN *conclusion* into the sentence. Doing so would involve adjusting the sentence as shown in the following: [Based on the above experiences, my **conclusion** is that] it was my best holiday.

Example (8) below provides the second whole essay selected for deeper scrutiny as mentioned previously. In this example, there are instances where the incorporation of SNs to the essay could have rendered a more effective essay. The possible SNs that could have been used by the learner are *holiday* and *problems*.

Example (8):

The learner's text:
Last year I had my worst summer holiday. I traveled to Paris. I went there with my father ,mother, sister and brother. We traveld by plane to Paris. Firstly, when we arrived Paris the streets were so crowded that we couldn't move easily and that took us a lot of time to reach the hotel. Moreover, some of our bags were lost in the airport so we should buy new clothes. Secondly, while I was shopping in one of Paris's shopping Center someone stole my bag I tried to run back him but I couldn't reach him because he was too fast . In addition to this my brother got lost in the garden. My father was afraid and he rushed to find him. Luckily, we found him crying on the chair. Finally, when we went to the airport for departurring the plane had taken of already, so we should wait for the other plane six hours more. We were so tierd in this holiday with a lot of bad things which happend to us. At the end I can say that it was the worst holiday I have ever spent in my life.
A suggested revision of the learner's text with the addition of SNs:

Last year I had my worst summer holiday. I traveled to Paris. I went there with my father, mother, sister and brother. We traveld by plane to Paris. [This **holiday** was the worst because we were faced up with many **problems**.]
Firstly, when we arrived Paris the streets were so crowded that we couldn't move easily and that took us a lot of time to reach the hotel. Moreover, some of our bags were lost in the airport so we had to buy new clothes. [Thirdly], while I was shopping in one of Paris's shopping Center someone stole my bag I tried to run back him but I couldn't reach him because he was too fast .
In addition to [these **problems/** dreadful **events**,] my brother got lost in a garden in paris. My father was afraid and he rushed to find him. Luckily, we found him crying on a chair in the garden. Finally, when we went to the airport for travelling back to our country, the plane had taken of already. So, we [had to] wait for the other plane six hours more. We were so tierd in this holiday with a lot of bad things which happend to us. At the end, I can say that it was the worst holiday I have ever spent in my life.

As shown in the suggested version of the learners' essay in example (8), the use of SNs could have made the learner's writing better. The addition of two SNs in the end of the first paragraph of the student's text may be said to have improved the learner's text in terms of cohesion and the organisation of the text. The use of the SN *holiday* links up with a previously stated proposition by means of encapsulation whereas the use of the SN *problems* prospects that the ensuing discourse is to present the problems that the learner encountered in his/her holiday. In the subsequent paragraphs, the learner presents the problems he/she encountered in his/her holiday. These problems are underlined as shown in the suggested text. The underlined propositions can be said to link up with the suggested SN *problems* in the first paragraph.

Another instance where an SN should have been used by the learner can be found in the first sentence of the third paragraph of the learner's text: "`In addition to this my brother got lost in the garden.`". The unattended "`this`" in this quote can be said to create ambiguity in the text because it is hard to locate its reference. One way of rectifying this issue is to follow the demonstrative with an SN. It is suggested that the learner could have used the following: "`In addition to [these` **`problems/`** `dreadful` **`events`**`,] my brother got lost in a garden in paris.`". By using either of the SNs *problem*/ *events* in the suggested pattern: *these problems*/ *dreadful events*, these SNs encapsulate the previously mentioned problems/ events and at the same time signal that what is coming next in the text is another problem/event. Thus, if the learner had used either of the SNs: *problems*/ *events*, his/her writing could have been more organised and cohesive.

The use of the SN *problems* could have also enabled the learner to incorporate his/her stance on the propositions he/she is talking about. If the learner had used the SN *problems* instead of the SN *events* in his/her essay, as shown in the suggested text of example (8), the learner would have enhanced the effectiveness of his/her writing in terms of stance by going beyond a mere presentation of experiences or propositions to the expression of personal attitudes towards these propositions. The following examples provide illustrative excerpts of cases where the addition of SNs seems to be enhancing the quality of the students' texts. These excerpts are taken from the other essays chosen for deeper qualitative scrutiny.

Example (9):

The original excerpt	The revised excerpt
After few miuntes they took her to the hospital. The Doctors checked the baby then he told us that the new baby born is sick. We were sad for this. Finaly my father determinde to return to the UAE.	After few miuntes they took her to the hospital. The Doctors checked the baby then he told us that the new baby born is sick. We were sad for this [worrying/ distressing **news**]. Finaly my father determinde to return to the UAE.

Example (9) above shows an instance where the use of an SN could have helped the learner to add more meanings (e.g. clarity, stance and organisation) to his/her writing. As shown in the original excerpt, the ambiguity of reference created by the unattended demonstrative in the third sentence of the excerpt could have been rectified by the use of an SN. It is suggested that the demonstrative could be followed by the following SN: *news*. This SN would have helped the learner to encapsulate part of the previous sentence and thus would have resulted in a more clear, organised and cohesive text. It could have been possible for the learner to use another SN instead of the SN *news*. One of the options is the SN 'thunderbolt'. However, to use this noun, the whole sentence would need to be adjusted. So, instead of writing "we were sad for this" the learner could have written the following: "We were [shocked by this thunderbolt]". By using this SN, the learner could have expressed his stance towards the news.

Example (10) provides two excerpts taken from two different essays. These excerpts are put together in the example for the reason that they are similar in that the two excerpts could have been written better if the writers of these excerpts had used SNs expressing causality.

Example (10):

The original excerpts:	The revised excerpts:
(1): About my working in Dubai Municipality as a trainer, I have enjoyed that experience so much becoose I have learned a lot about the job field and how it will be in the future and also I have made many strong relationships with the other trainers and this is the most important thing that we should concentrate on.	**(1):** About my working in Dubai Municipality as a trainer, I have enjoyed that experience so much [for two **reasons**]. [Firstly,] I have learned a lot about the job field and how it will be in the future. [Secondly,] I have made many strong relationships with the other trainers. This is the most important thing that we should concentrate on.
(2): In fact this computer is the bist thing I got, it means a lot to me, I gotit last summer, when I took my ICDL exam, also I got it because I needed a new computer to do my tesks. Adding to that I had many problems with the old computer, I think I really dissorve this gift. I love this computer because it is the bist thing I have ever got, not to forget it's a gift from my parents and I love it more the other possession because I had got it after the most difficul exam I ever had the ICDL exams. I stronglly belive that a gift comes from people you love is the most expinseve possession you have.	(**2):** In fact, this computer is the best thing I got. [For this **reason**/ as a **result**,]it means a lot to me. I got my computer last summer when I took my ICDL exam, also I got it because I needed a new computer to do my school tesks. [In **addition**,] I had many problems with my old computer. I think I really deserve this gift. I love my computer [for two **reasons**.] [The first **reason**] is that it is the bist thing I have ever got [due to the **fact** that] it's a gift from my parents. [Secondly,] I got it after the most difficult ICDL exam I have ever had. [In **conclusion**,] I strongly believe that a gift coming from people you love is the most expensive possession you have.

As illustrated in the revised excerpts in example (10), the proper use of SNs expressing causal meanings can be said to have transformed the learners' writing from being immature and less effective into being more sophisticated. For example, the learners' writing in excerpt one and the second paragraph of excerpt 2 in example (10) would have been written more effectively if the learners had used the SN *reasons* instead of the conjunct *because* for the following reason. The proper use of *reasons* could have been one way to resolve the issue of having a number of causal meanings being compressed in one long run-on sentence. As shown in the revised version of the excerpts in question, the use of *reasons* can be said to help in presenting these causal meanings in a more organised and mature way.

The first paragraph of the second excerpt of example (10) shows two instances where SNs could have been used to improve the quality of the paragraph. One case is related to the proper signalling of causal meanings. This can be observed in the following statement: "In fact this computer is the bist thing I got, it means a lot to me,". It is suggested that the link between the two clauses should have been made explicit. The explicitness in expressing causal meanings in the quoted statement can be achieved by the use of either the SN *reason* or the SN *result* as illustrated in the revised sentence below. In fact, this computer is the best thing I got. [For this **reason**/ as a **result**,] it means a lot to me.

By doing so, the two sentences are linked effectively and the causal relation being expressed is appropriately signalled.

The other case in the paragraph is related to the proper signalling of additional information to be found in the text. The following is illustrative of this instance:

also I got it because I needed a new computer to do my tesks. Adding to that I had many problems with the old computer,

The use of "Adding to that" to signal the writer's additional reason is inappropriate. One way to rectify this can be made possible by the use of the SN *addition*, as illustrated below:

,also I got it because I needed a new computer to do my school tesks. In **addition**, I had many problems with my old computer.

The qualitative analysis of the use of SNs in the selected examples discussed in this section suggests that the proficiency level of writers is linked to the significant difference between the learner and native speaker corpora. I now turn to the comparison of the two corpora in terms of which SN types are used.

4.5.2. SN types

Table 4-9 illustrates the difference between the learner corpus and the native speaker corpus used in the current study in terms of selection of SNs. The table lists all identified SN types in descending order of normalised frequency. Nouns that are found in one corpus only are in **bold** and underlined nouns represent nouns that appear in the two corpora.

Table 4-9 *List of SNs in the two corpora shared and unshared*

Rank	SNs in the learner corpus	Normalised frequency per 100,000 words	Rank	SNs in the native speaker corpus	Normalised frequency per 100,000 words
1	**Holiday**	239	1	Time	251
2	Place	148	2	Way	118
3	Time	144	3	Reason	117

Table 4-9 *(Continued)*

Rank	SNs in the learner corpus	Normalised frequency per 100,000 words	Rank	SNs in the native speaker corpus	Normalised frequency per 100,000 words
4	Thing	127	4	**Ban**	95
5	**Trip**	53	5	Thing	90
6	End	43	6	**Right**	71
7	**Story**	41	7	Idea	66
8	Reason	35	8	Opinion	60
9	Problem	27	9	Fact	54
10	**Vacation**	26	10	Example	53
11	Kind	23	11	**Issue**	52
12	Addition	20	12	**Case**	51
13	Way	19	13	Experience	46
14	Idea	19	14	Part	45
15	Opinion	19	15	**Statement**	44
16	**Moment**	17	16	**Choice**	43
17	Example	16	17	**Skill**	42
18	Part	16	18	Job	36
19	Activity	14	19	Place	36
20	**Group**	14	20	Problem	36
21	Experience	14	21	**Opportunity**	34
22	Chance	13	22	**Habit**	31
23	Fact	13	23	Kind	29
24	**Journey**	12	24	**Result**	29
25	**Conclusion**	11	25	Chance	28
26	**View**	10	26	**Freedom**	24
27	**Dream**	9	27	Point	24
28	**Feeling**	9	28	Addition	23
29	**News**	9	29	**Solution**	23
30	**Rest**	8	30	**Variety**	23
31	**Event**	6	31	**Benefit**	20
32	**Plan**	6	32	**Sense**	20
33	**Death**	6	33	**Majority**	19
34	**Word**	5	34	Matter	19
35	**Accident**	5	35	**Need**	19
36	Job	5	36	**Situation**	19
37	Point	5	37	End	18
38	**Information**	4	38	**Position**	16
39	**Favourite**	4	39	Activity	14
40	**Minute**	4	40	**Responsibility**	14
41	**Change**	3	41	**Smell**	14
42	Hope	3	42	**Cost**	13
43	Matter	3	43	**Risk**	13
44	**Nature**	2	44	**Effect**	10
45	Use	2	45	**Business**	10
46	Means	2	46	**Cause**	10
47	**Advice**	1	47	**Pressure**	6

Table 4-9 *(Continued)*

Rank	SNs in the learner corpus	Normalised frequency per 100,000 words	Rank	SNs in the native speaker corpus	Normalised frequency per 100,000 words
48	Thought	1	48	Use	6
			49	Thought	5
			50	**Deal**	4
			51	**Stress**	3
			52	**Culture**	2
			53	Hope	2
			54	Means	1

As illustrated by Table 4-9, two SNs: *time(s)* and *thing(s)* are among the top-5 SNs in the two lists. Nouns such as *reason(s)*, *way(s)*, *opinion(s)* and *idea(s)* are found to occur among the top-15 SNs that appear in the two lists. Therefore, these nouns will be investigated more closely in Chapter 5 (e.g. *time(s)*, *thing(s)* and *way(s)*), Chapter 6 (e.g. *opinion(s)* and *idea(s)*) and Chapter 7 (e.g. *reason(s)*).

Table 4-9 also shows a great deal of variation with regard to the SN types that are used in the two corpora. *Holiday(s)*, *trip(s)*, *story(s)*, *vacation(s)*, *moment(s)* are illustrative examples of SNs that are found within the top-20 SNs in the learner corpus and are not found in the native speaker corpus at all; *ban(s)*, *right(s)*, *issue(s)*, *case(s)*, *statement(s)*, *choice(s)* and *skill(s)* are among the top-20 SNs found only in the native speaker corpus. As discussed above, some of these frequently used SNs are topic-related.

In addition, it can be observed that the learner corpus stands out in terms of the use of SNs that are on the borderline between abstract and concrete nouns. The SNs with borderline status are used more often and found among the top-10 SNs. These nouns are *holiday(s)*, *trip(s)* and *vacation(s)*. This is not the case of the native speaker corpus where all SNs are clearly identifiable as abstract nouns.

Perhaps the most interesting observation is that there are more SNs that express stance and causality in the native speaker corpus compared with the learner corpus. That is demonstrated by the types and total frequency of nouns expressing stance and causality found to occur in the two corpora. Illustrative examples of SNs with stance function are *right, idea, fact* and *opinion*. Chapter 6 will be dedicated to this this type of noun in more detail. Illustrative examples of nouns signalling causality are *reason*, *cause* and *effect*. Chapter 7 will investigate this type of noun.

Having looked at Table 4-9 and having commented on the difference between the learner and the native speaker corpora, an interesting question has arisen and is indeed worthy of further investigation. The question is **what if** only the more advanced learners (CEPA 5 and CEPA 6) are included in the comparison; would one get the same differences? Figure 4-2 presented earlier shows that there is a substantial internal variation between all the levels comprising the learner corpus. In addition, it shows that there is a levelling out in both tokens and types of SNs used in CEPA 5 and CEPA 6. In the light of this finding, it would be more useful to only compare the native speaker corpus with the advanced learners (CEPA 5 and CEPA 6). Table 4-9 presented earlier is reproduced in Table 4-10 below to only include CEPA 5 and CEPA 6. It is important to note that the size of the two sub-corpora of the advanced learners only is considered when producing the average normalised frequency of SNs found to occur in these sub-corpora.

Table 4-10 *SNs as used in the highest-level learners (CEPA 5 and CEPA 6) and the native speaker corpus*

Rank	SNs in CEPA 5 and CEPA 6	Average normalised frequency per 100,000 words in the two sub-corpora	Rank	SNs in the native speaker corpus	Normalised frequency per 100,000 words
1	**Holiday**	140.69	1	Time	251
2	Time	85.33	2	Way	118
3	Place	72.84	3	Reason	117
4	Thing	71.59	4	**Ban**	95
5	**Trip**	39.54	5	Thing	90
6	Reason	25.80	6	**Right**	71
7	**Story**	24.14	7	Idea	66
8	End	23.72	8	Opinion	60
9	**Vacation**	21.22	9	Fact	54
10	Problem	16.23	10	Example	53
11	Way	14.56	11	**Issue**	52
12	Kind	14.15	12	**Case**	51
13	**Moment**	14.15	13	Experience	46
14	Addition	13.32	14	Part	45
15	Experience	12.07	15	**Statement**	44
16	**View**	7.07	16	**Choice**	43
17	**Chance**	11.23	17	**Skill**	42
18	**Group**	11.65	18	Job	36
19	Idea	11.23	19	Place	36
20	Fact	9.57	20	Problem	36
21	Part	10.40	21	**Opportunity**	34
22	**Journey**	8.74	22	**Habit**	31
23	Opinion	8.32	23	Kind	29
24	Example	7.07	24	**Result**	29
25	**Conclusion**	7.07	25	Chance	28
26	**Dream**	6.66	26	**Freedom**	24
27	**News**	6.24	27	Point	24
28	**Rest**	5.82	28	Addition	23
29	**Event**	5.41	29	**Solution**	23
30	Activity	5.41	30	**Variety**	23
31	**Feeling**	4.99	31	**Benefit**	20
32	**Word**	4.57	32	**Sense**	20
33	**Plan**	4.16	33	**Majority**	19
34	Job	4.16	34	Matter	19
35	**Favourite**	3.74	35	**Need**	19
36	**Minute**	3.74	36	**Situation**	19
37	**Death**	3.33	37	End	18
38	**Accident**	3.33	38	**Position**	16
39	Point	3.33	39	Activity	14

Table 4-10 *(Continued)*

Rank	SNs in CEPA 5 and CEPA 6	Average normalised frequency per 100,000 words in the two sub-corpora	Rank	SNs in the native speaker corpus	Normalised frequency per 100,000 words
40	**Change**	2.49	40	**Responsibility**	14
41	**Information**	2.08	41	**Smell**	14
42	Use	2.07	42	**Cost**	13
43	Matter	1.66	43	**Risk**	13
44	Means	1.66	44	**Effect**	10
45	**Nature**	1.66	45	**Business**	10
46	**Advice**	1.24	46	**Cause**	10
47	Hope	1.24	47	**Pressure**	6
48	Thought	1	48	Use	6
			49	Thought	5
			50	**Deal**	4
			51	**Stress**	3
			52	**Culture**	2
			53	Hope	2
			54	Means	1

Table 4-10 shows that despite the restriction of the learner corpus to include only the advanced level students the results obtained in terms of the differences between the two corpora are rather similar to the previous results obtained with the presence of lower level students. For instance, the first thirteen SNs in the list of SNs presented in Table 4-10 are the same as the first thirteen SNs in the list of SNs that occur in the learner corpus overall. This shows that the previous results were not skewed because of the inclusion of the lower level learners.

4.6. Summary of the findings and concluding remarks

The analysis of the frequency of SNs in the learner and the native speaker corpora has yielded some interesting findings. One of these findings is that the choice of SNs is often topic sensitive. The learner and the native speaker corpora are similar in that both corpora

showcase instances of SNs being influenced by the specific topic being addressed. The quantitative analysis of SNs across the four proficiency levels in the learner corpus suggests that there is an increase in the use of SNs as the learners move up in the proficiency scale.

The quantitative analysis of both tokens and types of SNs found to occur in the learner and native speaker corpora has revealed that the two corpora differ remarkably in many respects. One main finding is that the learner corpus contains fewer SN tokens and types; the difference is found to be statistically significant, more particularly with regard to SN tokens.

It is suggested that the significant difference in the use of SNs between the learner and the native speaker corpora is linked to proficiency. This has been demonstrated by the quantitative and qualitative analysis of the use of SNs.

The findings in this chapter will be compared with related literature in Chapter 8. It can be concluded that this chapter has identified the most important aspects in comparing the two corpora. Perhaps, one of the most important aspects of difference between the two corpora that has been identified in this chapter is the use of SNs that express stance and causality in discourse. This finding will be further addressed in more detail in the following Chapters: Chapter 6 and Chapter 7. The following chapter (Chapter 5) will examine in more detail the uses of three most frequent and shared SNs in the two corpora.

CHAPTER 5: CLOSE ANALYSIS OF THREE MOST FREQUENT AND SHARED SNs IN THE LEARNER AND NATIVE SPEAKER CORPORA

5.1. Overview

This chapter provides the results obtained from a closer investigation of the three most frequent and shared SNs in the learner and the native speaker corpora. These SNs are *time(s)*, *thing(s)*, and *way(s)*. The aim of this chapter is to give an account of the uses of these SNs in the two corpora. More specifically, the focus is on the specification of these SNs in terms of: (1) the reference category that the specification of these SNs might fall into, (2) the grammatical forms of the specification of the nouns, and (3) the meanings in context of the specification of these SNs. In the case of *thing(s)*, a discussion of pattern replaces the account of meaning. Section 5.2 provides a brief overview of both the reference categories identified for the SNs in question and the meanings in context associated with these SNs. The subsequent sections of Section 5.2 (5.3, 5.4, and 5.5) present the results for the SNs *time*, *thing*, and *way* respectively. A summary of the findings of this chapter is presented in Section 5.6.

5.2. Reference categories and textual meanings identified: a brief overview

Following Flowerdew and Forest (2015), the two main specification categories of an SN are: (1) within clause specification (WCS) and (2) across clause specification (ACS). The direction of reference between an SN and its specification in these two categories can be either anaphoric or cataphoric. The reference is established as anaphoric when the specification precedes the SN, and as cataphoric when the specification follows the SN. The following examples, (1) to (4), are taken from the learner corpus to illustrate the two

categories of an SN specification along with the possible reference relations. Anaphoric within-clause specification (ANWCS) is presented in (1), cataphoric within-clause specification (CAWCS) in (2), anaphoric across-clause specification (ANACS) in (3), and cataphoric across-clause specification (CAACS) in (4):

(1): I think that traveling is the best **thing** to do in Holidays... **(CEPA6)** (ANWCS)

(2): The best **time** of my life is the holiday. **(CEPA 5)** (CAWCS)

(3): During holidays, we travel abroad to have some fun. It's a **way** to break the routine and have a break from work. **(CEPA6)** (ANACS)

(4): At the morning my little cousin woke me up in a really annoying **way**. She put my uncle's monkey on my bed and then left the room. **(CEPA5)** (CAACS)

As illustrated in the above examples, the specification of SNs may take a form of a noun phrase as in (1) and (2) or a clause as in (3) and (4). The specification of the SNs in the above examples can be said to have specific meanings in context. For example, the meanings associated with specification of the SN *way*, as shown in (3) and (4), are not the same; the specification of *way* can be interpreted as a method in (3) and as a manner in (4). The most likely paraphrase of *way* in (3) is "It's a **method** to break the routine and have a break from work." while the most likely paraphrase of *way* in (4) is "At the morning my little cousin woke me up in a really annoying **manner**.".

In this section, I have explained the reference categories that a specification of SN might fall into and have touched briefly on the idea of textual meanings that can be associated with particular SNs; the SN *way* is a case in point. The next sections (5.3, 5.4, and 5.5) will address these two aspects in more detail. The aim is to compare the learners with the native speakers in terms of where to find the specification of the three SNs in question, how the specification is realised grammatically and the possible meanings associated with these nouns.

5.3. *Time(s)*

Time(s) is the most frequent and shared SN in the learner and native speaker corpora. It ranks third in the learner corpus and first in the native speaker corpus. The following two sections provide the results of how the SN *time(s)* is used in the two corpora in terms of the reference categories (5.3.1) and meanings in context identified (5.3.2).

5.3.1. Reference categories of *time(s)*

Table 5-1 provides the raw frequency of the specification categories of *time(s)* in the two corpora. The frequency is supplemented with the percentage of each category to the total number of all instances of the SN *time(s)* in each corpus.

Table 5-1 *Reference categories of the SN time(s) across the two corpora*

	Raw frequency in the learner corpus (%)		**Raw frequency in the native speaker corpus (%)**	
Within clause specification (WCS):	267 (88%)		217 (95%)	
Cataphoric and anaphoric	217 (81%)	50 (19%)	179 (82%)	38 (18%)
Across clause specification (ACS):	38 (12%)		11(5%)	
Cataphoric and anaphoric	0 (0%)	38 (100%)	0 (0%)	11 (100%)
Total frequency	305 (100%)		228 (100%)	

As illustrated in Table 5-1, *time(s)* shows a strong preference towards within clause specification than across clause specification. The raw frequencies and the percentages of this category of specification in the two corpora are indicative of this; 267 (88%) of all the instances of the SN *time(s)* in the learner corpus and 217 (95%) of the 228 occurrences of the SN in the native speaker corpus are found to be specified within clause boundaries. The difference between the two corpora in terms of the frequency of the within clause specification of the SN *time(s)* is not significant (The log-likelihood result is – 0.83). However, the difference of the instances of across clause specification of the noun in the two corpora is found to be statistically significant (The log-likelihood result is + 8.92).

A closer look at the directionality of reference of within clause and across clause specification of *time(s)*, as displayed in Table 5-1, also reveals two interesting findings. The first is concerned with the within clause specification. Table 5-1 demonstrates that there is a strong preference for cataphoric specification than anaphoric specification when the SN *time(s)* is specified within clause environment. The instances of cataphoric within clause specification are 217 (81%) in the learner corpus and 179 (82%) in the native speaker corpus. The following examples, (4) to (7), are illustrative cases of within clause specification in the two corpora; the reference is established as cataphoric in (4) and (5) and as anaphoric in (6) and (7).

(4): The best **time** for me in Al Kabah is the few hours before sunrises because... **(CEPA5)** (CAWCS)

(5): By the **time** students are old enough to go to college, it's my opinion that they're also old enough to understand the value of money,... **(Native speaker)** (CAWCS)

(6): Holiday is the most important **time** in our life,... **(CEPA 4)** (ANWCS)

(7): Secondly, College is an important **time** in the social and sexual growth an individual,... **(Native speaker)** (ANWCS)

The second finding is related to the instances of across clause specification of *time(s)*. As illustrated in Table 5-1, *time(s)* is always specified anaphorically in the learner and native speaker corpora. This suggests that the SN *time(s)* is less likely to refer to a stretch of language that may be present across clauses in the discourse to come. Examples (8) to (12) provide instances of anaphoric across clause specifics of the SN *time(s)* in the two corpora:

(8): However, the situation with respect to restaurants is more complicated, as a meal may last an hour or two, so one is committed to that **time**,... **(Native speaker)**

(9): All the people were staring at us. I got shy and mad that **time**. **(CEPA5)**

(10): When the weather started to be better we heard about the death of a dear person. I was sad at that **time**, I didn't stop crying. **(CEPA6)**

(11): I recently came back from a study abroad trip in China where I was there for one year. In this **time**, I saw that many restaurants in China had already banned smoking at restaurants and... **(Native speaker)**

(12): in the third time, she told him that she don`t know how to swim, he didn`t believe her, but she was said the truth this **time**, so... **(CEPA4)**

In all the examples above with the exception of example (8), the SN *time* is specified by clauses. Table 5-2 shows the frequency of the three grammatical ways with which the SN *time(s)* can be specified.

Table 5-2 *Specification of the SN time(s) in the learner and native speaker corpora*

Grammatical form	**Raw frequency in the learner corpus (%)**	**Raw frequency in the native speaker corpus (%)**
Clause	186 (61%)	152 (67%)
Prepositional phrase	72 (24%)	47 (20%)
Simple noun phrase	47 (15%)	29 (13%)
Total	305 (100%)	228 (100%)

As can be seen in Table 5-2, the SN *time(s)* is frequently specified in clauses in the two corpora. In the learner corpus, 61% of all instances of *time(s)* are specified in clauses. The corresponding figure in the native speaker corpus is 67%. Illustrative examples of *time(s)* specified as clauses are found in (13) and (14) presented below. The other two types of specification: prepositional phrase and simple noun phrase are less represented in the two corpora, as displayed in Table 5-2. The prepositional phrase and simple noun phrase providing the specifics of the SN *time(s)* are shown respectively in (15) and (16):

(13): The first **time** when I visit this country I was surprised because… **(CEPA5)**

(14): I think that is was the most fantastic holiday because it was the first **time** I get to travel with my father and sisters by ourselves and... **(CEPA6)**

(15): I use my imagination and give myself **time** for meditation. **(CEPA6)**

(16): Holiday is the best **time** for me. **(CEPA4)**

The above examples have illustrated the three ways with which the specifics of the SN *time(s)* can occur grammatically. The next section investigates the possible meanings that the specification of the SN *time(s)* can be classified into.

5.3.2. Meanings associated with the SN *time(s)*

A closer examination of the stretch of language providing the specification of *time(s)* in the learner and native speaker corpora has shown that there are three meanings, which can be associated with the SN *time(s)*. Figure 5-1 illustrates these meanings:

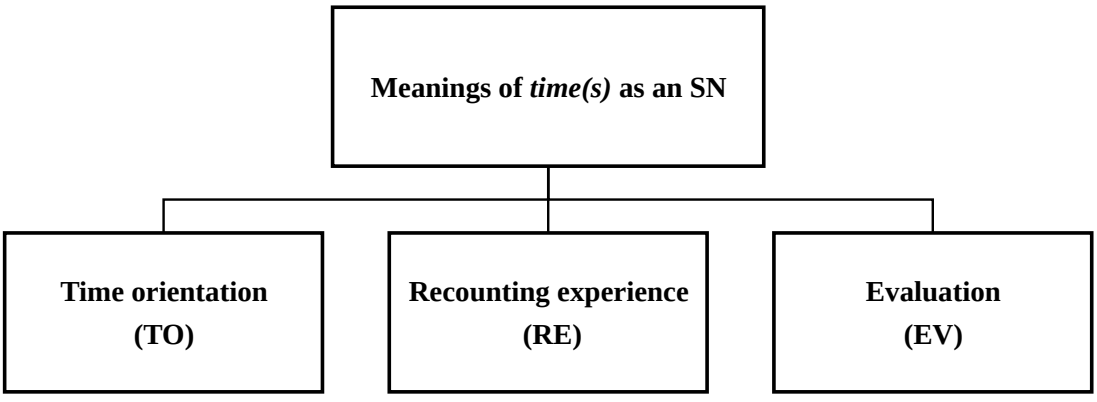


Figure 5-1 *Meanings associated with the specification of the SN time(s) in the learner and native speaker corpora*

The meaning of the SN *time(s)* is identified as a *time orientation* (TO) when its specification marks a specific point in time. When the specification spells out the general time when something happens or someone does something, *time(s)* is assigned to the *recounting experience* (RE) meaning. *Evaluation* (EV) is the third meaning associated with *time(s)* as an SN; this meaning is identified when the noun itself is used to express a sense of evaluation; either as a recommendation for desired actions or other evaluative senses. The following sets of examples are illustrative of the three meanings identified with the SN *time(s)* in the learner and native speaker corpora:

***Time orientation (TO)*: (Specific time)**

(17): Our departure **time** was at the early morning and… **(CEPA5)**

(18): The best **time** for me in Al Kabah is the few hours before sunrises because… **(CEPA5)**

(19): Holiday is the best **time** for every one to do whatever they wont. **(CEPA4)**

(20): Last month was the worst **time** of the year. **(CEPA6)**

(21): College is one of the most important **times** in many young people's lives. **(Native speaker)**

(22): Some of the best **times** to enjoy a cigarette were after a nice dinner at a restaurant and... **(Native speaker)**

(23): The amount of work that students have to do to graduate is more than enough to keep them busy during their **time** at college. **(Native speaker)**

***Recounting experience (RE)*: (General time)**

(24): One **time** we went at night to night Zoo. **(CEPA4)**

(25): It was the first **time** I had been to Belgium and France, **(CEPA6)**

(26): This rifle is very precious to me and every **time** I look at it I feel proud of my great Grandfather **(CEPA6)**

(27): Besides,my sister has final exams education and she didn't have enough **time** to talk or stay with me . **(CEPA5)**

(28): that it is sometimes difficult to find **time** to get proper rest, or time to eat properly. **(Native speaker)**

(29): ...and for the first **time** in a long time, I haven't been throwing my money away as much as I used to. **(Native speaker)**

(30): One **time** I even just had a stupid Marlboro hat, and... **(Native speaker)**

(31): ...without feeling the need to order an oxygen mask every **time** I go to a restaurant. **(Native speaker)**

***Evaluation (EV)*: (Recommendation and other senses)**

(32): I think that it is **time** Japan got off there behinds and became more proactive in their efforts. **(Native speaker)**

(33): I have had my books in **times** of bordem and in **times** of need when I thought my head would burst from studying. **(CEPA6)**

(34): The time for discussion is over and it is now **time** for serious action irrespective of the feeble arguments put forward by the smoking community. **(Native speaker)**

(35): Even though smoking has been a historical part of many cultures worldwide for many years, it is **time** for change. **(Native speaker)**

As shown in the examples of the SN *time(s)* in the *time orientation* set, the specification of the SN spells out specific points in time, as illustrated in the underlined stretch of discourse of these examples. The examples in the *recounting experience* group show cases where the specification of *time(s)* does not pinpoint a specific time as in the *time orientation* group but rather spells out occasions when something happens or someone does something. The specification of *time(s)* in the *evaluation* group provides the evaluated entity.

A closer look at all the instances of *time(s)* in each of the three meaning categories, as shown in the three sets above, has revealed that there is a link between the meaning identified and the grammatical form by which the SN is specified. When the SN is found to occur in the *time orientation* group, the specification of the noun can be in a form of a noun phrase as in (19) and (20), and a prepositional phrase as in (17) and (22). However, when the SN is found to occur in the *recounting experience* group, the specification of the noun is always found in a form of a clause as in (24) and (25). The specification of the SN in the *evaluation* category

is found to take two grammatical forms: a clause as in (32) and a prepositional phrase as in (34) and (35).

Table 5-3, below, provides the raw frequency of the three meaning groups of the SN *time(s)* as found to occur in the learner and native speaker corpora. As illustrated in Table 5-3, *time(s)* in the two corpora is used more frequently to serve two meanings: *time orientation* and *recounting experience*. The less frequently used meaning in the two corpora is that of *evaluation*; 6 (2%) of all the instances of the SN *time(s)* in the learner corpus and 13 (6%) of all the occurrences of the SN in the native speaker corpus are found to occur in this meaning. Therefore, it can be said that there are no differences between the learner and native speaker corpora in the use of the three meanings of the SN *time(s)*.

Table 5-3 *Raw frequency of the three meanings associated with the SN time(s) across the two corpora*

Meanings associated with the SN *Time(s)*	**Raw frequency in the learner corpus**	**Raw frequency in the native speaker corpus**
Time orientation (TO)	154 (50%)	114 (50%)
Recounting experience (RE)	145 (48%)	101 (44%)
Evaluation (EV)	6 (2%)	13 (6%)
Total frequency	305 (100%)	228 (100%)

5.4. *Thing(s)*

Thing(s) is the second most frequent and shared SN in the learner and native speaker corpora. It ranks fourth in the learner corpus and fifth in the native speaker corpus. Table 5-4, below, presents the raw and normalised frequencies of the SN *thing(s)* in the two corpora.

Table 5-4 *Frequency of thing(s) across the two corpora*

Frequency of *thing(s)*	**Learner corpus**	**Native speaker corpus**
Raw	268	86
Normalised per 100,000 words	127	95

As displayed in Table 5-4, the normalised frequency of *thing(s)* per 100,000 words in the learner corpus is 127. The corresponding figure in the native speaker corpus is 95. The difference in the frequency of *thing(s)* in the two corpora is significant. The log-likelihood result is + 5.76[7] with the probability value of $p<0.05$. The SN *thing(s)* is over-represented in the learner corpus. The following two sections provide the results of how the SN *thing(s)* is used in the learner and native speaker corpora in terms of the reference categories (5.4.1) and the uses of the SN *thing(s)* (5.4.2).

5.4.1. Reference categories of *thing(s)*

The raw frequency of the specification categories of *thing(s)* in the learner and native speaker corpora is given in Table 5-5. The raw frequency is supplemented with the percentage of each category to the total number of all instances of the SN *thing(s)* in each corpus.

[7] I have used the log-likelihood calculator created by Rayson (2017). For more clarification on how Rayson's calculator works and how to interpret obtained results, refer to Section 4.5 in Chapter 4.

Table 5-5 *Reference categories of thing(s) across the learner and native speaker corpora*

	Raw frequency in the learner corpus (%)		**Raw frequency in the native speaker corpus (%)**	
Within clause specification (WCS):	184 (69%)		55 (64%)	
Cataphoric and anaphoric	162 (88%)	22 (12%)	43 (78%)	12 (22%)
Across clause specification (ACS):	84 (31%)		31(36%)	
Cataphoric and anaphoric	47 (56%)	37 (44%)	12 (39%)	19 (61%)
Total frequency	268 (100%)		86 (100%)	

There is no significant difference between the two corpora in terms of the frequency of the two main reference categories of the specification of the SN *thing(s)*[8]. The log-likelihood result for the difference is + 0.22 for the WCS category and – 0.43 for the ACS category. Although the difference is not significant, the WCS is found to be slightly over-represented in the learner corpus and the ACS is found to be slightly under-represented[9]. When an SN is assigned the category of WCS, the SN itself and its specification are both contained in the same sentence. This indicates that learners show less preference towards the use of the SN *thing(s)* to organise a stretch of discourse where the specification of *thing(s)* spans beyond clause boundaries. One possible reason for the relative infrequency of *thing(s)* used to organise text is that the learners are more prone to think in terms of sentences only rather than longer stretches of discourse. A similar observation is also true for the other two nouns in this chapter. The following examples, (36) to (43), are taken from the learner and the native

[8] The log-likelihood test was run on the raw frequency of each reference category in the two corpora in proportion to the raw frequency of all instances of each category in the two corpora.

[9] Explanation of how the results obtained from the log-likelihood test is to be interpreted is explained in Section 4.5 in Chapter 4.

speaker corpora to show how the SN *thing(s)* is specified in discourse either within or across clause boundaries.

(36): The most important **thing** is to ban smoking at all public places. **(Native speaker)** (CAWCS)

(37): We rarely see each other outside Dubai, so we thought the best **thing** to do is to go out together. **(CEPA6)** (CAWCS)

(38): I think that smoking is a bad **thing**. **(Native speaker)** (ANWCS)

(39): For these reason My computer is the best **thing** for me. **(CEPA4)** (ANWCS)

(40): and it has given me a variety of good **things**. For example, I have been able to make a little bit of money for the weekends and other times. **(Native speaker)** (CAACS)

(41): I have learned many **things**, the most important is that not to rely on anybody to do my work. **(CEPA6)** (CAACS)

(42): Smoking is still allowed at many Australian restaurants, and this is one **thing** that I really hate about my home country. **(Native speaker)** (ANACS)

(43): and also I have made many strong relationships with the other trainers and this is the most important **thing** that we should concentrate on. **(CEPA5)** (ANACS)

Examples (36) to (39) are illustrative cases of the SN *thing(s)* being specified within a clause boundary; the reference between the SN and its specifics is established as cataphoric in (36) and (37) and as anaphoric in (38) and (39). In contrast, examples (40) to (43) show instances where the specification of the SN spans beyond a clause boundary; the reference category for (40) and (41) is established as cataphoric across clause and as anaphoric across clause in (42) and (43).

The specification of the SN *thing(s)*, as illustrated in examples (36) to (43), occurs in three grammatical forms: a clause as in (40) to (43), a nominalisation as in (38) and a simple noun phrase as in (39). Table 5-6 provides the raw frequencies and the percentages of these grammatical forms in the learner and native speaker corpora.

Table 5-6 *Specification of the SN thing(s) in the learner and native speaker corpora*

Specification type	**Raw frequency in the learner corpus (%)**	**Raw frequency in the native speaker corpus (%)**
Clause	118 (44%)	60 (70%)
Simple noun phrase	110 (41%)	8 (9%)
Nominalisation	40 (15%)	18 (21%)
Total	268 (100%)	86 (100%)

What the figures in Table 5-6 tell us is that the three grammatical forms with which the SN thing(s) is specified are not equally used in the learner and native speaker corpora. The SN is predominately specified in clauses in the native speaker corpus, representing 70% of all the instances of the noun in the corpus. This is not the case in the learner corpus where the specification of the noun is frequently realised in both clauses and simple noun phrases,

accounting for 44% and 41% respectively of the total number of the noun in the corpus. The log-likelihood test shows that there is a significant difference between the two corpora in terms of the frequency of the aforementioned grammatical forms. The results are – 7.96 for the difference in relation to the use of clauses and + 25.36 with regard to the use of simple noun phrases. What these figures mean is that in the learner corpus the use of clauses as a way to provide the specifics of the SN is significantly under-represented whereas the use of simple noun phrases is significantly over-represented. However, the log-likelihood result for the difference in relation to the use of nominalisation to spell out the specifics of the SN indicates that there is no significant difference between the two corpora, though it is found that this form of specification is slightly under-represented in the learner corpus.

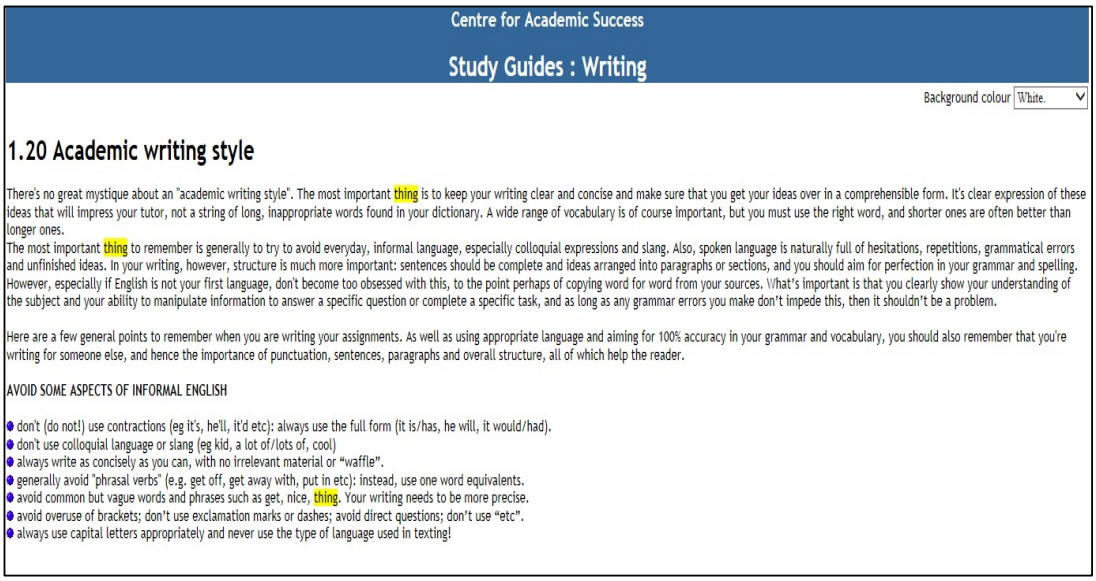
Having investigated the SN *thing(s)* in the learner and native speaker corpora in terms of the reference categories and the grammatical forms by which the specification of the noun is fulfilled, the next section sheds light on the uses of the SN in terms of the patterns associated with the noun as identified in the two corpora.

5.4.2. Uses of the SN *thing(s)*

Unlike the two SNs discussed in this chapter: *time(s)* and *way(s)*, *thing(s)* is a semantically neutral SN and thus can have a myriad of meanings in context that are difficult to categorise. It is for this reason that this section investigates the uses of the SN *thing(s)* in terms of the patterns that are found to occur with the noun rather than the meanings that can be associated with the noun in context. The aim of this section is to examine the patterns identified to occur with the SN *thing(s)* in the two corpora in order to show how the learners and native speakers compare with each other.

Before comparing the learners and native speakers, a point about the use of 'thing' and patterns needs to be made. 'Thing' is certainly useful as a means of establishing links between stretches of language in text either within or across clause boundaries (see Section 5.4.1). Yet, it is also true that there are instances where the use of another word in lieu of the SN *thing(s)* would make writing better. Equally, however, there are other instances where it is difficult to think of an alternative word to be used instead of the SN *thing(s)*. The notion of pattern is important here.

In language teaching, students are widely encouraged to avoid using the SN *thing(s)* in their writing. Figure 5-2 is an illustration of this:


Centre for Academic Success

Study Guides : Writing

Background colour White

1.20 Academic writing style

There's no great mystique about an "academic writing style". The most important thing is to keep your writing clear and concise and make sure that you get your ideas over in a comprehensible form. It's clear expression of these ideas that will impress your tutor, not a string of long, inappropriate words found in your dictionary. A wide range of vocabulary is of course important, but you must use the right word, and shorter ones are often better than longer ones.
The most important thing to remember is generally to try to avoid everyday, informal language, especially colloquial expressions and slang. Also, spoken language is naturally full of hesitations, repetitions, grammatical errors and unfinished ideas. In your writing, however, structure is much more important: sentences should be complete and ideas arranged into paragraphs or sections, and you should aim for perfection in your grammar and spelling. However, especially if English is not your first language, don't become too obsessed with this, to the point perhaps of copying word for word from your sources. What's important is that you clearly show your understanding of the subject and your ability to manipulate information to answer a specific question or complete a specific task, and as long as any grammar errors you make don't impede this, then it shouldn't be a problem.

Here are a few general points to remember when you are writing your assignments. As well as using appropriate language and aiming for 100% accuracy in your grammar and vocabulary, you should also remember that you're writing for someone else, and hence the importance of punctuation, sentences, paragraphs and overall structure, all of which help the reader.

AVOID SOME ASPECTS OF INFORMAL ENGLISH

- don't (do not!) use contractions (eg it's, he'll, it'd etc): always use the full form (it is/has, he will, it would/had).
- don't use colloquial language or slang (eg kid, a lot of/lots of, cool)
- always write as concisely as you can, with no irrelevant material or "waffle".
- generally avoid "phrasal verbs" (e.g. get off, get away with, put in etc): instead, use one word equivalents.
- avoid common but vague words and phrases such as get, nice, thing. Your writing needs to be more precise.
- avoid overuse of brackets; don't use exclamation marks or dashes; avoid direct questions; don't use "etc".
- always use capital letters appropriately and never use the type of language used in texting!

Figure 5-2 *Academic writing style*

Figure 5-2 is taken from a university website[10] in which students are provided with hints on how to write in a more precise and clear way and are told not to use the noun 'thing'. What is

[10]Study guides: writing as found at Birmingham City University website which can be accessed at http://library.bcu.ac.uk/learner/writingguides/1.20.htm (accessed 18 April, 2016).

interesting about the text shown in Figure 5-2 is the fact that the author of the text is explicitly warning students not to use 'thing' in their writing while at the same time he/she is using it as a way to create links between different parts of his/her text. The following are the two sentences used in Figure 5-2 where *thing* is used as an SN: "The most important thing is to keep your writing clear and concise and make sure that you get your ideas over in a comprehensible form." and "The most important thing to remember is generally to try to avoid everyday, informal language, especially colloquial expressions and slang.". The writer in these two quotes use the SN *thing* in the following pattern: "*The* + superlative + ADJ + *thing* + v-link + specifics: to-infinitive appositive clause". When *thing* occurs in this pattern, finding an alternative word to replace the SN *thing* can be difficult and using the SN *thing* in this pattern sounds natural. Although this pattern is found to occur in the learner and native speaker corpora used in the present study, the pattern is found to occur significantly less in the learner corpus. Out of the159 instances of the SN *thing* in the learner corpus, the pattern is observed in only five instances, representing only 3% of all the patterns identified to occur with *thing*.

Table 5-7 presents the most frequent patterns that are found to occur with the SN *thing* in its singular form in the learner and native speaker corpora. All the patterns that occur five times or more are identified as frequent. Table 5-7 also provides the raw frequencies and percentages of the most frequent patterns in proportion to all instances of the SN *thing* in each corpus. A list of all the frequent patterns, with all the instances of the noun illustrating these patterns, is provided in Appendix H (learner corpus) and Appendix I (native speaker corpus).

Table 5-7 *The most frequent patterns of the SN thing across the two corpora (* indicates some instances of divergent patterns; *** refers to redundant use)*

Patterns identified with the SN *thing* in its singular form in descending order of frequency in the learner corpus	**Raw frequency in the learner corpus (%)**	**Patterns identified with the SN *thing* in its singular form in descending order of frequency in the native speaker corpus**	**Raw frequency in the native speaker corpus (%)**
The + (superlative) + (ADJ) + *thing* + (post-modifier) + (v-link)* + specifics: that appositive clause	23 (14.46%)	*The* + (superlative) / another/ one + (ADJ) + *thing* + (post-modifier) + v-link + specifics: that-appositive clause	12 (20.68)
The + (superlative) + (ADJ) + *thing* + (post-modifier) + (v-link)* + specifics: other appositive clause (absence of that element (14) / wh-clause (6))	20 (12.57%)	This/ that/ which + v-link + a/an/the/one + (ADJ) + *thing* ***	9 (15.51%)
The + superlative + (ADJ) + *thing* + (post-modifier) + v-link + specifics: appositive NP/ nominalisation ***	19 (11.94%)	*The* + (superlative) + (ADJ) + *thing* + (post-modifier) + v-link + specifics: to-infinitive appositive clause	7 (12.06%)
It + v-link + a/the (superlative) + (ADJ) + *thing* ***	19 (11.94%)	Specifics: nominalisation/ NP + v-link + an/a + ADJ + *thing* ***	7 (12.06%)
Specifics: Nominalisation/NP + v-link + a/an/the + (superlative) + ADJ + *thing* ***	18 (11.32%)	*The/another/ one* + (superlative) + (ADJ) + *thing* + (post-modifier) + v-link + specifics: nominalisation (3)/ NP (2) ***	5 (8.62%)
The + (ADJ) + *thing* + (post-modifier) + v-link + specifics: appositive NP/ nominalisation ***	15 (9.43%)	*The* + *same* + *thing* (5) ***	5 (8.62%)
The + (superlative) + ADJ + *thing* + (post-modifier) + v-link + specifics: to-infinitive appositive clause	5 (3.14%)		
Other patterns total	40 (25%)	Other patterns total	13 (22%)
Total instances of the SN *thing* in the corpus	159 (100%)	Total instances of the SN: *thing* in the corpus	58 (100%)

It may be difficult to see how the two corpora compare with each other in terms of the patterns that are found to occur with the SN *thing* from Table 5-7. Therefore, the information provided in Table 5-7 will form the basis for my qualitative discussion below of how the SN *thing* is used in the learner and native speaker corpora.

As pointed out earlier, *thing* may occur in patterns where it may be difficult to find an alternative precise noun to be used instead of *thing*. Among the patterns presented in Table 5-7, three patterns in the learner corpus and two patterns in the native speaker corpus are identified to illustrate such a case. These patterns are listed below along with illustrative examples:

Learner corpus:

- ***The*** **+ (superlative) + (ADJ) +** ***thing*** **+ (post-modifier) + (v-link) + specifics: that appositive clause**

  (44): The best **thing** about that vacation was that I learned a lot from it, **(CEPA6)** (with the presence of v-link)

  (45): and the most **thing** that I loved in Germany that the people are very friendly and kind with tourists. **(CEPA5)** (With the absence of v-link)

- ***The*** **+ (superlative) + (ADJ) +** ***thing*** **+ (post-modifier) + (v-link) + specifics: other appositive clause (absence of that element / wh-clause)**

  (46): and the most important **thing** is the people that you love are surrounding you the whole time and **(CEPA6**) (absence of that element)

(47): The most funniest **thing** that happend to me there when my grand mother Lost there with my little sister, **(CEPA4**) (wh-clause)

- ***The*** **+ (superlative) + ADJ +** ***thing*** **+ (post-modifier) + v-link + specifics: to-infinitive appositive clause**

(48): The first **thing** I decided to do when I arrived Oman was to visit my friends in Oman. **(CEPA5)**

(49): We rarely see each other outside Dubai, so we thought the best **thing** to do is to go out together. **(CEPA6)**

Native speaker corpus:

- ***The*** **+ (superlative) / another/ one + (ADJ) +** ***thing*** **+ (post-modifier) + v-link + specifics: that-appositive clause**

(48): and the worst **thing** is that secondhand smokers do not have any choice in whether they will smoke or not. **(Native speaker)**

- ***The*** **+ (superlative) + (ADJ) +** ***thing*** **+ (post-modifier) + v-link + specifics: to-infinitive appositive clause**

(49): The most important **thing** is to ban smoking at all public places. **(Native speaker)**

The writers' use of *thing* in examples (44) to (49) is appropriate and natural. The two writer groups are similar in their use of the 'appropriate' patterns. A closer look at Table 5-7 shows

that these patterns account for 30% of all the identified patterns with *thing* in the learner corpus while the corresponding figure in the native speaker corpus is 33%.

The other remaining patterns in Table 5-7 are formally correct but the SN *thing* in these patterns is semantically redundant. The use of *thing* in these patterns is redundant in that the message could be effectively communicated without the use of *thing* either by replacing it with a more informative noun or by not using a noun head at all. Examples of the redundant uses of *thing* in the learner and native speaker corpora are provided below.

(50): They reached the island they didnt see anyperson The only **thing** that showed that there was life on this island was the skeletons on the ground and every where, **(CEPA 6)** (*The* + (ADJ) + *thing* + v-link + specifics: appositive NP/ nominalisation)

(51): Travilling around the world is a very exciting **thing**, **(CEPA 5)** (Specifics: nominalisation/NP + v-link + a/an/the + (superlative) + ADJ + *thing*)

(52): and it is really a slap in the face when you go to a party and someone is blowing smoke in your face. The same **thing** is true at restaurants, **(Native speaker)** (*The* + *same* + *thing*)

(53): Smoking is kind of an old fashioned **thing** and it doesn't really have a place in the modern world. (**Native speaker)** (Specifics: nominalisation/ NP + v-link + an/a + ADJ + *thing*)

As illustrated by examples (50) to (53), the use of *thing* does not add anything to the semantic content of the communicated propositions. Thus, the writers of examples (50) and (52) could have improved their writing by using more precise nouns instead of *thing*. In example (50), *thing* could have been replaced with 'sign' and 'analogy' would have been a better choice for example (52). The redundant use of *thing* in examples (51) and (53) could have been avoided if the writers of these examples had not used the noun at all. The redundancy in these examples stems from the fact that the message in each example could easily be conveyed without the use of the SN *thing* (`travilling around the world is very exciting, smoking is kind of old fashioned`).

So far, the learner and the native speaker writers show a large degree of similarity in terms of the patterns that occur with *thing* in its singular form. However, a closer look at the patterns associated with *thing* in its plural form indicates the two corpora are strikingly different. The learners heavily rely on two patterns. These patterns are:

- Quantifier + (ADJ) + *things*
- Quantifier + (ADJ) + *things* + specifics: comparative NP/ nominalisation

Out of 109 instances of *things*, 79 instances are found to occur in the above listed patterns, accounting for 72.47% of all the patterns identified in the learner corpus. Conversely, the native speaker writers' use of these patterns represent 46.42% of all the instances of *things*. The following are illustrative examples of the two most frequent patterns in the two corpora:

(54): `I have learned many **things**, the most important is that not to rely on anybody to do my work.` **(CEPA 6)** (Quantifier + (ADJ) + *things*)

(55): Some people spend they holiday in a good and usefull things like going to an institute to improve their skills in English or computer. **(CEPA 4)** (Quantifier + (ADJ) + *things* + specifics: comparative NP/ nominalisation)

(56): which is invastigating about different **things** like people or crimes. **(CEPA 6)** (Quantifier + (ADJ) + *things* + specifics: comparative NP/ nominalisation)

(57): and it has given me a variety of good **things**. For example, I have been able to make a little bit of money for the weekends and other times. **(Native speaker)** (Quantifier + (ADJ) + *things*)

(58): Having a part-time job would give them a little more credibility when applying for jobs and teach them **things** like organizing their time, responsibility and accountability which … **(Native speaker)** (Quantifier + (ADJ) + *things* + specifics: comparative NP/ nominalisation)

The use of the SN *things* in examples (54) to (58) is semantically redundant in that the writers of these examples could have conveyed their messages with more precise SNs. For instance, the writer of example (54) could have improved his/her writing by replacing the SN *things* with an SN that adds to the semantic content of his/her message. In example (54), the writer could have used the SN *lessons* instead of the SN *things*.

5.5. *Way(s)*

Way(s) is the last of the frequent and shared SNs to be discussed in this chapter; like the others it is one of the 15 most frequent SNs in the two corpora. This SN ranks thirteenth in the learner corpus and second in the native speaker corpus. The following two sections present the results of how the learner and the native speaker corpora compare with each other with regard to the use of the SN *way(s)* in terms of the reference categories (5.5.1) and the textual meanings identified (5.5.2).

5.5.1. Reference categories of *way(s)*

Table 5-8 shows the raw frequency of the specification categories of *way(s)* as found to occur in the learner and native speaker corpora. It also provides the percentages of the reference categories.

Table 5-8 *Specification categories of the SN way(s)*

	Raw frequency in the learner corpus (%)		**Raw frequency in the native speaker corpus (%)**	
Within clause specification (WCS):	27 (66%)		59 (55%)	
Cataphoric and anaphoric	19 (70%)	8 (30%)	23 (39%)	36 (61%)
Across clause specification (ACS):	14 (34%)		48 (45%)	
Cataphoric and anaphoric	7 (50%)	7 (50%)	11 (23%)	37 (77%)
Total frequency	41 (100%)		107 (100%)	

There is no significant difference between the two corpora in terms of the frequency of two reference categories of the specification of the SN *way*[11]. The log-likelihood results for the difference are + 0.57 for the WCS category and – 0.85 for the ACS category. Although the difference is not significant, the WCS is found to be slightly over-represented in the learner corpus and the ACS is found to be slightly under-represented[12]. As was found in the examination of the SN *thing(s)*, the learners here also show less preference towards the use of the SN *way(s)* to organise stretches of language beyond clause boundaries. As noted in the case of the SN *thing(s)*, it is also possible here that the less use of *way(s)* to organise text is because that the learners are more likely to think in terms of sentences only rather than longer stretches. The following examples: (59) and (60) are taken from the native speaker corpus to show how the SN *way(s)* can be used to organise texts.

(59): On that basis, I feel that it is crucial that society as a whole stands as one and rather than condemn smokers, continue to vote only for those representatives who will take up the cause in government and push for sufficient legislation to ban smoking in restaurants as well as in all public places. Only in this **way**, can long-term change take affect and society in general, like it usually does, will eventually toe the line.

(60): There are **ways** to achieve this end without the complete prohibition of smoking at Japanese restaurants. One viable

[11] The log-likelihood test was run on the raw frequency of each reference category in the two corpora in proportion to the raw frequency of all instances of the SN *way(s)* in the two corpora.

[12] Explanation of how the results obtained from the log-likelihood test is to be interpreted is explained in Section 4.5.

option is to mandate the use of separately ventilated smoking rooms.

In examples (59) and (60), the SN *way(s)* is specified by clauses. A closer examination of all the instances of the SN *way(s)* in the learner and native speaker corpora shows that the noun can also be specified by nominalisations and simple noun phrases. Table 5-9, below, shows the frequencies of the three grammatical forms in the two corpora.

Table 5-9 *Specification of the SN way(s) in the learner and native speaker corpora*

Specification type	**Raw frequency in the learner corpus (%)**	**Raw frequency in the native speaker corpus (%)**
Clause	25 (61%)	68 (63%)
Nominalisation	11 (27%)	18 (17%)
Simple noun phrase	5 (12%)	21 (20%)
Total	41 (100%)	107 (100%)

As illustrated by the figures in Table 5-9, the SN *way(s)* is much more frequently specified by clauses; 25 (61%) of all the occurrences of the SN in the learner corpus and 68 (63%) of the 107 instances of the SN in the native speaker corpus are specified by clauses. Illustrative examples of such instances are found in (59) and (60) presented earlier. The other two types of specification: nominalisations and simple noun phrases are less represented in the two corpora. These types of specification are shown in examples (61) and (62); the SN is specified by a nominalisation in example (61) and by a simple noun phrase in example (62).

(61): Sending emails and chatting with friends is another **way** to benefit from it. **(CEPA 5)**

(62): People’s taste in movies can differ, but at the end I think we all agree that movies are a nice **way** to spend one’s time. **(CEPA 6)**

5.5.2. Meanings associated with the SN *way(s)*

A closer look at the discourse providing the specifics of the SN *way(s)* in the learner and native speaker corpora has revealed that the specification can be classified into three meanings. Figure 5-3, below, presents the classification of these meanings:

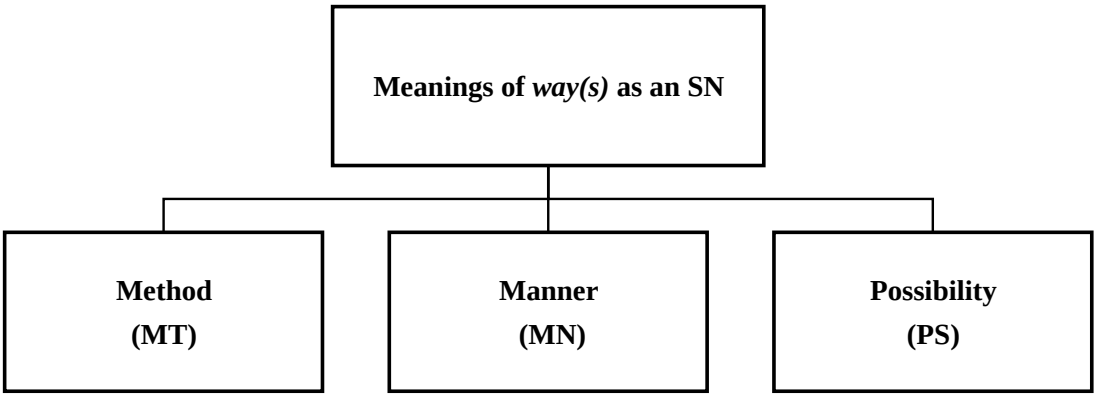


Figure 5-3 *Meanings associated with the SN way(s)*

The SN *way(s)* is identified as a *method* (MT) when its specification spells out how something is done or achieved. When the specification provides the manner in which someone does something or in which something happens, *way(s)* is identified as a *manner* (MN). *Possibility* (PS) is another meaning associated with *way(s)* as an SN; this meaning is assigned to the noun when the specification refers to possibility. The following are illustrative examples of the three meanings identified to occur with the SN *way(s)* in the learner and native speaker corpora:

(63): she said that the best **way** to spend the summer is to go somewhere overseas, which...**(CEPA 5)** (*Method*)

(64): I don't believe that legislation is the best **way** to ban smoking in restaurants. **(Native speaker)** (*Method*)

(65): Holidays are usually a **way** for someone to relax and enjoy their time. **(CEPA 6)** (*Method*)

(66): The best **way** to solve this problem is simply to have a smoking section and a nonsmoking section in every restaurant. **(Native speaker)** (*Method*)

(67): In the other hand what leaded me to Judge at the director harshly is the **way** he directed the killing events of the palestinian. He showed how bloody and having no mercy the palestinian were as if they are robot machine killing for him!! (**CEPA 6**) (*Manner*)

(68): A restaurant is usually a private concern and should be left that **way** if they meet the standard health regulations related to food and drink that I mentioned. (**Native speaker**) (*Manner*)

(69): After that when we was in ahospital the doctor came and told my mother that my father health it was very bad and it's no **way** to be alife. (**CEPA 4**) (*Possibility*)

(70): ...because there's just no **way** that smoking will be allowed in restaurants forever. **(Native speaker)** (*Possibility*)

As was found in the examination of the SN *time(s)*, a closer look at all the instances of *way(s)* in each of the three meaning categories as identified in the learner and native speaker corpora has also revealed that there is a link between the meanings identified and the grammatical forms by which the SN is specified. When the SN is found to occur in the *method* group, the specification of the noun can be realised as a simple noun phrase as in (65), a nominalisation as in (64) and a clause as in (66). However, when the SN is found to occur in both the *manner* group and the *possibility* group, the specification of the noun is always realised as a clause, as illustrated in examples (67) to (70).

The raw frequencies of the three meaning categories of the SN *way(s)* are given in Table 5-10. As shown in Table 5-10, 'method' is by far the most frequent meaning in the two corpora, followed by 'manner'. The least frequent meaning is 'possibility'. As found in the discussion of the meanings of the SN *time(s)*, the learner and native speaker corpora here also show a high degree of similarity with regard to the use of the three meanings in context that are associated with the SN *way(s)*.

Table 5-10 *Meanings of the SN way(s) across the two corpora*

Meanings associated with the SN *way(s)*	Raw frequency in the learner corpus (%)	Raw frequency in the native speaker corpus (%)
Method (MT)	25 (61%)	73 (68%)
Manner (Man.)	13 (32%)	27 (25%)
Possibility (PS)	3(7%)	7 (7%)
Total frequency	41 (100%)	107 (100%)

5.6. Summary of the findings of the chapter

This chapter has focused on three SNs: *time(s)*, *thing(s)* and *way(s)*. The analysis has revealed a number of interesting findings with regard to the uses of these SNs. Perhaps, the most interesting one is that there is a wide range of uses of these SNs in terms of grammatical forms of specification, meanings in context, reference categories and in terms of patterns as shown in the examination of the SN *thing(s)*. In a way, it is not surprising because these SNs are very frequent in the two corpora but in a way it is surprising that there is quite very much variation in the uses of these words as SNs.

The findings presented in this chapter will be discussed in Chapter 8. Overall, it can be concluded that the chapter has contributed to the existing literature on the phenomenon of SNs by examining in more detail how particular SNs are used.

CHAPTER 6: SNs AND STANCE

6.1. Overview

This chapter is concerned with SNs as a potential resource of stance in discourse. The first part of the chapter presents a review of the literature about this phenomenon of stance signalling nouns (SSNs henceforth). The second part of the chapter provides an account of findings related to SSNs in the two corpora used in the current study. The findings are presented in two main sections. The aim of the first section (6.3) is to explore how stance through nouns can be constructed and also to identify the patterns associated with the construction of stance. The second section (6.4) compares the two corpora used in the current study in terms of SNs with an inherent stance function. The chapter concludes with a summary of the findings, presented in Section 6.5.

6.2. SSNs: a literature review

Signalling nouns (SNs) and similar phenomena, as reviewed in Chapter 2, have been extensively studied, albeit under a variety of different labels. The labels assigned to SNs are *general nouns* (Halliday and Hasan, 1976), *anaphoric nouns* (Francis, 1986), *retrospective* and *advance nouns* (Francis, 1994), *shell nouns* (Hunston and Francis, 1999; Schmid, 2000) and *signalling nouns* (Flowerdew, 2003; Flowerdew and Forest, 2015). No matter how SNs are labelled, it can be said that the focus placed on this type of nouns has largely been on the role these nouns play in the organisation of texts. This is because of the fact that SNs are capable of encapsulating and prospecting the meaning of either a previously mentioned or succeeding discourse (Rego, 2001, Flowerdew and Forest, 2015).

While SNs as an organising discourse resource have received considerable attention in the literature, their potential in constructing stance towards a stretch of discourse has attracted less attention. The exceptions to this are the studies by Halliday and Hasan (1976), Francis (1986, 1994) and Hunston (2011). Although Halliday and Hasan (1976) make mention of the possibility of expressing evaluative meanings by the use of *general nouns*, a full picture of such potential remains not elucidated. Francis (1986; 1994) points out that many *anaphoric nouns* and *retrospective labels* have an evaluative meaning. In her studies, the main focus was on the role that nouns play in text organisation and thus she has briefly addressed the evaluative potential of nouns. Hunston (2011) argues that evaluation of propositions can be achieved by using 'status nouns', nouns such as *fact* and *assumption*. The investigation of status nouns, however, is limited to those nouns that are followed by that-appositive clauses. The present chapter aims to contribute to the literature on stance nouns by investigating how SNs offer an important stance resource. The examination the phenomenon of SSNs is not restricted to given patterns. The phenomenon is looked at from a discourse perspective and patterns associated with SSNs are identified accordingly.

Stance has been defined as a linguistic resource by which writers can express their "personal feelings, attitudes, value judgments or assessments" towards propositions (Biber, Johansson, Leech, Conrad, and Finegan, 1999, p.966). By applying this resource appropriately, writers are able to foreground their opinions, attitudes or evaluation of a given stretch of discourse. Writers, as stated by Jiang and Hyland (2015, p.530), use culturally accepted resources when they produce texts, selecting acts of stance that locate those texts within given communities or disciplines. This means that by employing stance, writers not only express their personal views towards propositions but they also align these propositions with the epistemic values of

the world around them: communities or disciplines (Hunston, 1989, 1994; Jiang and Hyland, 2015).

Having briefly touched on the importance of stance, the obvious question is how stance is constructed. To answer this question, the literature on the resources associated with stance will be addressed below.

Stance as a linguistic resource available at the disposal of writers has attracted much attention in the literature. The attention has mostly been devoted to the construction of stance in academic writing (Biber and Finegan, 1989; Hunston, 1989, 1994 under the term *evaluation*; Hyland, 2006 under the label *hedging*; Conrard and Biber, 2000; Biber, 2006a, 2006b). The shift in how academic texts are viewed can be regarded as a plausible explanation for this emphasis. It is contended that the commonly held perspective that views academic writing as "an objective faceless and impersonal form of discourse" has lost its currency (Hyland, 2005, p.173). Thus, academic texts are to be seen as interactive in that academic text writers not only offer a representation of the world around them but also they construct language that evaluates this representation (Hyland, 2005; Jiang and Hyland, 2015).

Biber and Finegan (1989, as cited in Thompson and Hunston, 2000, p.19) provide one list of possible linguistic markers by which stance can be constructed. These markers are:

1- verbs, adverbs and adjectives, showing certainty, doubt and affect;

2- hedges indicating vague language;

3- emphatics;

4- modals that expresses possibility, necessity, and prediction.

What is noticeable about Biber and Finegan's list of stance markers is the absence of the resource of nouns. Using a similar corpus methodology, Conrad and Biber (2000) investigate adverbials that indicate stance and quantify them in different registers (Biber in a later work includes nouns in his list of stance markers, a point that will be addressed below). The authors identify three types of stance; *epistemic stance* (e.g. *probably*), *attitudinal stance* (e.g. *amazingly*) and *style stance* (e.g. *briefly*). The authors are able to show the importance of adverbials to the construction of stance in texts both written and spoken.

Hunston and Sinclair (2000) have devised a local grammar of evaluation in English. The aim was to identify possible patterns associated with the concept of evaluation. They have pinpointed patterns that can be employed to indicate evaluation within adjectives and nouns. However, the two authors have chiefly limited their focus to adjectives as an available resource of evaluation in English.

Biber in his most recent work (e.g. Biber, 2006a, 2006b) has extended the investigation of the linguistic features by which speakers and writers can express their stance. Biber (2006a, 2006b) is able to identify a wide range of stance establishing features. The features are 'modal verbs' (e.g. *may*), 'adverbs' (e.g. *possibly*), 'verbs' (e.g. *believe*), 'adjectives' (e.g. *interesting*) and 'nouns' (e.g. *claim*). Unlike Biber and Finegan (1989) and Conrad and Biber (2000), Biber (2006a, 2006b) is able to draw attention to nouns as a potential marker of stance.

Having reviewed some of the literature on the resources through which stance can be conveyed, it is evident that nouns as a potential resource of evaluation have attracted relatively little attention until recently. As mentioned earlier, nouns are not among the list of

stance markers by Biber and Finegan (1989) and Conrad and Biber (2000). Though Biber (2006a, 2006b) includes some nouns among the list of features used to convey stance in academic genres. The little attention given to stance nouns is, in fact, not surprising since the use of nouns as stance resource can be regarded as the less traditionally obvious resource of stance.

It is not until recently that the construction of stance through nouns has been investigated. This can be found in Charles' studies (2003, 2007), Biber (2006a, b), Hunston (2011) and Jiang and Hyland (2015). These studies have examined the use of the resource of stance nouns across academic disciplines and genres. Despite the fact that these studies have acknowledged the importance of nouns in the construction of stance in texts and have brought it to the fore, these studies are not without disadvantages.

One drawback is that these studies have restricted their investigation to stance nouns that belong to few numbers of patterns. The first is the '*noun* + appositive-that clause' pattern. This can be found in Biber (2006a, b), Charles (2007), Hunston (2011) and Jiang and Hyland (2015). The second pattern associated with stance nouns is 'This + *noun'*. It is this pattern that Charles (2003) has based her analysis and findings on. The third pattern is '*noun* + to appositive clause'. This is only found in Biber (2006a, 2006b). By restricting the investigation of stance nouns to particular patterns, many instances of stance nouns may be left undiscovered. Thus, it is an aim of this chapter not to restrict SSNs to already prescribed patterns in order to yield a comprehensive picture of how the resource of stance through SSNs is constructed.

Another limitation of the aforementioned studies is that these studies are exclusively devoted to SSNs in academic contexts. The present chapter broadens this in two ways. It investigates SSNs in non-academic contexts. It also compares SSNs in learner and native speaker English.

6.3. Patterns and stance nouns

As mentioned previously, this chapter aims to explore and identify the patterns associated with the construction of stance through nouns. To this end, I have focused on the inherent SSNs identified in the two corpora, as shown in Table 6-2 presented later in Section 6.4. By inherent SSNs, I mean SNs that signal evaluative meanings in themselves, as opposed to those that act as carriers for stance-indicating adjectives. For example, *problem* is an inherent SSN, but *point* (as in *a good point*) is not. In the cases where there is an inherent SSN, the specification of the noun provides the proposition that is evaluated by the noun.

The concordance lines of SSNs are thoroughly examined to identify the patterns that can be said to frequently occur with SSNs. The main finding is that there are many ways of constructing stance through the resource of SSNs. This finding will be addressed in the next sections. Before I present the patterns, two methodological points should be noted. The first is that the patterns are divided into two main categories. I follow Flowerdew and Forest's (2015) categorisation of patterns. The two groups of patterns that they use are within clause boundary and across clause boundary. The second is that the patterns that are associated with within and across clause boundary are selected based on their frequent occurrences in the two corpora. All the patterns that occur five times or more in the learner and native speaker corpora are identified as frequent.

6.3.1. Construction of stance within clause boundary

Stance through nouns can be constructed within a clause boundary. This means that both the stance signalling noun (SSN) and the signalled evaluated entity are to be found within the same clause. The most frequent patterns associated with this type of stance construction, based on the definition above, are presented below.

6.3.1.1. SSN + specifics: other appositive clause

Although a SSN followed immediately by a **that** appositive clause is a common way of constructing stance (e.g. Charles (2007); Hunston (2011); Jiang and Hyland (2015)), in my data the pattern: *SSN + specifics: other appositive clause* is the most frequent of all the patterns identified to occur with SSNs. An appositive clause in English is not only expressed in a that-complement clause. In fact, it can be expressed in a range of grammatical choices such as an appositive clause with the absence of **that** element, to-infinitive appositive clause, Wh-clause, prepositional appositive clause and gerund appositive clause. It should be noted that, unlike some researchers, I define a that-appositive clause as including the word *that* and I classify clauses without *that* as other appositive clauses.

The following examples are illustrative of the pattern: SSN + specifics: other appositive clause. In these examples, all the aforementioned grammatical options available when using SSNs in this pattern are illustrated:

(1): What made this trip such a memorable & wndoerfull trip, is the **fact** I have had the chance to meet new people, see new places and look at buildings that have made history over time. **CEPA 6** (SSN + specifics: appositive clause with the absence of that element)

(2): For these reasons, I think it is a good **idea** to ban smoking at restaurants in Japan **Native speaker** (SSN + specifics: to-infinitive appositive clause)

(3): If they are having **problems** getting a job then they should look at doing some volunteer type **Native speaker** (SSN + specifics: gerund appositive clause)

(4): We are adults and have the **right** to make decisions for ourselves, **Native speaker** (SSN + specifics: to-infinitive appositive clause)

(5): We eventually realised that the **idea** of making a plan in the first place was actually a mistake. **CEPA 6** (SSN + specifics: prepositional appositive clause)

(6): remember a man passing very near our table but have no **idea** how he stole it so easily. After those two incidents happened, **CEPA 6** (SSN + specifics: wh- appositive clause)

In each of the above examples, the appositive clause underlined provides the proposition that is evaluated by the SSN in question. Both the evaluated propositions and the SSNs are found to occur within clause boundaries.

6.3.1.2. Construction of stance through a that-appositive clause

Stance can be constructed with a SSN followed by a that-appositive clause. There are two patterns that involve the construction of stance in a that- appositive clause. These patterns are listed below:

- SSN + specifics: that appositive clause
- SSN + v-link + specifics: that appositive clause

The that-appositive clause spells out the evaluated proposition as shown in the following examples taken from the learner and native speaker corpora. The following examples (7) to (13) illustrate the first pattern listed above.

(7): I agree with the **idea** that smoking should be banned at restaurants in Japan... **Native speaker**

(8):...and they have an **idea** that all the gulf people are rich. That wasn't true. **CEPA 5**

(9): Nevertheless, I do think it is a very serious **problem** that the country allows smoking in public places... **Native speaker**

(10): I simply loved the **fact** that the city remained the same and took pride of what they had. **CEPA 6**

(11): Smokers can no longer ignore the **fact** that by smoking around others, smokers or non smokers alike, they are endangering everyone’s health in that enclosed PUBLIC space. **Native speaker**

(12): What makes this vacation so special is the **fact** that I travelled to a wonderful country with my bestest friend and we had the opportunity to do everything that we love so much together. **CEPA 6**

(13): There is merit to the **opinion** that college students should have part time jobs. **Native speaker**

As can be seen from examples (7) to (13), the writers use SSNs: *fact, idea, problem* and *opinion,* followed by a that-appositive clause to construct their stance towards the propositions that they make.

Compared to the first pattern, the second pattern listed above is less frequent in the two corpora. Both patterns, though, are widely recognised in the literature that dealt with the phenomenon of SNs as major syntactic patterns upon which SNs can be identified (Charles (2007); Hunston (2011); Jiang and Hyland (2015); Schmid, 2000; Flowerdew and Forest, 2015). The following are examples of SSNs identified in the two corpora to occur in the pattern: *SSN + v-link + specifics: that appositive clause*.

(14)…and the greatest **problem** was that my father and my old brother died in this accident. **CEPA4**

(15): And the big **problem** was that I didn't know how to speak French. **CEPA5**

(16): One **problem** they will face is that many restaurants sell cigarettes and make part of their profit from doing that… **Native speaker**

(17): The **fact** is that we have a right to smoke if we want to. **Native speaker**

The evaluated propositions in the examples (14) to (17) are found in the complement clause presented after the verb to BE. When SSNs are used with this pattern, the SSNs are presented in the subject position of the clause. The following pattern is just the opposite.

6.3.1.3. Specifics + v-link + SSN

This pattern is observed when the v-link is preceded by the evaluated proposition and is succeeded by the SSN. Flowerdew and Forest (2015) note that this pattern is not the prototypical syntactic pattern associated with the phenomenon of SNs. However, the examination of SSNs in the current study has revealed that this pattern is widely used in the two corpora. The following examples are illustrative of this finding:

(18): Freedom is a basic **right** of the individual in any civilized society. **Native speaker**

(19): Being able to work in a healthy environment is a basic **right**. **Native speaker**

(20): Choosing to work in college really is a personal **choice**. **Native speaker**

(21): Going to Eygpt is great **idea** for having a great summer holidy. **CEPA6**

(22): Due to the daily exhaustion throughout the week, the fact of spending a relaxing vacation was the ideal **choice**. **Native speaker**

(23): It seems like these days work ethics is a big **issue**. **Native speaker**

The evaluated propositions in the above example are marked since the prototypical specification of SSNs when constructed within clause boundary, as noted above, is to succeed the SSN category. This marked use of SSNs can be said to emphasise the stance taken towards the evaluated stretch of language.

6.3.1.4. Patterns of SSNs occurring in adjunct phrases

There are two patterns identified with SSNs when occurring in adjunct phrases. These patterns are listed below:

- SSN in adjunct phrases + specifics: appositive clause
- Specifics: appositive clause + SSN in adjunct phrases

The two patterns are only found to occur with the SSNs *opinion* and *fact*. The following examples (24) to (29) illustrate the patterns listed above, examples (24) to (27) are illustrative cases of the first pattern while the second pattern are demonstrated in examples (28) and (29):

(24): In my **opinion**, Oman is the best country to visit it. **CEPA 5**

(25): In my **opinion**, it is very important for students to have a part-time job. **Native speaker**

(26): In **fact** this computer is the bist thing I got, **CEPA 5**

(27): In **fact**, in many countries, smoking has been completely banned in restaurants, **Native speaker**

(28): The Movie was funny and sad in my **opinion. CEPA 4**

(29): Banning smoking in all restaurants in Japan would be over-regulating in my **opinion**. **Native speaker**

6.3.1.5. SSN + specifics: prepositional phrases

Unlike the previous pattern where the specification of the SSN is specified in clauses, this pattern represents instances where the specification is provided within clauses in

prepositional phrases. This pattern and the two remaining patterns of within clause patterns, which will be shown in the following two headings, are the least frequent patterns that occur with SSNs in the two corpora. Examples (30) to (32) provide instances of SSNs specified in prepositional phrases.

(30): However, this individual right to expose oneself to risk cannot supersede the collective **right** to safety of all citizens within a given public space. **Native speaker**

(31): Finally, there is the **issue** of health costs. **Native speaker**

(32): If given a **choice** between money and classes, many people especially early on in one's life would choose money. **Native speaker**

6.3.1.6. SSN + v-link + specifics: other appositive clause

This pattern is the second least frequent in the two corpora. The following are examples demonstrating the pattern.

(33): The biggest **problem** that slipped our mind was who will drive the yacht? **CEPA 6** (SSN + v-link + specifics: wh- appositive clause)

(34): The best **choice** would be to look for work which complements one's studies. **Native speaker** (SSN + v-link + specifics: to-infinitive appositive clause)

(35): I don't understand why anyone thinks they have the **right** to tell me what I can or cannot do in a public place. **Native speaker** (SSN + v-link + specifics: to-infinitive appositive clause)

6.3.1.7. SSN + specifics: comparative

This is the least frequent pattern in the two corpora. The specification of the SSN is spelled out by means of exemplification. Illustrative examples are provided below:

(36): How about the **rights** of others though, like say, thousands of non-smokers who like to dine out? **Native speaker**

(37): Smoking can cause many health **problems** such as lung cancer, heart disease, and emphysema. **Native speaker**

Having examined the construction of stance through SSNs within a clause boundary, I will now move on to examine cases when SSNs are used in a way in which the evaluated text goes beyond clause boundary. The following section deals with across clause construction of stance through SSNs.

6.3.2. Construction of stance across clause boundaries through prospection and encapsulation

Stance can be constructed through SNs by means of prospection and encapsulation. When SSNs serve these two mechanisms, the evaluated entity that has been signalled by the stance noun is either in the preceding or in the following discourse. In this sense, the evaluated discourse (the specification of SSN) in relation to the evaluating category (SSN) is tied in with the notion of prospection and encapsulation proposed in Sinclair's model of text organisation (Sinclair, 1993).

6.3.2.1. Patterns associated with the construction of stance through SSNs by means of encapsulation

As mentioned previously, Charles (2003) has investigated the construction of stance through *noun* preceded by the deictic *this*. When used with this pattern, nouns encapsulate previous discourse with the element of stance being incorporated to show how this discourse is to be interpreted (Charles, 2003, p.313). The pattern: *This + SSN* as identified in Charles' study also occurs in my data. Examples of this pattern are presented below:

(38): We do know that cigarettes contain chemicals that are toxic and not natural for human consumption. Based on this **fact** alone, it says a lot about the deadly effect of smoking whether you're a smoker or a naive non-smoker. **Native speaker**

(39): I disagree with this **opinion**. I have three opinions. **Native speaker**

(40): The best way to solve this **problem** is simply to have a smoking section and a nonsmoking section in every restaurant. **Native speaker**

(41): ...and we called the police. to solve this **problem**. **CEPA 4**

The pattern: *This + SSN*, as shown in the examples, is used by the writers to encapsulate a previously mentioned proposition, helping them to construe their stance towards the proposition either epistemically as in examples (38) and (39) and attitudinally as in examples (40) and (41).

The examination of the concordance lines of the SSNs in the two corpora has also revealed that SSNs can be also constructed by means of encapsulation if the SSNs occur in the pattern: *that + SSN*. The following examples are illustrative of this:

(42): Last I heard, Japan's voters, the majority of who are smokers, weren't too keen on that **idea** and cracked it down. **Native speaker**

(43): Why take away that **right**. Smoking is not dangerous like drinking a driving. **Native speaker**

(44): The rights of smokers do not extend to allowing them to cause physical damage to nonsmokers. In this, I agree wholeheartedly. As a smoker, I do not have that **right**. **Native speaker**

The writers use SSNs preceded by the deictic determiner **that** to encapsulate previously stated propositions. This encapsulating pattern is found to occur less frequently with SSNs in the two corpora of the current study compared to that of *This + SSN*. This might explain why Charles (2003) has limited her study to SSNs preceded by the deictic *this*. However, if one to get a wider picture of the patterns associated with SSNs that can be constructed through encapsulation, it is important to include *That + SSN* pattern as an available pattern by which stance can be established.

Another similar encapsulating pattern is *It/ this/that + v-link + SSN*. This pattern is frequently used with SSNs in the two corpora. The following three examples illustrate the pattern:

(45): I am not saying that this is true for all students, it is just my **opinion** based on the comparisons that I have made between different categories of students who I happen to know personally. **Native speaker**

(46): So a part-time job at university is not just a way to make a bit of extra pocket money, but a way to secure a great career and a great future. This is an **opportunity** which should not be passed up. **Native speaker**

(47): The story begens when my brother who works in dubai mnusipality as a maneger, desided that the whole family should go to Abudabi to see his activities there. We thought that was agrate **idea** and... **CEPA 5**

Before I continue to present other patterns that accompany SSNs, it is useful to note that there is a strong tendency in the literature on SSNs to focus on the singular form of SSNs. This has led to patterns associated with SSNs when used in their plural form left unidentified. Thus, the current thesis takes into consideration both forms to gain a more complete picture of the phenomenon of SSNs.

SSNs with plural forms can encapsulate previously mentioned stretches of discourse. By stretches of discourse, I mean that plural SSNs may encapsulate a couple of sentences rather than a single sentence or proposition, as shown in the following examples taken from the native speaker corpus:

(48): Based on those **facts** alone, it is difficult to see how smokers feel so offended and picked on.

(49): It is therefore necessary to take steps to reduce these **issues** as much as possible

The authors in the above examples use the SNs *facts* and *issues* to encapsulate a number of already stated propositions that they believe to be facts as in example (48) and issues as in example (49). The evaluated propositions in the two examples are expressed in previous sentences. From these examples, two patterns associated with plural form of SSNs can be drawn. The patterns are *these* + *SSNs* and *those* + *SSNs*.

The definite article preceding SSNs can indicate that the proposition being evaluated has already been mentioned in the text. When doing so, *The* + *SSNs* pattern can be said to construct stance by means of encapsulation. To illustrate this, the following examples are provided:

(50):**...**they decided to watch a film, I suppoted the **idea** because they will stop talking and watch the movie. **CEPA 6**

(51): It therefore does not surprise me that the **issue** has begun to arise in Japan as well. **Native speaker**

(52):...I feel that the arguments that were presented from both sides would be similar to the arguments that will be presented from both sides if the same **issue** comes under discussion in Japan. **Native speaker**

(53): Based on the above reasons, I support the **opinion**. **Native speaker**

The SSNs in these examples refer back to earlier propositions. This is indicated by the use of the definite article preceding the SSNs. This is only true when the SSN is not followed immediately by an appositive clause, as illustrated in examples (50) to (53). However, when the SSN is followed by an appositive clause the definitive article is best understood as a *structural cataphora* (see Halliday and Matthiessen (2004, p.552)). Illustrative cases are given in the following examples:

(54): agree with the **idea** that smoking should be banned at restaurants in Japan and ... **Native speaker**

(55): What made it very dear to me was the **fact** that it was a gift from my father. **CEPA 6**

(56): There is merit to the **opinion** that college students should have part time jobs. **Native speaker**

(57): Finally, there is the snowball **issue**: if we ban smoking due to health concerns, what will be next? **Native speaker**

SSNs preceded by possessive determiners can encapsulate a previously stated proposition. Examples (58) and (59) show instances of SSNs preceded by possessive determiners:

(58): I think it is not important for students to have a part time job while they are in college. There are two main reasons for my **opinion**. **Native speaker**

(59): After he had cup of water he said that Australia is a long journey and it takes may be 10 hours to reach the capital. I started with my sister to change his **idea**. **CEPA 5**

6.3.2.2. Patterns associated with SSNs constructed by means of prospection

SSN construction through prospection can be said to occur when the noun signals that an evaluated stretch of language is to be present in the following discourse. Example (60) is an illustration of this:

(60): In addition, to that I faced another **problem**. One day my grand mother went to the masjid alone and she losted her way. **CEPA 5**

The underlined sentence represents the prospected proposition, which the author views to be a problem. The writer's use of "*another*" preceding the SSN *problem* makes it expected that a problem is to be specified in the evolving discourse.

Closer inspection of the patterns associated with the construction of stance by means of prospection in the two corpora has revealed that there are two patterns, which more frequently co-occur with SSNs. The patterns are listed below:

- Quantifier + SSN
- Patterns including prospecting verbs:

  Prospecting verb + (quantifier) + SSN or (quantifier) + SSN + prospecting verb

Quantifier + SSN is one of the most common patterns identified with SSNs that are found to prospect a stretch of language. This pattern is found in the following examples:

(61): I agree with this opinion. I have three **opinions**. First, a part time job in their field of study would be beneficial. **Native speaker**

(62): I disagree with this opinion. I have three **opinions**. First, if smoking is banned in all restaurants this may begin a trend in other establishments such as bars and nightclubs. **Native speaker**

(63): We had many **problems** we arrived there like loosing some suitcases. **CEPA 5**

The second pattern identified with SSNs, more particularly with the SSN *problem*, when acting to prospect is interesting. It is a combination of SSNs followed by a prospecting verb (SSNs occur in the subject position of clauses) or SSNs preceded by a prospecting verb (SSNs occur in the object position of clauses). Prospecting verbs are those verbs that indicate that something is likely to be told more particularly when occurring with SSNs. The following sets of examples are illustrative of this:

- SSN + prospecting verb:

  (64): We went there for one week but we wished that we had come back earlier than that because **problems** just kept happening to us. **CEPA6**

  (65): When we reached to the airport, the **problems** began, we lost our bags and everything we brought with us expect the money which was in a small pocket that my father took it in his hands. **CEPA5**

  (66): Our deal was to divide the coast of renting the room. Here is the **problem** begun, When I tried to pay I noticed that I forgot all my money in my home. **CEPA5**

❖ Prospecting verb + SSN:

(67): At last we went to the hottle, but her comes the terrible **problem**. My mothers gold was stolen and also our clothes and my fathers important paper. **CEPA5**

(68): Two days later we faced some **problems**. Big problems well both my dad and my brother got sick. **CEPA5**

(69): Even though this idea may not pertain to Japanese college students as much as it does for students in the U. S. I will voice my **opinion** anyway. **Native speaker**

In addition to the previous prospecting patterns, other patterns in the two corpora are also identified. However, they have not met the pattern frequency definition mentioned earlier and thus will not be addressed. One example of these infrequent prospecting patterns is *there + v-link + SSN*. The existential pattern preceding SSNs indicates that SSNs exist and that their specifics are more likely to be supplied in the evolving discourse. Examples of SSNs with this pattern are presented below:

(70): Finally, there is the snowball **issue**: if we ban smoking due to health concerns, what will be next? **Native speaker**

(71): and Their is some big **problem** our car is stopped. **CEPA4**

Having addressed all the frequent patterns occurring with SSNs either within clause or across clause boundaries, Table 6-1, below, lists all these patterns along with their frequencies in the learner and native speaker corpora.

Table 6-1 *Patterns associated with SSNs in the learner and native speaker corpora*

Pattern	Raw frequency in the learner corpus	Raw frequency in the native speaker corpus	Total frequency
Within clause boundary patterns			
SSN + specifics: that-appositive clause	14	36	50
SSN + specifics: other appositive clause	39	146	185
SSN + specifics: comparative	0	6	6
SSN + specifics: prepositional phrases	3	21	24
SSN + v-link/ ø v- link + specifics: that-appositive clause	8	18	26
SSN + v-link + specifics: other appositive clause	7	2	9
Specifics + v-link + SSN	3	36	39
Patterns of SSNs occurring in adjunct phrases: - SSN in adjunct phrases + specifics: appositive clause - Specifics: appositive clause + SSN in adjunct phrases	52 1	17 3	73
Across clause boundary patterns			
(Prep.) + *this/that* + SSN	10	43	53
(Prep.) + *these/those* + SSN	0	5	5
The + SSN	7	27	34
Poss. + SSN + (*above*)	4	13	17
It/this/that + v-link + SSN	16	19	35
Patterns including prospecting verbs: - Prospecting verb + (quantifier) SSN - (Quantifier) SSN + prospecting verb	 13 5	 3 0	21
Quantifier + SSN	6	5	11
Others (patterns occurring less than three times)	4	5	9

So far, I have investigated how stance is constructed through SSNs and have identified the patterns associated with the SSNs used in the learner and native speaker corpora. The section that follows aims to compare the two corpora in terms of the use of SSNs.

6.4. Analysis and findings of SSNs in the two corpora

The aim of this section is to compare how SSNs are used in the learner and native speaker corpora in order to find out how the two corpora resemble or differ from each other. SSNs are to be seen as a potential resource by which the stance of writers of texts can be signalled. This resource is used in the learner and the native speaker corpora. The findings on how SSNs are used in the two corpora are provided below.

Table 6-2 demonstrates the raw and normalised frequencies of the most frequent SSNs that occur in the two corpora. The table lists SSNs that are found to occur among the top-25 SNs in the learner and native speaker corpora used in the present study.

Table 6-2 *Stance signalling nouns appearing in the top-25 SNs in the two corpora*

Rank	**Lemmatised SSNs in the learner corpus**	**Raw frequency**	**Normalised frequency per 100,000 words**	**Rank**	**Lemmatised SSNs in the native speaker corpus**	**Raw frequency**	**Normalised frequency per 100,000 words**
1	Problem	57	27	1	Right (obligation)	65	71
2	Opinion	40	19	2	Idea	60	66
3	Idea	40	19	3	Opinion	54	59
4	Fact	28	13	4	Fact	49	54
5	Chance	27	13	5	Issue	47	52
6	-	-	-	6	Choice (preference)	39	43
7	-	-	-	7	Problem	33	36
8				8	Opportunity	32	35
9				9	Chance	26	28
		Total= 192	Total= 91			Total=405	Total= 444

It can be seen from Table 6-2 that there are considerable variations in both the normalised frequency and the choices of individual SSNs across the two corpora. Of the top-25 SNs that occur in the learner corpus, there are only five that are inherently SSNs. The number of inherent SSNs occurring among the top-25 SNs in the native speaker corpus is almost double that number. Another discrepancy between the two corpora is that there are many more tokens of SSNs found in the native speaker corpus than that of the learners. The table shows that the normalised frequency of SSNs in the native speaker corpus is almost five times higher than the normalised frequency of SSNs in the learner corpus. There are two possible explanations for this marked difference. One is that as learners of English the writers in the learner corpus are less familiar with the resource of SSNs than are native speakers. The other possibility is that although the learners may be familiar with the resource they may lack the level of expertise in writing that allow them to incorporate their stance on the entities and propositions that they write about using the resource of SSNs. The learners may not have reached that level of expertise in the target language and thus their use of SSNs is remarkably less (see Section 8.2.4.1 for further discussion). Thus, it can be argued that the resource of SSNs is a more defining feature of texts produced by native speakers of English than those produced by speakers of Arabic who are learning English.

It can be seen from Table 6-2 that the huge difference with regard to the frequency of individual SSNs in the learner corpus compared to the same SSNs in the native speaker corpus is accounted for by three SSNs: *opinion(s), fact(s) and idea(s)*. Thus, these SSNs will be closely examined in the following sections, starting with the SSN *opinion(s)*.

6.4.1. *Opinion(s)*

Opinion(s) is the second most frequently used SSN in the learner corpus and the third most frequently used SSN in the native speaker corpus. However, its normalised frequency per

100,000 words in the native speaker corpus is much higher than in the learner corpus. As shown in Table 6-2 presented earlier, the normalised frequency of *opinion(s)* per 100,000 words in the learner corpus is 19 while it is 59 in the native speaker corpus. The difference in the use of *opinion(s)* in the two corpora is significant[13] with the probability value of $p<0.0001$. The SSN *opinion(s)* is less represented in the learner corpus.

The difference between the learner corpus and the native speaker corpus in terms of the identified patterns that are found to occur with the SSN *opinion(s)* is also noteworthy. Table 6-3 below shows the patterns associated with the SSN *opinion(s)* as identified in the two corpora along with their raw frequency:

Table 6-3 *Patterns associated with the SSN opinion(s) as identified in the two corpora (* refers to an erroneous use of preposition and possessive pronouns found to occur once in the learner corpus)*

Patterns identified with the lemmatised SSN *opinion*	Frequency of the identified pattern in the learner corpus (%)	Frequency of the identified pattern in the native speaker corpus (%)
Opinion in adjunct phrase: ❖ *In/from*/on* my/mine* *opinion* + specifics: appositive clause ❖ Specifics: appositive clause + *in my opinion*	39 (97.5%) 1 (2.5%)	9 (17%) 3 (6%)
(Prep) + (*The/my*) + *opinion* + specifics: that appositive clause	0 (0%)	5 (9%)
My + *opinion* + v-link + specifics: that appositive clause	0 (0%)	1 (2%)
Poss. + *opinion* (*above*)	0 (0%)	11 (20%)

[13] I have used the log likelihood calculator created by Rayson (2017). The likelihood result obtained is -30.28. For more clarification how Rayson's calculator works and how to interpret obtained results, refer to Section 4.5 in Chapter 4.

Table 6-3 *(continued)*

(Prep.) *This/that* + (ADJ) *opinion*	0 (0%)	10 (18%)
These + *opinions*	0 (0%)	1 (2%)
It/ that + v-link + poss. + *opinion*	0 (0%)	2 (4%)
The + *opinion*	0 (0%)	5 (9%)
Quantifier + *opinions*	0 (0%)	3 (6%)
Prospecting verb + *opinions*	0 (0%)	3 (6%)
Such + *an* + *opinion*	0 (0%)	1 (2%)
	Total = 40 (100%)	Total = 54 (101%)

As shown in Table 6-3, *opinion(s)* is always found to occur in the following adjunct phrase pattern: *in/from/on my/mine opinion* in the learner corpus while in the native speaker corpus a number of patterns besides the fixed phrase: *in my opinion* are identified to occur with the SSN *opinion*. In other words, the native speakers use this SSN as the subject, object and complement of a clause as well as in the adjunct. The following are examples that show how the SSN *opinion* is used in the learner and native speaker corpora.

(72): In my **opinion** the prefect holiday is very important for me. **CEPA3**

(73): In my **opinion**, I'm without my room sim like fish without sea. **CEPA4**

(74): In my **opinion**, if you want to be a profisor, you should make the library as a second home for you. **CEPA5**

(75): In my **opinion**, I think movies should be more about moral things than immoral things which is shown in alot of movies. **CEPA6**

(76): I disagree with this **opinion** and feel that holding a part-time job during one's college days will negatively impact long term goals. **Native speaker**

(77): There is merit to the **opinion** that college students should have part time jobs. **Native speaker**

(78): Why do I hold such an **opinion** in today's age of scientific proof that so clearly shows how harmful second hand smoke is to others? **Native speaker**

The examples above illustrate the use of the SSN *opinion(s)* in the two corpora. They exemplify the students' use of *opinion* in the prefabricated pattern: *in my opinion + specifics: appositive clause*, as shown in example (72) to example (75). Conversely, they show the wider variety of patterns used by the native speakers, as illustrated in examples (76) to example (78). In these examples, the SSN *opinion* is used in the following patterns: *with + this + opinion* as in (76), *to + the +opinion + specifics: that appositive clause* as in (77) and *such + an + opinion* as in (78). This discrepancy between the two groups of writers, as shown in the aforementioned examples and by the figures presented in Table 6-2, clearly suggests that the speakers of Arabic who are learning English tend to use the SSN: *opinion* in its prefabricated phrase. This tendency can be explained by one possibility. The possibility is that the learners are less familiar with the SSN *opinion(s)* used as a noun outside the fixed phrase, so their use of it is less flexible than that of the native speakers. Hasselgren (1994) argues that learners of English even at advanced levels of proficiency cling to certain lexical items that they feel familiar and safe with. She called this phenomenon "lexical teddy bears". Hasselgren (1994, p. 237) offers two explanations for this phenomenon. One is that L2

lexical teddy bears are likely to stem from ‘words and phrases closely resembling [learner’s] first language’. The second is that they come about because they are ‘learnt early’.

Considering Hasselgren’s arguments above, the overdependence on the use of the SSN *opinion* in its fixed patterns as shown by the Arabic speakers who are learning English is a lexical teddy bear par excellence. In Arabic, the noun *opinion* is widely used in the fixed pattern “fy rāʾyy”, which is literally translated into English as “in my opinion”. Given this fact, the use of the SSN: *opinion* by the Arabic speakers who are learning English in the present study are more likely to have been influenced by their L1. This is reflected in these learners’ restricted repertoire of patterns that can be used with SSNs. For instance, the Arabic learners of English could have signalled their personal stance towards given propositions by using *opinion* with patterns like *my opinion + specifics: that appositive clause*, *my opinion + v-link + specifics: that appositive clause* and *it/ that + v-link + my + opinion*. These patterns are examples of the patterns used by the native speaker writers to refer to their own personal opinions.

A relevant study that has investigated the noun *opinion* is Tåqvist (2016). In this study, Tåqvist (2016, p.119) identifies two classifications of the noun uses in terms of reference; one is *self-reference* and the other is *other-reference*. The first category refers to instances where the noun is used to make reference to the writer’s personal opinion while the other category refers to cases when other opinion is referred to (Tåqvist, 2016). Following Tåqvist’s definition, the uses of the SSN *opinion(s)* with regard to reference are investigated below.

Table 6-4 shows how the two reference functions are used in the learner and native speaker corpora used in the present study in terms of raw frequencies and percentages.

Table 6-4 *Raw frequencies and percentages of reference uses of opinion across the two corpora*

Reference type	Raw frequency in the learner corpus (%)	Raw frequency in the native speaker corpus (%)
Reference to the writer's personal opinion	40 (100%)	35 (65%)
Reference to other opinion	0 (0%)	19 (35%)
	Total= 40 (100%)	Total= 54 (100%)

As shown in Table 6-4, all the uses of the SSN *opinion(s)* in the learner corpus is self-referring. On the other hand, both reference uses are found to occur in the native speaker corpus. However, a large majority of the instances of *opinion* is self-referring. The predominant use self-referring instances of *opinion* in the learner and the native speaker corpora may be partly explained by the topics that these two groups of writers are asked to write about. As noted in Chapter 4, both the learners and native speakers are asked to give their personal opinions, a point which may give rise to the self-referring instances of the SSN *opinion(s)*. Another possibility is that the noun 'opinion' in English is more likely to be used to refer to one's own opinion and not of others. For example, in the Bank of English (BoE) corpus, there are almost 2000 instances of "in my opinion" while there are fewer than 400 instances of "in his/her opinion".

Examples (79) to (82) are illustrative of the two uses of the SSN *opinion* in terms of reference; examples (79) and (80) illustrate self-reference uses of the noun, and examples (81) and (82) illustrate other-reference uses of the noun.

(79): In my **opinion**, Oman is the best country to visit it. **CEPA5**

(80): However, I am comfortable in stating my **opinion** that smoking is a terrible habit, **Native speaker**

(81): There is merit to the **opinion** that college students should have part time jobs. However it is important not to overlook the main purpose of a college education. **Native speaker**

(82): I disagree with the **opinion** that smoking should be completely banned from all restaurants in the country. **Native speaker**

In examples (79) and (80), the SSN *opinion* makes reference to the writer's personal opinion, illustrated by the patterns*: In my opinion + specifics: appositive clause* and *my + opinion + specifics: that appositive clause*, respectively. In examples (81) and (82), on the other hand, the noun is used to make reference to someone else's opinion (probably the writer of the writing prompt). In each of these examples (79) to (80), the writer's opinion is specified within clause boundary. This brings us to the next point about the SSN *opinion(s)* that I want to investigate, that is how the learner and native speaker corpora compare with each other with regard to the within and across clause construction of stance.

As illustrated in the first part of this chapter, the patterns associated with SSNs can determine as to whether stance is constructed within or across clause boundaries. When the SSN *opinion* is used in its fixed adjunct pattern, the specification of the evaluated proposition is always observed within a clause boundary, as illustrated in the following examples:

(83): In my **opinion** I think Salalah is more nice than anther city in the world. **CEPA4**

(84): In my **opinion**, smokers should be free to smoke if they choose to do so. **Native speaker**

(85): In my **opinion**, Oman is the best country to visit it. **CEPA5**

Example (83) is particularly interesting. The learner uses two stance expressions; *in my opinion* followed by *I think*. The learner could have used one expression or the other to construct an act of stance towards the given proposition: "*Salalah is more nice than anther city in the world*". The use of the double stance construction by the learner can be regarded as a redundant style in expressing stance. In fact, this repetition is found to account for 13% of all the uses of the SSN *opinion* in the learner corpus while it is non-existent in the native speaker corpus.

As shown in examples (84) to (85) above, the fixed pattern identified to co-occur with the SSN *opinion* can be easily identified as a within clause construction of stance. As mentioned before, the learners only use the SSN *opinion* in its fixed pattern and thus construction of stance through the SSN *opinion* in the learner corpus is limited to that of within clause construction. Unlike the learners, the usage of the SSN *opinion* in the native speaker corpus is not restricted to within clause construction of stance, as shown in the following examples:

(86): I completely agree with the statement that smoking should be banned from all restaurants in the country. The

reasons for my **opinion** are as follows: second hand smoke, distinctiveness of Japanese food and social etiquette.

(87): I agree with this **opinion**. I have three opinions.

(88): I have three **opinions**. First, if smoking is banned in all restaurants this may begin a trend in other establishments such as bars and nightclubs.

The native speaker writers use the SSN *opinion(s)* to construct stance beyond the clause boundary. Examples (86) and (87) are illustrative of cases where stance is constructed by means of encapsulation; the evaluated proposition is found to occur in the previous sentence. Example (88), however, shows an instance of stance that is constructed by means of prospection; the evaluated propositions are found to occur in the following sentences. Table 6-5 presents the frequency of within and cross clause construction of stance through the SSN *opinion(s)* in the two corpora:

Table 6-5 *Construction of stance with the SSN opinion in the two corpora*

Construction of stance through the lemmatised SSN *opinion*	Total raw frequency and the percentage in the learner corpus	Total raw frequency and the percentage in the native speaker corpus
Within clause boundary	40 (100%)	18 (33%)
Across clause boundary	0 (0%)	36 (67%)

The two corpora, as can be seen in Table 6-5, differ in terms of the construction of stance through the SSN *opinion*. It is apparent that the learners exclusively rely on within clause construction of stance. By contrast, the native speakers make use of both within and across clauses construction when they use the SSN *opinion*. However, there is a pronounced

difference between the two ways of stance construction in the native speaker corpus; the occurrences of across clauses construction of stance are double those of within clause construction.

The differences illustrated by Table 6-5 give us an insight into the learners' understanding of the SNN *opinion*. Their restricted use of *opinion* within a clause boundary indicates that the learners are less familiar with the noun being used in a variety of contexts. Thus, they do not seem to be aware of the potential that the SSN *opinion* has as a discourse organising item. It is only when the SSN *opinion* is constructed beyond clause boundary that the SSN *opinion* can be said to organise stretches of text. The conclusion might be that the use of the SSN *opinion* to organise text can be said to characterise writing by native speakers of English than that of novice learners of English.

6.4.2. *Fact(s)*

Fact(s) is the third most frequent SSN in the learner corpus. As is the case with *opinion(s)*, there is a substantial difference in the normalised frequency of this SSN between the two corpora. There are 13 occurrences of *fact(s)* per 100,000 words in the learner corpus whereas there are 54 instances of *fact(s)* per 100,000 words in the native speaker corpus. There is significant[14] difference between the two corpora in the use of the SSN *fact* with the probability value of $p<0.0001$. *Fact(s)* is a less represented SSN in the learner corpus compared to that of native speakers.

Similar to the SSN *opinion* discussed in the previous section, the identified patterns that are found to occur with *fact* are noteworthy, though not as striking as the figures for *opinion*.

[14] The likelihood result obtained from Rayson's tool (2017) is -36.96. Refer to Section 4.5 to understand how this result is to be interpreted.

Table 6-6 shows the patterns associated with the SSN *fact* as identified in the two corpora along with their raw frequency:

Table 6-6 *Patterns associated with fact in the two corpora*

Patterns identified with the lemmatised SSN *fact*	Raw frequency of the identified pattern in the learner corpus (%)	Raw frequency of the identified pattern in the native speaker corpus (%)
The + *fact* + specifics: that appositive clause	13 (46%)	24 (49%)
Fact in adjunct phrase: *In fact* + specifics: appositive clause	13 (46%)	8 (16%)
The + *fact* + specifics: other appositive clause (absence of that element)	2 (7%)	0 (0%)
The + *fact* + v-link + specifics: that appositive clause	0 (0%)	9 (18%)
The + *fact* + *v-link* + *specifics: other appositive clause (absence of that element)*	0 (0%)	1 (2%)
(Prep.) + *this* + *fact*	0 (0%)	3 (6%)
Those + *facts*	0 (0%)	1 (2%)
It/ this + v-link + *fact*	0 (0%)	2 (4%)
ADJ + *facts*	0 (%)	1 (2%)
	Total= 28 (99%)	Total= 49 (99%)

As illustrated by Table 6-6, there is a difference between the two corpora. The difference lies in the number of the patterns that are found to occur with the SSN *fact* and also in the frequency of the identified patterns. For instance, only three patterns associated with *fact* are found to occur in the learner corpus: *the fact + specifics: that appositive clause*, *the fact + specifics: other appositive clause* and the fixed adjunct pattern: *in fact + specifics: appositive clause*. The native speaker writers, on the other hand, are found to use the SSN *fact* in seven patterns. The identified patterns in the learner corpus are demonstrated in the following examples:

(89): In **fact** in this holiday it was worst holidy ever. **CEPA3**

(90): In **fact**, the beautiful place for me is Dubai… **CEPA3**

(91): My holiday it was very nise in **fact** it was awonder ful. **CEPA 4**

(92): In **fact**, this computer is the bist thing I got. **CEPA5**

(93): despite the **fact** that we were really happy, we were really sad… **CEPA5**

(94): It could be the **fact** that I'm getting older in mind… **CEPA6**

(95): I simply love the **fact** that the city remained the same and took pride of what they had… **CEPA6**

(96): Despite of the **fact** he gave him bisson but maximos killed him but after that maximos died because the bison killed him. **CEPA5**

The learners rely almost exclusively on both the fixed pattern *in fact* and *fact* followed by *that appositive clause*. The first four examples, (89) to (92) are illustrative of the SSN *fact* occurring in its adjunct phrase pattern. This pattern represents 46% of all the patterns identified to occur with the noun in the learner corpus. Examples (93) to (95) are examples of cases where the SSN *fact* is followed by that-appositive clause specifics. These cases also account for 46% of the total number of patterns occurring with the noun. Example (96) represents cases used by the learners where *fact* is followed by an appositive clause with the absence of the that element. These cases are rare, accounting only for 7% of the total.

Unlike the learners, the native speakers make use of a wide range of patterns when constructing stance with *fact*. For instance, the pattern: *The + fact + specifics: that appositive clause* is found to be the most frequent pattern. This pattern represents 49% of all the patterns used by the native speakers. In comparison to the learner corpus, the prefabricated pattern: *In fact + specifics: appositive clause* in the native speaker corpus is less commonly used. The pattern represents 16% of the overall patterns occurring with the noun in the native speaker corpus. The other remaining patterns in the native speaker corpus are non-existent in the learner corpus. The following examples illustrate some of the patterns that uniquely occur in the native speaker corpus:

(97): In light of this **fact**, I think that it is necessary for students to get part time job so that they may contribute to the economy in some small way. **Native speaker**

(98): The job-market is so cutthroat and students need to be aware of this **fact**. **Native speaker**

(99): Based on those **facts** alone, it is difficult to see how smokers feel so offended and picked on. **Native speaker**

(100): We smokers do not live in denial of scientific **facts**, we readily accept that smoking is not a healthy habit (it causes all sorts of lung and respiratory ailments). **Native speaker**

As shown in the above examples (97) to (100), a number of patterns are used to construct stance across clause boundaries. In examples (97) to (99), the pattern: *fact preceded by a deictic determiner* is used to evaluate a previously mentioned stretch of text. The pattern identified in example (100): *ADJ + facts* is used to evaluate propositions that are to appear in

the evolving discourse. Table 6-7 shows the frequency of within and across clauses construction of stance through the SSN *fact(s)* in the two corpora:

Table 6-7 *The construction of stance through the SSN fact in the two corpora*

Construction of stance through the lemmatised SSN *fact*	Total raw frequency and the percentage in the learner corpus	Total raw frequency and the percentage in the native speaker corpus
Within clause boundary	28 (100%)	42 (86%)
Across clause boundary	0 (0%)	7 (14%)

The two corpora, as shown in Table 6-7, are similar in that they both rely extensively on the within clause construction of stance through the SSN *fact*. However, the two corpora are strikingly different with regard to the use of the noun to express stance beyond clause boundaries. This way of stance construction is non-existent in the learner corpus. This finding is similar to that identified earlier when the SSN *opinion* is discussed. The native speaker corpus, by contrast, consists of instances where across clause construction of stance using the SSN *fact* is observed. However, these instances are markedly different compared to the instances of within clause construction of stance; the occurrences of *fact* being specified within clauses are six times higher than that of across clauses.

6.4.3. *Idea(s)*

Another striking difference between the two corpora is apparent in the usage of the SSN *idea*. The normalised frequency of *idea* is indicative of this discrepancy. There are 19 occurrences of *idea* per 100,000 words in the learner corpus while in the native speaker there are 66

instances of *idea* per 100,000 words. The difference is significant[15] with the probability value of p<0.0001. Thus, *idea* is less represented in the learner corpus compared to that of native speakers.

Unlike the two previously discussed SSNs, the learners are found to use *idea* with a wide range of patterns. Table 6-8, below, lists all the patterns identified to occur with *idea* in the learners and native speaker corpora:

Table 6-8 *Patterns identified to occur with the SSN idea(s) in the learner and native speaker corpora*

Number	Patterns identified with the lemmatised SSN *idea*	Raw frequency of the identified pattern in the learner corpus (%)	Raw frequency of the identified pattern in the native speaker corpus (%)
1	*The* + *idea* + specifics: other appositive clause (of appositive clause)	4 (10%)	6 (10%)
2	*Idea* + specifics: other appositive clause (to-infinitive appositive clause)	2 (5%)	19 (31 %)
3	*Idea* + specifics: that appositive clause	1 (2.5%)	6 (10%)
4	Specifics + v-link + *idea*	3 (7.5%)	17 (28%)
5	*This/that* + *idea*	4 (10%)	5 (8%)
6	*It/his/that* + v-link + determiner + (ADJ) *idea*	10 (25%)	1 (2 %)
7	*Idea* + specifics: prepositional phrases	2 (5%)	1 (2%)
8	*The* + *idea*	3 (7.5%)	4 (7%)

[15] The likelihood result obtained from Rayson's tool (2017) is -38.34. Refer to Section 4.5 in Chapter 4 to understand how this result is to be interpreted.

Table 6-8 *(continued)*

9	*The* + *idea* + specifics: other appositive clause (Wh- appositive clause)	3 (7.5%)	0 (0%)
10	*Idea* + specifics: other appositive clause (about + appositive clause)	3 (7.5%)	0 (0%)
11	*Idea* + v-link + specifics: other appositive clause (gerund clause)	1(2.5%)	0 (0%)
12	Poss. + *idea*	1 (2.5%)	0 (0%)
13	Prospecting verb + *idea*	3 (7.5%)	0 (0%)
14	*These* + *ideas*	0 (0%)	1(2 %)
		Total= 40 (100%)	Total= 60 (100%)

The patterns displayed in Table 6-8 show that the learners use more patterns when constructing stance through the noun *idea(s)*. The number of patterns identified in the learner corpus is 14 while there are 10 patterns in the native speaker corpus. This indicates that the Arabic learners who are learning English are familiar with the SSN *idea(s)* as an ordinary noun that can be used in a variety of contexts. Thus, the learners' repertoire of the possible patterns that can accompany the SSN *idea* is not restricted. This finding is not in line with my earlier findings about the SSNs *opinion* and *fact* where the learners' use of these SSNs is found to occur in very limited patterns.

As shown in Table 6-8, the first eight patterns are used by the two groups of writers. The subsequent patterns are either found to occur in the learner corpus: 9 to 13 or in the native speaker corpus: 14. Overall, the frequency of the shared patterns in the two corpora shows

that native speaker writers make use of these patterns more. For instance, the native speakers rely extensively on two patterns: *Idea + specifics: other appositive clause (to-infinitive appositive clause)* and *Specifics + v-link + idea*. These patterns account for 32% and 28.33% of all the identified patterns in the native speaker corpus respectively. Conversely, these two patterns are used strikingly less in the learner corpus with the former accounting for 5% and the latter represents 7.5% of all the identified patterns in the corpus. The following examples are illustrative of these patterns:

(101): In the afternoon we came back home. with an **idea** to go thier agien. **CEPA4**

(102): For these reasons, I think it is a good **idea** to ban smoking at restaurants in Japan. **Native speaker**

(103): Going to Eygpt is great **idea** for having a great summer holidy. **CEPA6**

(104): Part time jobs for college students are a great **idea** and… **Native speaker**

The evaluated propositions, as shown in examples (101) and (102), are spelled out in the to-infinitive clause that immediately comes after the SSN category. The other two examples, (103) and (104), are illustrative of the pattern: *specifics + v-link + idea*. When this pattern is used, the construction of stance within a clause boundary can be said to be marked; i.e. the evaluated discourse precedes the SSN category. This marked usage of the SSN *idea* can be said to distinguish learners' writing from that of native speakers'.

The two groups of writers are found to construct stance through *idea(s)* within and across clause boundary. Table 6-9 provides the total number of both ways of stance construction as found in the two corpora:

Table 6-9 *The construction of stance through idea across the two corpora*

Construction of stance through the lemmatised SSN *idea*	Total raw frequency and the percentage of in the learner corpus	Total raw frequency and the percentage in the native speaker corpus
Within a clause boundary	20 (50%)	49 (82%)
Across clause boundary	20 (50%)	11(18%)

Table 6-9 shows that there is a pronounced difference in the way the SSN *idea(s)* is used in the two corpora. The learners equally use the two ways of stance construction while the native speakers' usage of the SSN shows a marked preference for the within sentence construction of stance. The finding that the native speakers tend to construct the SSN *idea* within a sentence boundary is also found to be the case with the uses of the SSN *fact* in the native speaker corpus (see Table 6-7). The following are two sets of examples that illustrate cases of within and across clause construction of stance using the SSN *idea* as found in the two corpora.

❖ **Within-clauses**

(105): They knowed that we are from the Gulf, and they have an **idea** that all the gulf people are rich. That wasn't true. **CEPA 5**

(106): I remember a man passing very near our table but have no **idea** how he stole it so easily. **CEPA 6**

(107): The main **idea** of the movie is kidnapping and rapping the children of wealthy families. **CEPA 6**

(108): I think it is a great **idea** for college students to have a part time job. **Native Speaker**

(109): I think that it would probably be a pretty good **idea** to completely ban smoking in all the restaurants in Japan. **Native speaker**

(110): Hefty penalties put many businesses off the idea of allowing people to smoke. **Native speaker**

❖ **Across-clauses**

(111): I started with my sister to change his **idea**. **CEPA5**

(112): We thought that was agrate **idea** and in the next day we headed there… **CEPA5**

(113): I use my diary on a daily basis to write in my daily events, in order to look back through its pages when I grow older, as I think it is a nice **idea** to help me recall my memories, the good and the bad. **CEPA6**

(114): All of these **ideas** are equally valid in any country, so I think that smoking should be banned at restaurants in Japan. **Native speaker**

(115): Too many people these days think only of what they can take a nothing of what they can give. As John F. Kennedy once said, "Ask not what your country can do for you but

what you can do for your country.” The same **idea** applies to schools, and student should be active members of the college community in order to be successful community members. **Native speaker**

(116): the majority of who are smokers, weren't too keen on that **idea** and cracked it down. **Native speaker**

The examples above illustrate some of the patterns that are associated with the stance function that *idea* serves in discourse. The first set of examples shows patterns that can be said to indicate that stance is to be communicated within the same sentence where the SSN category is used. The second set of examples demonstrates the patterns that are when found to occur with *idea* the stance construction is to go beyond a sentence boundary. A closer look at the cross clauses construction of *idea* in the two corpora indicates that it is more likely that the stance construction is done by means of encapsulation where *idea* is used to evaluate a preceding stretch of discourse than by that of prospection. In fact, all instances of cross clauses construction of *idea* in the native speaker corpus is fulfilled by the reference act of encapsulation whereas in the learner corpus 17 out of 20 instances of *idea* constructed across clauses are found to evaluate an already stated proposition by means of encapsulation. The three instances where idea is found to prospect in the learner corpus are shown below:

(117): I siggest agood **idea** it is to go my aunt and my brother to bring car,so every bady agree with me.

(118):...and my mother got an **idea** which was the best. she said that the best way to spend the summer is to go somewhere overseas.

(119): I suggested an **idea** to my family and they loved it. Unfortunately, fishing was not exactly the best way to spend your holiday.

6.5. Summary of the chapter

This chapter has investigated the phenomenon of SSNs with respect to two main aspects. The first part of the chapter has focused on how stance is constructed in the learner and native speaker corpora. The second part of the chapter has focused on how SSNs are used in the two corpora.

The examination of SSNs in the first part of the chapter has broadened our understanding of how stance is constructed through SSNs in three ways. First, it has been found that stance through SSNs can be constructed either within or across clause boundary. The most frequent patterns associated with these two categories of construction have been identified. What this part has demonstrated is the advantage of not limiting the search of SSNs to specific patterns if one is to understand how stance is constructed through SSNs.

The investigation of SSN usage in the learner and native speaker corpora in the second part has shown that there are many differences across the two corpora. One remarkable difference is that the learners use the resource of SSNs less. The patterns used with *opinion, fact* and *idea* as well as the two ways of stance construction of these nouns also show variation. The learners, unlike the native speakers, tend to exclusively use the SSN: *opinion* in its fixed phrase pattern: *in my opinion + specifics: appositive clause* and also tend to almost rely on *in fact + specifics: appositive clause* when they use the SSN: *fact*. As a result, the stance construction through these two SSNs in the learner corps is always signalled within clause

boundaries. However, this is not the case with the SSN *idea* where the learners make use of a wider range of patterns with the noun, a point which have enabled them to use the noun to construct stance within and across clause boundaries.

The findings of this chapter will be returned to in Chapter 8. The next chapter will be concerned with SNs as a resource to signal causal meanings in discourse.

CHAPTER 7: SNs AND CAUSALITY

7.1. Overview

This chapter is concerned with SNs as a potential resource for establishing causal relations among propositions expressed in texts. The first part of the chapter briefly provides a definition of the concept of causality and how this definition relates to causal SNs (CSNs henceforth). The second part gives an account of the findings of CSNs in the two corpora used in the current study. The findings are presented in two main sections. The aim of the first section (7.3) is to explore how causality through CSNs can be constructed and also to identify the patterns associated with the construction of causality. The second section (7.4) compares the two corpora with each other in terms of the use of CSNs. The chapter ends with a summary of the findings and concluding remarks. This is provided in Section 7.5.

7.2. The concept of causality

Within the discipline of Philosophy, causality is a concept surrounded by a great deal of debate and technical terminology (e.g. Russel, 1913; Pearl, 2000; Illary and Russo, 2014). This chapter, however, takes a more general view of the concept, using Schmid's definition of it as "a relation between two events or facts, in which one entity, the CAUSE, is conceived by a human conceptualizer as being prior to, contributing to and/or being responsible for the occurrence or existence of another, called EFFECT" (Schmid, 2000, p.102, emphasis in original). To establish the causal relation between the element of cause and the element of effect, writers can draw on observable linguistic resources such as conjunctive adverbs, verbs and nouns. Before presenting illustrative examples of these resources and before moving further into the chapter, important points need to be addressed. In any causal relation, there would be an element of 'cause' and an element of 'effect'. Some words tell us about the cause such as the verb 'cause', the noun 'cause' and the noun 'reason' while some other

words tell us about the effect such as the nouns ‘effect’, ‘result’ and ‘consequence’. Since the focus of the present thesis is on SNs, it will be noticed that the focus is placed on nouns that tell us about either the cause or the effect.

The following are examples of the three resources capable of establishing causal relations between propositions, as mentioned above. Depending on what the signalling element is, it can be realised/specified as either the cause or the effect. As displayed in the following examples, the causal signalling element is **in bold**; the specification of the signalled causal element is underlined.

(1): *well on that day.. I got the chicken box.. Oh yes I was sick and had dots all over my face. I had to stay at Grama’s house while my family were having fun in London*. **Thus** I decided to tag this as the worst holiday I ever spent in my whole life! **CEPA 6** (Causal element is effect)

(2):...accidently she triped over a rigid stone that **caused** *her to fall off and unfortunatelly break her leg*. **CEPA6** (causal element is cause)

(3): The **reason** *why I called this my perfect holiday* was because I got to spend great time with my family. **CEPA6** (causal element is cause)

As illustrated in the above examples, the causal relation between the two propositions are signalled by an adverbial conjunct *thus* in example (1); a causal verb *cause* in example (2) and a causal noun *reason* in example (3). The use of *thus* in example (1) signals that what is coming in the discourse is an effect/result of a stated causal events previously presented. The

causal element in example (2) and (3), on the other hand, is the cause event. The causal element in example (2) is signalled by a verb and by a noun in example (3).

As demonstrated above, adverbs, verbs and nouns can signal causality. However, only causal-signalling nouns (CSNs) will be investigated since the current thesis is only concerned with the phenomenon of SNs of which CSNs are classic examples. More particularly, the focus will be on inherent CSNs found to occur in the two corpora used in the current study. By inherent CSNs, I mean those SNs that are inherently causal—their use is largely understood to that of signalling causality such as 'reason', 'cause', and 'result'. The SN *thing*, for example, is sometimes found to signal causality as shown in the following example. However, the SN *thing* is not an inherent CSN and thus is not examined in this chapter. This SN is investigated in Chapter 5.

(4): The **thing** that made me like the movie very much is the actors that were in it and because the movie contained both funny scenes and action scenes so that the movie will not be boring. **CEPA 6**

7.3. Patterns and CSNs

As mentioned previously, one aim of this chapter is to investigate the construction of causality through the resource of nouns. To this end, I have focused on the inherent CSNs identified in the two corpora, as shown in Table 7-3 presented later in section 7.4. The concordance lines of these nouns are thoroughly examined to identify the patterns that frequently occur with CSNs. The main finding is that there are many patterns that can be said to accompany CSNs, which in turn can indicate where to find the specification of the causal signalling element. This finding will be addressed in the next sections.

Before proceeding to the patterns, important methodological points should be noted. Similar to the points reported in Section 6.3 that dealt with the patterns associated with SSNs, I have chosen to focus on the frequent patterns associated with CSNs and have also followed Flowerdew and Forest's (2015) classification of patterns (within clause boundary and across clause boundary). All the patterns that occur with CSNs three times or more are identified as frequent.

Another important point is concerned with the representation of CSNs patterns. It will be noticed from the examples used to illustrate the identified patterns throughout this section that both the 'effect' and the 'cause' in many cases are present as shown in the following example:

(5): I guess the main **reason** *why I'd choose Boston* is because I grew up there, and miss it terribly. **CEPA6**

In example (5), both the effect (shown in italics) and cause elements (underlined) are present. The possible pattern that can represent the example is **CSN + Wh effect clause + v-link + specifics: because appositive clause**. However, since my focus is on CSNs, I only include CSNs and their specifications in the pattern. So, the pattern representation I am using for example (5) is **CSN + v-link + specifics: because appositive clause**.

I will now move on to present the patterns associated with CSNs, starting with within clause boundary patterns.

7.3.1. Signalling of causality within clause boundary

Causality through nouns can be constructed within a clause boundary. This means that the causal category (CSN) and its signalled causal entity are to be found within the clause where the CSN occurs. The most frequent patterns associated with within-clause construction of causality through CSNs, based on the definition noted earlier, are presented below.

7.3.1.1. CSN + v-link + specifics: that-appositive clause

This pattern can be used to signal causal meanings within a clause boundary. This means that the specifics of the CSN is found within the same clause. The that-appositive clause in the pattern lexicalises the signalled causal meaning as shown in the following examples taken from the two corpora:

(6): The first **reason** is that it is the house of allah where we can be ask allah what we want and forgiveness. **CEPA 5**

(7): The **reason** I love this necklace is that it showed me that I have two more sisters in this world that both love and cherish me... **CEPA 6**

(8): The final **reason** why I think smoking should be banned is that the attempts to separate smokers and non-smokers in restaurants in Japan are completely useless. **Native speaker**

(9): The fact of the matter is the standards for college graduates have been rising for some time, in one **result** of this is that students must accomplish more in the same amount of time as they would have had to in the past. **Native speaker**

As can be seen from the above examples, the writers use the CSNs nouns: *reason and result* to signal causal meanings: stretches of language that are seen as the reasons for given effects as in examples (6) to (8) or as the effects of given causes as in example (9).

The examination of the inherent CSNs in the two corpora indicates that *CSN + v-link + specifics: that-appositive clause* is found to occur with only two CSNs: *reason* and *result*, as shown in the above examples. The reason why this pattern is not observed with the other two CSNs: *effect* and *cause* is possibly explained by the fact that these nouns are the least used CSNs, as illustrated in Table 7-3 presented in Section 7.4 below. Evidence from a corpus of general English (the BNC) shows that these nouns in general can be used in the pattern to signal causal meaning within a clause boundary. However, these nouns are less frequently used in this pattern in the BNC compared to the CSNs: *reason* and *result*. The total instances of *reason* and *result* occurring in this pattern in the BNC are found to be 264 and 301, respectively. However, the total instances of *cause* and *effect* occurring in this pattern are strikingly fewer; there are 62 instances of *effect* while *cause* is found to only occur 6 times.

Thus, this pattern is one of the ways that causal meanings are constructed within clause boundaries.

7.3.1.2. CSN + v-link + specifics: other appositive clause

As mentioned in the previous chapter dealt with the construction of stance through signalling nouns, appositive clauses providing the signalled stance are not limited to that-appositive clauses when stance is expressed within a clause boundary. The same finding holds true for causality expressed through nouns within a clause boundary. The appositive clause providing the specification of the signalled causal meanings is found to occur in a range of grammatical choices such as an appositive clause with the absence of **that** element, to-infinitive appositive clause, because appositive clause.

The following examples (10) to (14) are illustrative of the aforementioned options available when constructing causality through appositive clauses other than that-appositive clause:

(10): My first **reason** is there were to many kids. What I mean by that is the people there were very imiture. **CEPA 6** (CSN + v-link + specifics: appositive clause with the absence of that element)

(11): The only **reason** anyone would really have a part time job is to make money so they can go out at night and party with their friends. **Native speaker** (CSN + v-link + specifics: to-infinitive appositive clause)

(12): The second **reason** is to enjoy our vacation there and have a really quality time there. **CEPA 5** (CSN + v-link + specifics: to-infinitive appositive clause)

(13): I guess the main **reason** why I'd choose Boston is because I grew up there, and miss it terribly. **CEPA6** (CSN + v-link + specifics: because appositive clause)

(14): I disagree with this statement. One of the main **reasons** is because it creates billions of revenue for the government that can be then recycled and used to restore old buildings or some other areas of need in our cities. **Native speaker** (CSN + v-link + specifics: because appositive clause)

7.3.1.3. Specifics + v-link + CSN

Unlike the other patterns identified earlier where the CSN specifics occur after the v-link in the complement position of the clause, the CSN specifics in this pattern precede the causal

signalling category. The examination of the inherent CSNs in the current study has revealed that this pattern is only found to occur with *reason* and *cause* as illustrated below in examples (15) to (19):

(15): Entertainment is one of the **reasons** I use my computer. **CEPA 5**

(16): And Tom cruise also was the **reason** of making the film successful. **CEPA 5**

(17): In addition, having the freedom to enjoy a drink, or having a cigarette are added **reasons** why we pay to support restaurants. **Native speaker**

(18): Safety is another **reason** why smoking should be banned. **Native speaker**

(19): It has been proven that smoking is the number one preventable **cause** of death in the world and still, smokers don't seem to care. **Native speaker**

The causal signalled propositions in the above examples can be said to be marked since the prototypical specification of CSNs when constructed within clause boundary is to follow the CSN category (e.g. Flowerdew and Forest, 2015, p.20). In fact, this marked pattern is the third least frequent pattern of all the within clause patterns identified to occur with CSNs in the two corpora. However, the use of this pattern as shown in examples (15) to (19) can be said to emphasise the causal stretch of language being signalled.

7.3.1.4. CSN + v-link + specifics: noun phrase

This pattern is the second least used pattern in the two corpora. The pattern is only found to occur with the CSN *reason* as shown in the following examples:

(20): The **reason** for this is the danger of secondhand smoke. **Native speaker**

(21): the **reason** why I where late was the poor on the stairs, **CEPA 5**

7.3.1.5. Patterns of CSNs occurring in adjunct phrases

There are two patterns identified with CSNs when occurring in adjunct phrases. These patterns are listed below:

- Specifics: appositive clause + CSN in adjunct phrases
- CSN in adjunct phrases + specifics: appositive clause

The two patterns are found to occur only with two CSNs: *effect* and *result*. The latter CSN is the most frequently used in these patterns. In the first pattern (specifics: appositive clause + CSN in adjunct phrases), the specification of the CSN precedes the CSN category and thus can be said to be marked. The following are illustrative examples of this pattern:

(22): Also, I feel that I've become more responsible as a **result** of working part time... **Native speaker**

(23): Statistics have shown that over one hundred thousand people die every year as the **result** of passive smoking. **Native speaker**

(24): The liver began to develop sores as a **result** of trying to filter out all the harmful chemicals from the blood. **Native speaker**

In the second pattern (CSN in adjunct phrases + specifics: appositive clause); however, the specification appears after the CSN category. The following are illustrative cases of this pattern:

(25): First of all, smoking is extremely dangerous for the health of everyone around it. As a **result**, I do not think it is fair for the habit of a few to affect the health of many. **Native speaker**

(26): Too many'non-smoking' sections of restaurants are arbitrarily, haphazardly designed. In **effect**, smoke drifts from smoking to non-smoking sections, harming all its way. **Native speaker**

(27): For example, I work at the rec center, and as a **result** I spend a lot of time thinking about health and fitness and how to become stronger, faster, and more fit. **Native speaker**

7.3.1.6. CSN + specifics: comparative

This pattern is the least frequent of the within clause patterns. When CSNs occur in this pattern, they are specified by means of exemplification. Examples (28) and (29) show this:

(28): Also, there are the many negative **effects** of smoking including lung cancer, decreased lifespan, and many other symptoms. **Native speaker**

(29): The all muslime loved to death to go there for many **reasons** like: Al Haj and Al_Omra. **CEPA 3**

Having examined the patterns associated with the signalling of causality within a clause boundary, I will now move on to examine cases when CSNs are signalled in a way in which the causal signalled entity goes beyond clause boundary. The following section deals with across clause construction of causality through CSNs.

7.3.2. Signalling of causality across clause boundary

Causality through CSNs can be used to signal causal meanings across clause boundary. This means that the causal category (CSN) and its specifics are to be found beyond the clause where the CSN category occurs either by means of encapsulation or prospection. When CSNs serves these two functions, the causal specifics that have been signalled by the CSN are either in the preceding or the following discourse. Consequently, the specifying stretch of discourse in relation to the causal category (CSN) is tied in with the notion of encapsulation and prospection proposed in Sinclair's model of text organisation (Sinclair, 1993). The most frequent patterns associated with across clause specification of CSNs are presented below.

7.3.2.1. Patterns associated with CSNs when encapsulating causal meanings

There are many patterns which signal that the causal specification of the CSN is to be found in a previously mentioned stretch of discourse by means of encapsulation. These patterns are presented below with illustrative examples taken from the two corpora of study.

7.3.2.1.1 (*For*) *these* + CSN

The CSN found to occur in this pattern can be used to encapsulate more than one causal meanings stated previously. The following examples are illustrative of CSNs found to occur in this pattern:

(30): There are a number of reasons why I believe college students should have a part time job. Firstly, students need to learn the value of money. Throughout their time at school they are fully dependent on their parents and grandparents for all their financial needs. Many youngsters...Secondly, having a part time job allows students to develop practical skills, for example communicating with a variety of age groups, teamwork, dealing with problems, business etiquette, and using formal spoken language. These are skills which students may have studied about, but not had a chance to put into practice. Thirdly, students in a part time job can gain valuable experience in their chosen (future) field of employment ... Having prior work experience may be the key to securing a good job. In addition, students can develop professional contacts, leading to opportunities of full time employment in the same organization upon graduation. For these **reasons**, I believe that having a part time job is very important for students. **Native speaker**

(31): There are many reasons why that specific holiday was the best. Firstly, we travelled as a big family with our cousins. Secondly, we went to Italy which was and still is the best country I have ever been to. Last but not least, Angelo and his Family who were extremely kind to us. I think these **reasons** altogether made this trip the best vacation I have ever had. **CEPA 6**

As shown in the above examples, the pattern *(for) these + CSN* is used to encapsulate all the causal meanings (reasons) mentioned in the previous clauses. Although this pattern is found to frequently occur with the CSN *reason*, it is less likely to occur with other CSNs, at least in the two corpora of the study. The pattern is found to occur once with the CSN *result* and is non-existent with the CSNs *effect* and *cause*.

7.3.2.1.2 (*For*) *this/ that* + CSN

Another pattern associated with CSNs with the function of encapsulating previously mentioned causal meanings is *(for) this/ that + CSN*. Unlike the previous pattern: *(for) these + CSN* whose function is to encapsulate more than one causal meaning, CSNs occurring in the pattern: *(for) this/that + CSN* are found encapsulate only one causal meaning as shown in the following examples:

(32): Smoking is a filthy and disgusting habit. For this **reason**, it is my sincere hope that one day smoking will become much rarer than it is now. **Native speaker**

(33): But something was happened, my passport was lost. I was really angry and confused For this **reason** we stayed at home. **CEPA 5**

As shown in examples (32) and (33), the CSN *reason* is preceded by the deictic *This* to encapsulate one causal meaning that is seen to be the reason or the cause of given effects. The other inherent CSNs: *result*, *cause* and *effect* are not used in this pattern in the two corpora of the study.

Having consulted the BNC corpus, it seems that the CSNs: *result* and *effect* are used quite frequently in the pattern. *This + effect* occurs 496 times while the occurrences of *This + result* are 287. The use of *This + cause* in the BNC, however, indicates that this pattern is the

least used pattern with the CSN *cause;* there are only 37 occurrences of this pattern. Thus, it can be concluded that *This + CSN* is a potential pattern which can be used to construct causal meanings beyond clause boundaries with all inherent CSNs as identified in the two corpora with the exception of the CSN *cause*.

7.3.2.1.3 *It/this/that* + v-link + CSN

It/this/that + v-link + CSN is another pattern associated with CSNs with the function of encapsulating a previously mentioned causal meaning. This pattern is found to occur with only two CSNs: *reason* and *cause*. The pattern is illustrated in the following examples.

(34): Finally we arrived to the Global village . But we didn't see any body but my father park the car. he went to ask the guard of the village and he said to him that the village was clossed. Then we had been very disapointed .At the end we returned to our home. Every body was very sad. So It was the **reasons** to be the worst trip I have ever had. **CEPA 4** (It + v-link + CSN)

(35): Firstly, they will be gaining real business experience which is an added plus when students search for jobs after graduation and it is one of the **reasons** why I believe that part-time jobs are important. **Native speaker** (It + v-link + CSN)

(36): The largest and most undisputable reason is that smoking shortens people's lives. It is definitely the largest **cause** of lung cancer in the world and is a plague on mankind. **Native speaker** (It + v-link + CSN)

(37): Also, all the boys I talked to say that they do not like to kiss girls who smoke, so I think that is a good **reason** to have abstained from smoking too. **Native speaker** (That + v-link + CSN)

(38): All my favourite actors were in the movie and that's one of the **reasons** that encouraged me to go watch it. **CEPA 6** (That + v-link + CSN)

(39): of course, there are still restaurants where this set-up is non-existent, and this is undoubtedly a **cause** for grievance among non-smokers and smokers who prefer a clean environment while eating. **Native speaker** (This + v-link + CSN)

7.3.2.1.4 Patterns including the word *above*

There are two patterns that include the word *above*. These patterns are listed below:

- *For/based on* + *the* + *above* + CSN
- *For* + *the* + CSN + v-ed + *above*

These two patterns are found to occur only with the CSN *reason*. The following examples illustrate these patterns:

(40): For the above **reasons**, I agree 100% with the statement. **Native speaker** (*For* + *the* + *above* + CSN)

(41): Based on the above **reasons**, I support the opinion. **Native speaker** (*Based on* + *the* + *above* + CSN)

(42): Either way, I think that all students should have a part-time job for the **reasons** mentioned above and... **Native speaker** (*For* + *the* + CSN + v-ed + *above*)

(43): For the **reasons** stated above, I believe that a college student needs the experience of part-time work. **Native speaker** (For + the + CSN + v-ed + above)

As demonstrated in the above examples, the CSN *reason* serves to signal causal meanings by means of encapsulation. The specification of the CSN *reason* in each of these examples is to

be found throughout the previous parts of the whole essay where the CSN occurs. It is for this reason that the stretch of language specifying the CSN in the above examples is not included.

7.3.2.2. Patterns associated with CSNs when prospecting causal meanings

CSNs serve the function of prospection when occurring in particular patterns**.** Compared with the encapsulating CSN patterns described above, prospecting CSNs are found to occur with fewer patterns. These patterns are presented below.

7.3.2.2.1 Patterns with quantifier

There are four patterns identified to occur with CSNs when functioning to prospect causal meanings in text. These patterns are listed below:

- *There* + v-link + quantifier + CSN
- *For* + quantifier + CSN
- *Because of* + quantifier + CSN
- *I* + *have* + quantifier + CSN

The above patterns are the most frequent prospecting patterns identified to occur with CSNs in the two corpora of the present study. More specifically, these patterns are found to only occur with the CSN *reason*. The patterns are illustrated in the following examples:

(44): There are several **reasons** why this is a good idea. Firstly, a college education is a significant expense for a college student and their family. **Native speaker** (*There* + v-link + quantifier + CSN)

(45): There are many **reasons** why that specific holiday was the best. Firstly, we travelled as a big family with our cousins. **CEPA 6** (*There* + v-link + quantifier + CSN)

(46): I like it very much and it is my favorite for many **reasons**. First,It helps me in my homework. Second, I can chat with my friends and exchange messages with them. **CEPA 5** (For + quantifier + CSN)

(47): I agree with the above statement for a number of **reasons**. The most important reason however, is that I believe that anyone who goes to work, even part time, will better learn the value and responsibility of money and how hard people have to work to get it. **Native speaker** (For + quantifier + CSN)

(48): I have three **reasons** to support my position. Firstly, regarding the health of customers, more and more compelling evidence suggests that the ingestion of second hand smoke is a direct cause of various lung disorders, predominant of which is lung cancer. **Native speaker** (I + have + quantifier + CSN)

(49): Its special for two **reasons**. First, it was the last gift I got from my grandfather before he passed away. Second because it is supernatural. Shocked? Allow me to explain. **CEPA6** (For + quantifier + CSN)

The examples above are illustrative cases of the patterns involving quantifier occurring with the CSN *reason*. The CSN found in these patterns serves to prospect that more than one causal meaning is to be specified in the subsequent discourse. Examples (44) to (47) illustrate instances where the CSN: *reason* is pre-modified by vague quantifiers: *several*, *many* and *a number of*. Whenever that is the case, the causal meanings that are to be specified in the text to come are not exactly predicted as the writer could make mention of two, three or even more causal meanings. Examples (48) and (49) are , however, the opposite. By using a precise quantifier preceding *reason* in each of these examples, the number of the causal meanings that are to be present in the coming text are often predicted.

As mentioned previously, of the four inherent CSNs identified to occur in the two corpora of the current study, *quantifier + CSN* is uniquely associated with *reason*. The pattern is not found to occur with other CSNs: *cause*, *effect* and *result*. Thus, the pattern is less likely to occur with these CSNs, at least in the two corpora under investigation. To rule out as to whether these CSNs are less likely to be preceded by quantifiers in general compared to the CSN *reason*, the use of these CSNs in the BNC is examined. The results are shown in Table 7-1 below:

Table 7-1 *The frequency of quantifier + CSN in the BNC*

Observed pattern: *quantifier*[16] *+ CSN*	**BNC corpus**
Quantifier + reason	838
Quantifier + cause	51
Quantifier + result	30
Quantifier + effect	59

As illustrated in Table 7-1 above, the pattern: *quantifier + CSN* is found to frequently occur with the CSN *reason* and is less likely to occur with the other CSNs. There are 838 occurrences of the CSN *reason* preceded by a quantifier. Compared to the CSN *reason*, the other CSNs occur strikingly less: *cause* (51), *result* (30) and *effect* (59). These findings strongly support the findings proposed earlier with regard to the use of the pattern in the two corpora of the current study. Thus, it could be concluded that the pattern *quantifier + CSN* is more likely to be associated with the CSN *reason* when prospecting that causal meanings are to be supplied in the ensuing discourse.

[16] I have chosen the following quantifiers in my enquiry: **several, many, two and three.**

7.3.2.2.2 *For + the + following* + CSN

For + the following + CSN is another pattern that is found to occur only with the CSN: *reason*. When the pattern is observed, the CSN *reason* signals that the causal meanings (reasons) are to be mentioned by means of prospection as shown in the following examples:

(50): I do support the idea of college students working part-time, or even full-time for that matter for the following **reasons**. To begin, I believe by working it teaches students many valuable lessons about life that will also be a part of his/her educational process. **Native speaker**

(51): Smoking in restaurants should be allowed for the following **reasons**: First, smoking is legal in Japan, second smoking and food and drinks go hand in hand, and third, smoking should be left to the businesses. **Native speaker**

This pattern is only found to occur with the CSN *reason* in its plural form, as shown in examples (50) and (51). The BNC corpus lends strong support to this finding. In the BNC, the pattern: *For + the + following + CSN* is only found to occur with the CSN *reason* and is non-existent with the other CSNs: *cause*, *result*, and *effect*.

7.3.2.2.3 Prospecting verb + CSN

Like the previous pattern, this pattern is only found to occur with the CSN *reason*. As noted previously in Chapter 6, prospecting verbs are called so because they indicate that something is likely to be talked about. Examples (52) to (54) present cases of prospecting verbs occurring with the CSN *reason*.

(52): Which brings me to my second **reason**: no self-respecting chef should allow their culinary creations to be ruined by cigarette smoke. **Native speaker**

(53): In this essay I'll discuss the **reasons** that made me enjoy the movie. The movie talks about the life of a geisha in Japan, which is a bit mysterious to us. It reveals the life of a geisha and what it takes a normal japanese girl to become a geisha. **CEPA 6**

(54): Before I address this question I will give my **reasons** why I believe smoking should be banned in ALL restaurants. To begin, let's start with the obvious. It has been scientifically proven that second hand smoke has enough toxins to cause lung cancer and many other diseases such as heart disease. **Native speaker**

7.3.2.2.4 Other

There are cases where there are not enough examples for me to confidently identify specific patterns that can be said to occur with prospecting CSNs in the two corpora. The following is an illustration of such cases.

(55): I think smoking should be completely banned in all restaurants in Japan. The **reasons** are quite simple. First and most important smoking is dangerous to health, secondly, second had smoke is unfair to non smokers and finally, smoking, particularly in restaurants is bad for the taste of food. **Native speaker**

Example (55) presents an instance where it is difficult to identify the pattern as a potential pattern associated with the CSN *reason*. The reason is that it occurs once in the two corpora. However, it can be said that the whole sentence where the CSN *reason* is located serves to prospect the causal meanings. In the first example, "*The reasons are quite simple*" can be said to prospect that causal meanings (reasons) are to be supplied in the ensuing text.

Having completed all the frequent patterns occurring with CSNs both within and across clause boundaries, Table 7-2 lists all these patterns along with their frequencies in the learner and native speaker corpora.

Table 7-2 *List of all the patterns associated with CSNs along with their frequencies in the two corpora*

Pattern	Raw frequency in the learner corpus	Raw frequency in the native speaker corpus	Total frequency in the two corpora
Within clause boundary patterns			
CSN + v-link + specifics: that-appositive clause	10	7	17
CSN + v-link + specifics: other appositive clause	16	13	29
CSN + v-link + specifics: noun phrase	1	3	4
Specifics + v-link + CSN	2	9	11
Patterns of CSNs occurring in adjunct phrases:			
- Specifics: appositive clause + CSN in adjunct phrases	0	13	26
- CSN in adjunct phrases + specifics: appositive clause	0	13	
CSN + specifics: comparative	1	2	3
Across clause boundary patterns			
(For) these + CSN	3	21	24

***Table 7-2** (continued)*

(For) this/ that + CSN	5	4	8
It/ this/ that + v-link + CSN	3	8	11
Patterns including the word *above*:			
- *For/based on* + *the* + *above* + CSN	0	5	8
- *For* + *the* + CSN + v-ed + *above*	0	3	
Patterns including quantifiers:			
- *There* + v-link + Quantifier + CSN	6	12	
- *For* + Quantifier + CSN	20	13	59
- *Because of* + Quantifier + CSN	4	0	
- *I* + *have* + Quantifier + CSN	0	5	
For + *the* + *following* + CSN	0	3	3
Prospecting verb + CSN	1	4	5
Others (patterns occurring less than three times)	3	14	17

Having explored how CSNs are constructed and having identified the patterns associated with the construction of causality, I will now move on to the second section of the second part of this chapter. The aim of this section is to compare the learner corpus and the native speaker corpus in terms of the use of CSNs.

7.4. Analysis and findings of CSNs in the two corpora

Table 7-3 demonstrates the frequency with which SNs with inherent causal function are used in the two corpora. The table lists the CSNs that are found to occur among the list of the SNs in the two corpora of the study:

Table 7-3 *Causal signalling nouns in the two corpora along with their raw and normalised frequencies*

Rank	Lemmatised CSNs in the learner corpus	Raw frequency	Normalised frequency per 100,000 words	Rank	Lemmatised CSNs in the native speaker corpus	Raw frequency	Normalised frequency per 100,000 words
1	Reason	75	35	1	Reason	106	117
2	-	-	-	2	Result	27	29
3	-	-	-	3	Effect	10	11
4	-	-	-	4	Cause	9	10
		Total=75	Total= 35			Total=152	Total= 167

Table 7-3 shows that there are strikingly discrepancies between the two corpora in terms of the normalised frequency of the CSNs identified and the choices of individual CSNs across the two corpora. Of all the SNs that occur in the learner corpus, only one SN is identified to be an inherent CSN. This noun, as displayed in Table 7-3, is *reason*. Unlike the learner corpus, there are four inherent CSNs occurring among the list of SNs in the native speaker corpus: *reason, result, cause* and *effect*. Only the CSN *reason* is used in both corpora. Thus, this noun will be closely looked at in the following section.

7.4.1. *Reason(s)*

Reason is the most frequently used CSN in the two corpora; it ranks third in the list of SNs in the native speaker corpus, and eighth in the list of SNs in the learner corpus. However, the normalised frequency of the noun per 100,000 words, as shown in Table 7-3 presented earlier, indicates a striking difference between the two corpora—there are 106 occurrences (117 per 100,000 words) in the native speaker corpus whereas there are 75 occurrences (35 per 100,000 words) in the learner corpus. The difference in the use of *reason* in the two

corpora is significant[17] with the probability value of $p<0.0001$. One plausible reason why *reason* is under-represented in the learner corpus is that the learners are less familiar with nouns as a potential linguistic resource by which causal meanings are signalled. Thus, the learners rarely use nouns at all to signal causality. The only noun they use with this function is *reason*. The question then arises as to whether this is because they do not talk about causality very much or because they find other ways to express it. In order to investigate this, I have chosen to investigate the causal conjunct *because* whose function is very close to that of the CSN *reason*. To test the hypothesis that the learners tend to use *because* instead of *reason* I have counted the number of times that it is used in the learner and native speaker corpora.

Table 7-4 provides both the raw and normalised frequency of both causal-establishing elements in the learner and native speaker corpora.

Table 7-4 *Raw and normalised frequencies of reason(s) and because across the two corpora*

Source Corpus	**Raw and (normalised frequency per 100,000 words) of the CSN *reason***	**Raw and (normalised frequency per 100,000 words) of the causal conjunct *because***
Student corpus (CEPA)	75 (35)	1309 (620)
Native speaker corpus	106 (117)	315 (347)

[17] I have used the log likelihood calculator created by Rayson (2017). The likelihood result obtained is -63.04. For more clarification how Rayson's calculator works and how to interpret obtained results, refer to Section 4.5 in Chapter 4.

As shown in Table 7-4, there is a striking difference between the two causal resources in terms of their frequency in the learner corpus. Unlike English native speakers, the learners tend to rely heavily on the causal conjunct *because* when signalling causal meanings between clauses. Thus, this might partly explain the lower frequency of the CSN *reason* in the students' writing. The explanation for the learners' extensive use of the conjunct *because* in lieu of the CSN *reason* could be related to the degree of abstraction involved in the construction of casual relations between propositions. To explain what I mean by abstraction here, Halliday's (1985, 1994) notions of 'congruency' and 'metaphoricity' in the representation of meanings in text are highly relevant. According to Halliday (1985, 1994) and Martin (1993), there is a relationship between grammar and semantics in the representation of meanings in discourse. The authors point out that this relationship can be either congruent or incongruent (metaphorical). When things are represented in language by nouns; actions are represented in verbs and logical relations between ideas are represented in conjunctions, information is communicated congruently (Halliday, 1985; 1994; Martin, 1993). When causality is expressed through CSNs, causal meanings are abstracted in the sense that they are not often linked in a congruent way. The corollary is also true; causality expressed by the causal conjunction *because* directly and invariably links effect and cause. Thus, it can be concluded that the abstract nature of CSNs can be held responsible for the underuse of causality expressed through CSNs in the learner corpus compared to that of native speakers'.

The grammatical complexity arising from the use of CSNs in expressing causal meanings can be further explained by the following two examples; the first is an illustrative example of causality signalled by the use of *because*; the subsequent examples represent my personal attempts to rephrase the first example by using the CSN *reason* instead of *because*.

(56): I liked this movie **because** it was funny, interesting and different. **CEPA 6**

(57): I liked this movie for the **reason** that it was funny, interesting and different. **Within clause Paraphrase 1**

(58): The **reason** why I liked this movie was that it was funny, interesting and different. **Within clause Paraphrase 2**

In example (56), the two elements of causality: the 'cause' and the 'effect' are linked congruently by the use of the conjunct *because.* Halliday and Matthesion (1999, p.245) point out that logical relations are "construed congruently by the grammar as a clause nexus joined by a conjunction" either paratactically or hypotactically. What we have in example (56) are two clauses joined hypotactacally, representing what Halliday and Matthesion (1999, p.245) refer to as a clause nexus. These two clauses are linked congruently by the conjunct *because*. However, if writers were to use the CSN *reason* in lieu of the conjunct *because*, inevitable grammatical abstraction is to take place. Examples (57) and (58) are illustrative of this grammatical complexity. In these examples, a number of grammatical changes have taken place. The causal meanings are no longer linked congruently by conjunctions, but are linked incongruently by another grammatical class: nouns. Another change is that the clauses used to construe the causal meanings are rankshifted (see Halliday, 1994, p.242, for more detail about rankshifting). In example (57), we have one clause that is rankshifted into a noun phrase: *for the reason* [[*that it was funny, interesting and different*]]. In example (58), the situation is even more complicated. Here, we have two clauses that are rankshifted: The **reason** [[why I liked this movie]] and [[that it was funny, interesting and different]]. The first of these is rankshifted into a noun phrase while the latter has been turned into a complement clause forming an element in another

clause. This shows that choosing ‘reason’ rather than ‘because’ entails a more complex grammar in the sentence, particularly in the noun phrase.

In addition, the use of CSNs rather than the use of the conjunct *because* to signal causality enables writers to add more meanings to the causal meaning(s) being signalled. The use of CSN *reason* is a case in point. Hoey (1993) points out that the grammar of the CSN *reason* when occurring in subject position allows writers “to comment on the importance or relevance of the reason to be offered” (p.75) and “to count one’s reasons and thereby alert the reader to either the presence or the absence of others” (p.76). A closer look at the patterns identified with the CSN *reason* in the two corpora of the current study lends support to Hoey’s observation (see, for instance, the examples provided in Section 7.3.1.2). However, the two functions pointed out by Hoey (1993) are also found when *reason* occurs as complement rather than subject, as shown in the following examples taken from the two corpora used in the present study:

(59): Down here in the south, I think that a lot more people smoke than in other areas of the country, and I also think this is a primary **reason** for the port image of southern health. **Native speaker**

(60): There are mnay beautiful places in my country, but my best one is my school. I really like it for many **reasons**. **CEPA5**

What the two examples illustrate is that the grammar of the CSN *reason* even when occurring in the complement position allows writers to evaluate the offered reason as in example (59) and to enumerate the reasons to be offered as in example (60). As noted above, The CSN *reason* can be preceded by adjectives whose function is to evaluate the signalled reason(s) or

can also be preceded by quantifiers whose role is to enumerate the signalled reason(s). However, when the conjunct *because* is used to signal causality, there is no opportunity for enumerating or evaluating the offered *reason(s)*. To illustrate this, examples (61) and (62) are presented; the first is an illustrative example of causality signalled by the use of the CSN *reason* preceded by an enumerative quantifier and by an evaluating adjective (I have added this adjective to the native speaker example for illustrative purposes); the subsequent example represents an attempted paraphrase using the conjunct *because*.

(61): College students should not work while they are in school for several (main) **reasons**. First, it is difficult to focus, second, it takes time away from studies, and third, sometimes a part time job can lead to other paths which are distracting to the path of a degree which was original goal. **Native speaker**

(62): College students should not work while they are in school. This is **because** it is difficult to focus, and **because** it takes time away from studies, and **because** sometimes a part time job can lead to other paths which are distracting to the path of a degree which was original goal. **Paraphrased version**

Examples (61) and (62) illustrate the fact that to paraphrase a CSN preceded by a quantifier or an adjective using the conjunct *because* leads to a cumbersome sentence. Therefore, it could be argued here that the use of the causal conjunct *because* in lieu of other possible CSNs may have made the learners' writing fall short in terms of the addition of more meanings to the causal meaning being signalled (e.g. organisation and evaluation). This

argument will be further discussed in Chapter 8 under the heading of pedagogical implications. Next, the adjectival pre-modifiers that are found to occur with the CSN *reason* in the two corpora will be addressed in Section 7.4.1.1.

7.4.1.1. Adjectival pre-modification of the CSN *reason(s)*

CSNs, as any other nouns, can be pre-modified with an adjectival phrase. The adjectival phrase may consist of one or more adjectives. Examples of adjectives phrases occurring with the CSN *reason* in the corpora of the present study include *main reason*, *two main reasons* and *the following three reasons*, where the last example also consists of a determiner (*the*) and is only found in the native speaker corpus. These examples illustrate the two main categories of adjectival modifiers found to occur in the two corpora, as noted above; (1): adjectives such as *main* and *following* and (2) quantifiers such as *two* and *three*. Quantifiers are generally accepted as adjectival modifiers of nouns (e.g. Tåqvist, 2016). Table 7-5 shows the total number of adjectival pre-modifiers that occur with CSN *reason* in the two corpora along with number of the different types and the Type-Token Ratio (TTR)[18].

Table 7-5 *Total raw frequency of adjectival modifiers along with the number of types*

	The learner corpus	**The native speaker corpus**
Raw Frequency (tokens) of adjectival modifiers	54	78
Number of types	14	33
TTR	0.26	0.42

[18] Type-Token Ration (TTR) = (number of types/number of tokens) * 100

As Table 7-5 indicates, the difference between the learner and native speaker corpora in terms of both the total number of instances of adjectival pre-modification and the number of adjectival types is substantial. The considerable difference in the frequency of the adjectival pre-modifiers is not surprising, given the significant difference between the instances of the CSN: *reason* in the two corpora overall reported earlier in the beginning of Section 7.4.1. However, the difference in the total instances of adjectival types in relation to the overall frequency of adjectival modification instances, as displayed by the TTR figures in Table 7-5, is striking. The learner corpus is characterised by a TTR that is remarkably lower than that of native speakers'. That is, the Arabic speakers who are learning English use substantially fewer adjective types in relation to the overall frequency of the adjectives that they use to pre-modify the CSN *reason*. The native speakers, on the other hand, make use of a wide variety of adjectival modification. The most likely explanation for the substantial variation between the two corpora with regard to the number of the adjective types identified is that the learners have a very restricted repertoire of adjectives that can be used to precede the CSN *reason.*

As mentioned previously, the adjectival pre-modification of the CSN *reason* in the learner and native speaker corpora falls into two main categories: adjectives and quantifiers. A closer analysis of these types of pre-modification has revealed interesting findings. These findings will be presented below in Sections 7.4.1.1.1 and 7.4.1.1.2.

7.4.1.1.1 Adjective modifiers

Table 7-6 provides the frequency of the adjectives that are found to precede the CSN *reason* in the two corpora along with Type-Token Ratio (TTR):

Table 7-6 *Frequency of the adjectives preceding the CSN reason along with the TTR*

	Learner corpus	**Native speaker corpus**
Total frequency (tokens)	21	39
Number of ADJ types	8	24
TTR	0.38	0.62

As Table 7-6 demonstrates, there is no substantial difference between the two groups in the total frequency of the adjectives that are found to pre-modify the CSN *reason*. However, the Type-Token Ratio (TTR) figures, as illustrated in Table 7-6, shows a remarkable difference between the learner and native speaker corpora in terms of the number of the adjective types identified in relation to the overall frequency of the adjectives used. This finding is similar to that noted above when all adjectival modifiers occurring with the noun are considered. The Arabic speakers who are learning English are less varied in the adjectives they use to pre-modify the CSN *reason*. Compared with the native speaker writers, the learners use fewer adjective types. The learners are found to only use the following adjectives to pre-modify *reason*: *main, important, another, first, second, other, only* and *personal*. The native speakers, on the other hand, make use of a wide variety of adjectival modification. Examples of adjectives that are only found to occur in the native speaker corpus are *primary, main, good, biggest, obvious, excellent, real, significant, important, enough, undisputable, stupid, largest, added, sole, only, above, following, first, second, another, other, final* and *last*. The frequency of the adjective types in the learner and native speaker corpora can be found in Appendix J. What the aforementioned adjective types used by the native speakers of English suggest is that these writers, unlike the learners, use adjective pre-modifiers to communicate a wider range of meanings. The following examples, taken from the native speaker corpus,

illustrate this finding; the examples are presented first and comments on the examples are addressed afterwards.

(63): Of course, there are the arguments of health and second hand smoke and more, but the real **reason** I think that smoking should be banned in restaurants is because it negatively impacts the experience.

(64): Many people might think these are some stupid **reasons** to have a part time job, but to me they're perfectly good,

(65): The largest and most undisputable **reason** is that smoking shortens people's lives.

(66): I have most certainly heard of many deaths and illnesses related to smoking. In itself, that should be enough **reason** for Japan to ban smoking in restaurants and all public places without argument or debate.

(67): Based on the above **reasons**, I am against the opinion.

(68): I do support the idea of college students working part-time, or even full-time for that matter for the following **reasons**. To begin, I believe by working it teaches students many valuable lessons about life that will also be a part of his/her educational process.

In the above examples, the native speaker writers make use of different adjectives to pre-modify the CSN *reason*. By using these adjectives, they are able to express a wide range of meanings. For instance, examples (63) and (64) both enact a dialogue between the writer and a presumed reader. In (63), for example, the writer assumes that the reader would hold a view about previously given reasons and then the writers provides a new reason that the writer thinks is the 'real' reason for the issue being discussed. In (65), the writer provides their stance strongly. The reason offered in this example is 'undisputable'. By using this adjective, the reason is presented as something not to be argued about. Similarly, the adjective used in (66) is used to help the writer cast their stance strongly. By using the adjective 'enough', the writer strengthens the force of their argument by emphasising the causal meaning being signalled. Finally, in (67) the writer uses the adjective 'above' with the CSN *reason* as an organising adjective to help encapsulate previously mentioned causal meanings.

7.4.1.1.2 Enumerative modifiers

The difference between the learner corpus and the native speaker in terms of the enumerative modification of the CSN *reason* is noteworthy. Table 7-7 below shows the total tokens of enumerative items that are found to precede the CSN *reason* in the two corpora and the number of the different types of these items along with the Type-Token Ratio (TTR):

Table 7-7 *Frequency of enumerative tokens preceding the CSN reason along with the TTR*

	Learner corpus	**Native speaker corpus**
Absolute frequency (tokens)	33	39
Number of different modifiers (types)	7	9
TTR	0.21	0.23

Unlike the adjectival modification where it was noted that the learners use strikingly fewer adjective types to modify the CSN *reason*, the TTR of enumerative items that are found to modify the noun, as shown in Table 7-7, is a less distinguishing factor between the two corpora. The learners use only slightly fewer enumeration types preceding the CSN *reason* compared to that of native speakers'. The frequency of all the enumerative types that are found to occur with the CSN *reason* in the two corpora is provided in Appendix K.

A closer look at the enumerating items used by the two group writers has revealed that these items can be classified into two categories: **vague** enumerative modifiers and **precise** enumerative modifiers. The vague enumerative modifiers preceding the CSN *reason* are defined as those modifiers that fail to indicate the exact number of the offered or to be offered reason(s). Precise enumerative modifiers, as the name suggests, are defined as those modifiers that precisely indicate the number of the reason(s) either previously mentioned in text or to be present in the ensuing text.

The learners and the native speakers strikingly differ in relation to the use of these two categories of modifiers, as displayed in Table 7-8, below.

Table 7-8 *Total number of vague and precise modifiers preceding the CSN reason along with percentages*

	Learner corpus	**Native speaker corpus**
Absolute frequency (tokens) of vague modifiers (%)	27 (82%)	16 (41%)
Absolute frequency (tokens) of precise modifiers (%)	6 (18%)	23 (59%)
Total	33 (100%)	39 (100%)

As the figures in Table 7-8 demonstrate, the learners differ considerably from the native speakers in the use of the two categories of enumerative modifiers. Out of the 33 occurrences of the enumerative modifiers in the learner corpus, 27 of these modifiers, which represent 82% of all enumerative modifiers, are found to be vaguely modifying the CSN *reason*. The native speakers, on the other hand, use less vague enumerative modifiers; their use of vague modifiers accounts for 16 instances comprising 41% of the total occurrences of enumerative modifiers found in the corpus. Unlike the learners, the native speakers predominantly use precise enumerative modifiers to count the offered or to be offered reason(s); out of 39 instances of the CSN modified by enumeration, 23 (59%) of these instances are classified as precise enumerative modifiers. Thus, it can be concluded that the learners are vaguer than the native speakers when signalling the number of causal meanings (reasons) in their writing.

Examples of vague enumerative modifiers of the CSN *reason* that are found in the two corpora are *many* and *several*. The following are illustrative examples of these vague modifiers:

(69): There are many **reasons** to judge this as worst. First, my lovely friend Aisha who always goes out with me was sicked. **CEPA 5**

(70): Despite this, there are many **reasons** why discouraging smoking especially in restaurants makes sense given the patronage of other clients and especially small children. It is well known that secondhand smoke is much more dangerous than first hand smoke. **Native speaker**

(71): I enjoyed the movie a lot for several **reasons**, one because it was based upon a true story, a story ...**CEPA 6**

(72): I agree with the statement above for several **reasons**. First, secondhand smoke harms non-smokers. **Native speaker**

As displayed in the above examples, the vague enumerative modifiers: *many* and *several* are found to occur in the two corpora. However, there are other vague modifiers that are found to uniquely occur either the learners' corpus or the native speakers' corpus. The vague modifiers that are only used by the native speakers are *a number of*, *various* and *countless* whereas the vague modifiers that are only used by the learners are *some* and *a lot of/lots of.*

By modifying the CSN *reason* with vague enumerative modifiers, writers avoid committing themselves to a particular number of reasons that are needed to lexicalise the CSN *reason.* As mentioned previously, the learners stand out in terms of the use of vague modifiers that are found to precede the CSN *reason* and thus it could be said that the learners show strong preference towards not committing themselves to signalling the exact number of causal meanings (reasons) that are to be found in text.

By preceding the CSN *reason* with precise enumerative modifiers, writers are committed to a given number of reasons that are needed to supply the specification of the CSN *reason.* As mentioned before, the native speakers stand out in terms of the use of precise enumerative modifiers that are found to precede the CSN *reason.* Examples of precise enumerative modifiers that are found to occur in the two corpora are *one* and *two*. The native speakers are found to make use of other modifiers, which are non-existent in the learners' corpus. Examples of these modifiers are *three* and *four*. The following are illustrative of cases when

precise enumerative modifiers are found to precede the CSN *reason* as taken from the native speaker corpus and the learner corpus:

(70): I agree with this statement for two **reasons**. Firstly there is the issue of finance. **Native speaker**

(71): It was the worst holiday I have ever had for two main **reasons**. First, we had a mechenical problem in the car.

CEPA 5

(72): One **reason** I think this is because I have been working at a part time job for most of my college career. **Native speaker**

(73): Entertainment is one of the **reasons** I use my computer. **CEPA5**

(74): I have three **reasons**. First, all non smokers are paying customers and have the right to enjoy their meal in a smoke free environment. **Native speaker**

7.5. Summary of the findings and concluding remarks

This chapter has focused on the use of CSNs in the learner and native speaker corpora. It has enumerated the most frequent within and across clause patterns associated with this type of nouns in the two corpora. The chapter has also compared the two corpora with regard to the use of CSNs. In comparison with the native speaker writers, the use of CSNs is one of the greatest areas of disparity. It has been found that the learners use CSNs significantly less in terms of tokens as well as types. It has been argued that the lack of CSNs does not occur

because the learners do not talk about causality but rather that they use other resources to signall it. The stylistic consequences of that have been discussed. It can be concluded that CSNs are one of the most important points for language development for learners. So, this is something of pedagogic importance (see Section 8.3.2 in Chapter 8).

CHAPTER 8: DISCUSSION AND CONCLUDING REMARKS

8.1. Overview

This final chapter of the present thesis is organised into four sections. The aim of the first section (8.2) is to summarise and discuss the findings of the thesis. The second section (8.3) explores the pedagogical implications of the results. An assessment of the methodology used in the present thesis is addressed in Section 8.4. The final section of this chapter (8.5) presents my concluding remarks/thoughts and possible directions for further research on the phenomenon of SNs.

8.2. Summary and discussion of findings

The overall aims of the present thesis were to investigate how SNs are used in texts written by Arabic speakers who are learning English and to find out in what ways the uses of SNs by those learners either resemble or differ from those found to occur in texts produced by native speakers of English. Another aim was to examine how the resources of stance and causality through SNs were constructed. Four main questions were addressed. Each of these is presented in the following sections. The aim is to summarise and discuss the findings related to each question.

8.2.1. What is the frequency of SNs in the learner and native speaker corpora, and how the two corpora compare?

This question is concerned with the overall frequency of SNs across the learner and native speaker corpora. The question was addressed in Chapter 4. The analysis of SNs in terms of frequency in the two corpora yielded a number of interesting findings. These findings showed that there are similarities and differences between the two corpora, with the latter being more prominent. The main point of similarity is addressed first below.

In both corpora, the choice of SNs is often sensitive to the topic being addressed. Although this finding is generally in line with previous literature, which indicates that there is a link between genres and disciplines and the choice of certain SNs (see, Flowerdew and Forest, 2015), the frequency analysis of SNs in the two corpora showed that it was possible to link the choice and the frequency of particular SNs to the specific topic being addressed.

The frequency analysis in terms of the tokens and types of SNs in the learner and the native speaker corpora demonstrated that there were remarkable differences. The learner corpus contained fewer SN tokens and types, with the first in comparison to the native speaker corpus was found to be statistically significant (see Table 4-8 and its commentary). This finding is interesting in view of previous research in the existing literature, particularly Flowerdew (2010) and Tåqvist (2016). The finding corroborates Flowerdew's study (2010), which found that Chinese speakers who were learning English used fewer SN tokens and types. However, the finding that English L2 learners use fewer SN tokens is contrary to Tåqvist's study on *discourse organising nouns (DONs)* (2016). In her study, she found that Swedish speakers who were learning English used more DONs tokens compared to English native speakers writers. The possible reason for the discrepancy between the present study's finding with regard to SN tokens and that of Tåqvist (2016) is that in Tåqvist's study the learners being examined were more advanced and were writing in a more academic context. All previous research has shown that SNs are particularly prevalent in academic discourse (e.g. Flowerdew, 2003; Flowerdew, 2006; Flowerdew and Forest, 2015).

The quantitative analysis with regard to which SNs are used in the learner and native speaker corpora uncovered clear differences. The learners were found to use SNs with borderline status (such as *holiday(s)*, *trip(s)* and *vacation(s)*). This was not the case with the native

speaker writers where all the instances of SNs used were clearly identified as abstract nouns. The frequency regarding borderline SNs in the learner corpus is interesting for the following reason. Although my inclusion of these nouns has been justified (see Section 3.4), the high proportion of borderline SNs does suggest that if we were to remove these borderline cases we would, in fact, have far less use of SNs. This in turn would make the contrast with Tåqvist (2016) more striking, as the nouns investigated in her study were standard abstract nouns.

8.2.2. What is the frequency of SNs in the learner corpus across the proficiency levels of the learners?

This question was addressed and answered in Chapter 4. The frequency analysis of SNs across the four proficiency levels (CEPA 3, CEPA 4, CEPA 5, and CEPA 6) of the Arabic speakers who were learning English supported the hypothesis raised in Chapter 4. The hypothesis was that the learners' use of SNs would increase as they become more competent in English. So, as expected, the number of SN tokens as well as SN types used by the Arabic speakers who were learning English was found to increase as the proficiency level of these learners increased (see Table 4-4 and Figure 4-2). This finding in terms of SN tokens is in line with previous research which found that there was a correlation between graded written texts by Chinese speakers who were learning English and the frequency of SNs; texts of lower grades showed a lower frequency of SNs (Flowerdew, 2006). The finding obtained in the present study with regard to SN types used across the four proficiency levels adds a new, though unsurprising, insight to the existing literature; it is that not only SN tokens are used more as learners get better but also learners use more SN types as they improve. What this finding suggests is that the control over SNs is a developmental resource.

SNs are part of lexical cohesion as Flowerdew (2006) points out. Therefore, improving learners' control over SNs would improve their use of lexical cohesion. This in turn would enhance the cohesiveness of their writing.

8.2.3. How are the three most frequent and shared SNs in the learner and native speaker corpora used in relation to the reference categories these nouns may fall into, the grammatical forms of specification and what meanings of these nouns can be identified? How do the two corpora compare with each other with these regards?

To answer this question, I focused on three most frequent and shared SNs in the learner and native speaker corpora. These SNs were *time(s)*, *thing(s)* and *way(s)*. The detailed examination of each of these SNs was addressed in Chapter 5. The summary and discussion of the findings of this chapter are presented below under the following three headings: reference category, grammatical forms of SN specification, meaning in context and the SN *thing(s)* and the notion of pattern.

8.2.3.1. Reference category

The data analysis of the reference categories (within clause specification (WCS) and across clause specification (ACS)) of the three SNs in the learner and native speaker corpora showed more similarities than differences. The instances of within clause specification were by far the most frequent reference categories in the two corpora. This finding is interesting in view of the relevant literature in two respects. First, the literature on cohesion is always restricted to reference relations that exist between segments of text beyond clause and sentence boundaries (e.g., Halliday and Hasan, 1976, p.9; Francis, 1986, p.5; Martin, 1992). The examination of the three nouns in question in Chapter 5 in terms of across and within clause reference relations showed that within clause reference is the most frequent reference category. This finding is also echoed with other SNs investigated in the present study such as *opinion(s)*

and *fact(s)*, as found in Chapter 6. Thus, it can be said that the present study in terms of reference has demonstrated that it is also a value to consider cohesive links within clause and across clause boundaries as they are frequent in discourse.

Second, in view of the literature on the phenomenon of SNs where within clause reference is included the finding of the present study is interesting. It is not corroborated in Flowerdew and Forest's study (2015, p.185) which found that SNs were used more frequently in the across clause reference pattern. However, it is in line with Tåqvist's study (2016).

A closer look at the sub-categories of the WCS (cataphoric (CAWCS) and anaphoric (ANWCS)) and ACS (cataphoric (CAACS) and anaphoric (ANACS)) revealed a key finding. It was that the SNs examined in Chapter 5 showed strong preference towards certain reference sub-categories. For example, two of the three SNs (e.g. *time(s)* and *thing(s)*) when used within clause boundaries they were predominantly found to occur in the CAWCS reference pattern across the two corpora. It was also found that the SN *time(s)* when used across clause boundaries in the two corpora it never occurred in the CAACS pattern. In other words, the SN *time(s)* was always used to refer to a previously stated stretch of discourse. Viewing these findings in light of the literature dealt with the phenomenon of SNs (e.g. Flowerdew and Forest (2015) and Tåqvist (2016)), it can be said the present study has contributed to the knowledge of how SNs are used in terms of reference patterns. These studies are able to tell us about the overall frequency of reference patterns associated with SNs found to occur in these studies but it is not clear whether such findings are representative of all nouns. This raises the possibility that the very general SNs (e.g. *way(s), time(s)* and *thing(s)*) may behave in ways that other SNs (e.g. *reason(s), problem(s)* and *issue(s)*) do not. Therefore, the very general SNs may need to be investigated in ways that other SNs may not. I will return to this point later when I offer suggestion for further research in Section 8.5.

8.2.3.2. Grammatical forms of SN specification

The investigation of the three SNs gave us insight into how these nouns were specified in discourse. The specification of the SN *time(s)* was found to occur in three grammatical forms: clauses, prepositional phrases and noun phrases. The other two nouns (e.g. *thing(s)* and *way(s)*) were specified as clauses, noun phrases and nominalisation. These findings are interesting in view of the literature on SNs. As noted in Section 2.5 in Chapter 2, there is no total agreement on what constitutes the specification of SNs. Some are strict and only include clauses (e.g. Francis, 1994; Schmid (2000)) while others include nominalisation instances along with clauses (Flowerdew, 2003; Flowerdew and Forest, 2015). Therefore, the present study broadens the picture by including phrases and makes a contribution to the knowledge in that it shows how SNs are used in discourse.

8.2.3.3. Meaning in context

The close scrutiny of the meanings in context associated with *time(s)*, *thing(s)* and *way(s)* revealed that these nouns are used with a wide range of meanings. In the learner and native speaker corpora, three meanings were identified to occur with each SN. The learners and native speakers showed a large degree of similarity in their uses of these meanings. For example, the meanings identified with the SN *time(s)* were “time orientation”, “recounting experiences” and “evaluation”. Both groups of writers used these meanings with almost the same frequency (see Table 5-3 in Chapter 5). *Time orientation* and *recounting experience* were the two most frequent uses of the SN *time(s)*, followed by the meaning of evaluation which represented the least uses of the noun in discourse. In terms of the SN *way(s)*, the three meanings identified for the noun were “method”, “manner” and “possibility”; the first two of these meanings represented the most frequent uses of the SN whereas the possibility meaning constituted the least frequent uses of the noun (see Table 5-10 in Chapter 5). As far as the third SN *thing(s)* is concerned, the noun was found to be the most semantically neutral SN

and thus meanings associated with the noun were found to be difficult to categorise. It was for this reason that the uses of the noun were investigated in terms of pattern instead of meaning. The findings are presented and discussed in the next section (Section 8.2.3.4).

8.2.3.4. The SN *thing(s)* and the notion of pattern

The data analysis of the SN *thing(s)* produced a wealth of insights into how the noun was used in the two corpora. The noun was found to be used significantly more in the learner corpora. Perhaps, the most interesting findings were that there was a wide range of patterns of the SN *thing(s)* and that the notion of pattern that the noun occurred in was important in determining whether the use of the noun could be regarded as redundant or as natural (see Chapter 5). These two key findings are not without their pedagogical value and thus are addressed again in the pedagogical implication section, below (see Section 8.3.1).

8.2.4. What are the patterns that can be associated with SNs signalling causality and those signalling stance? How comparable are the two corpora in their uses of SNs to signal those two discourse functions?

This question addressed two important roles of the phenomenon of SNs: stance and causality. Findings related to the stance function of SNs were addressed in Chapter 6 and those related to the causality role of SNs were provided in Chapter 7. The findings are summarised and discussed in the following two sections.

8.2.4.1. SNs and stance

The examination of the inherent stance signalling nouns (SSNs) that occurred in the learner and native speaker corpora revealed a number of useful information with regard to how stance was constructed in discourse in general and how SSNs were used by the two groups of writers. The central findings were that (1) stance through SNs were constructed in two ways:

either within or across clause boundaries, that (2) there were a wide range of patterns for each way of construction, and that (3) there were remarkable discrepancies between the learners and native speakers in their uses of SSNs.

The first two findings (1) and (2) demonstrated the advantage of looking at SSNs from a discourse perspective as a way to understand how the resource of stance could be constructed through SNs. In view of the relevant literature, the findings with regards to stance construction can be said to have broadened the picture. As noted in Chapter 6, when stance was looked at from the perspective of nouns the literature tended to restrict the investigation to very few patterns (e.g. 'noun + appositive-that clause' pattern as in Charles (2007), Hunston (2011) and Jiang and Hyland (2015) and 'this + noun' as in Charles (2003)). The present study was able to identify a wide range of patterns that occurred with the SSNs examined either when stance through these nouns were constructed within or beyond clause boundaries. Thus, the study has made an important contribution to the field.

As noted above, the third key finding was that learners as compared with native speakers were strikingly different in terms of the uses of SSNs. There were two main areas of difference. The first difference was that the learners used SSNs significantly less. Although a possible explanation for this marked difference was given in Section 6.4, this discrepancy raises the question as to whether learners may have signalled stance in other ways, or as to whether they had not expressed stance in general as much as native speakers. Inevitably, it may be difficult to answer this question since in any corpus-based study like the present one we can only search for words and not concepts. Searching for words signalling stance (SSNs), the present study suggested that stance through nouns are significantly less used. However, stance can be signalled by other words such as adjectives, adverbs, verbs and so

on. It was beyond the focus of the present study to look at these resources and thus this may be another area where stance in learner language can be investigated.

The second difference was revealed by the detailed examination of the SSNs: *opinion(s)* and *fact(s)*. The learners used these nouns to construct stance within-clause boundaries whereas the native speakers used the nouns to construct stance within and across clause boundaries. The learners also tended to use these SSNs in their fixed patterns (*in my opinion* and *in fact*). The prevalence of these two fixed patterns in the learners' writing is interesting for two reasons. First, it showed the learners' tendency to cling to what they felt safe and familiar with. This interpretation is congruent with Hasselgren's study (1994) on 'lexical teddy bears', that is words and phrases that learners feel familiar with either because they resemble L1 words and phrases or because they are learnt early. Thus, the learners were found to be restricted in their uses of these nouns. Compared to the learners, the native speaker writers were found to use 'in my opinion' and 'in fact' less often. Second, it showed the importance of including SNs in fixed patterns in order to yield a more comprehensive picture of the phenomenon of SNs in learner language.

Flowerdew and Forest (2015, p.59) exclude SNs occurring in prefabs (their terminology for fixed phrases) for two reasons. One is that they would vastly inflate the number of SNs. The more important reason is that they regard prefabs as close to the grammatical end of the grammaticalisation cline. I have taken the opposite view for two reasons. One is the practical reason that if I exclude these very frequent items, in some cases there is very little left for me to investigate. The more important reason is that learners may well learn these prefabs at the grammatical end of the continuum and extend that knowledge towards the lexical end of the cline.

Interestingly, although the SSN *idea(s)* was found to occur significantly less in the learner corpus the data analysis of this noun showed that the learners used a wide range of patterns and used this noun to construct stance within and across clause boundaries. In fact, they used more patterns than those used by the native speakers. This finding is not line with the findings about the SSN *opinion(s)* and *fact(s)* where the learners showed less flexibility in using these two nouns in a variety of contexts. The possible explanation for this different finding in relation to the SSN *idea(s)* is that there is no 'lexical teddy bear' for the SSN *idea(s)* that is the same as the fixed phrases: *in my opinion* and *in fact*.

8.2.4.2. SNs and causality

The analysis of inherent causal signalling nouns (CSNs) used in the learner and native speaker corpora provided many interesting insights into how causality was signalled in discourse. One of these insights was that causality through SNs was constructed within and across clause boundaries and that there were many patterns associated with these two ways of construction. The second main finding was that the resource of CSNs was used significantly less by the learners compared with the native speaker writers. In fact, the only CSN used by the learners was *reason(s)*, whereas the native speakers used four CSNs. This finding raised a question as to whether the learners did not talk about causality as much as native speakers, or whether the learners used other ways to express the concept. The present study found that the learners used the conjunct *because* strikingly more than the native speakers (see Table 7-4 in Chapter 7). This suggests that the learners signalled causality in their texts, but through a resource other than a noun. The conclusion as to why the learners extensively used the conjunct *because* in lieu of the CSN *reason(s)* to signal causality was that the use of the CSN *reason(s)*, like other CSNs, would involve a degree of grammatical abstraction (see Chapter 7). The final key finding was related to the adjectival modification of the CSN *reason(s)*. The

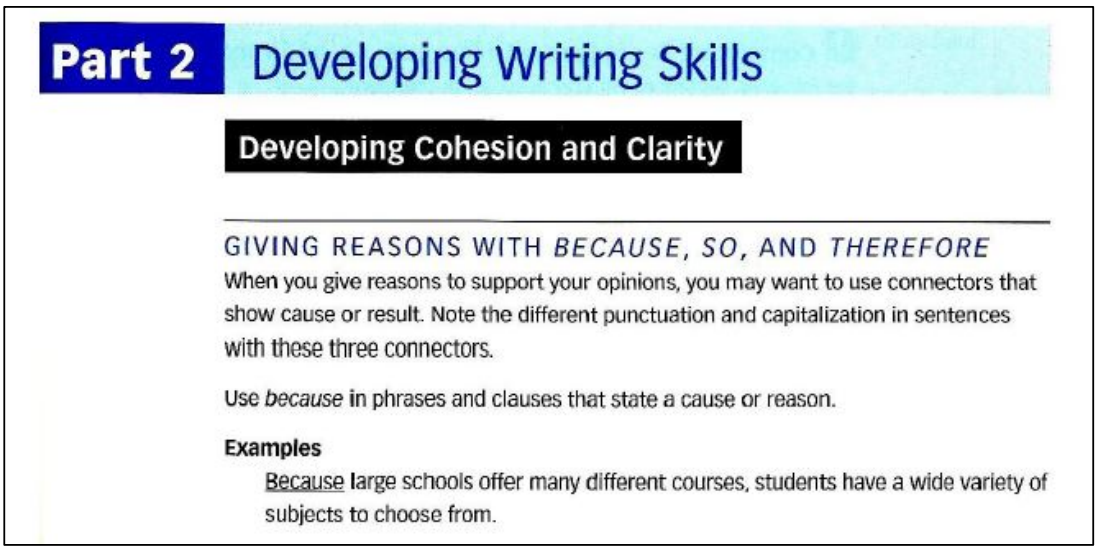
learners' use of adjective pre-modification in terms of token and types was significantly less than the native speaker writers.

8.3. Pedagogical implications

Having summarised and discussed the findings in Section 8.2 above, I now move on to consider the pedagogical implications that may be drawn from the findings of the present thesis. The findings have showed several major disparities between the Arabic speakers who were learning English and native speakers of English in terms of how the resource of SN was used in written texts. These discrepancies in the usage of the resource between the writers of the two groups largely suggested that the learner group was less appreciative of the phenomenon of SNs. It is reasonable to suppose that other similar learners would have similar limitations. The insufficient awareness of the phenomenon on the part of learners of English in general is widely raised by teachers of English writing (Flowerdew and Forest, 2015, p.188). As noted by Flowerdew and Forest (2015, p.188), one of the things that teachers of English are usually concerned with is that "second-language learners (writers in particular) are not familiar with SN usage, that they do not use SNs enough, if at all, or that they may avoid using them". The question that arises then is what should be done in order to help second-language learners of English become familiar with SNs, so that they use these nouns appropriately in their writing.

I strongly believe that SNs are useful for learners of English and are thus worthy of attention in terms of pedagogy. However, the phenomenon of SNs, as rightly pointed out by Flowerdew (2002), has not received the attention it deserves in pedagogic materials. Based on personal experience as a teacher of English writing, SNs are not talked about much in writing textbooks. Figure 1-1 presented in the introduction chapter of the present thesis is

repeated below as Figure 8-1 as an illustrative example of the issue. The advice on causality exemplified in Figure 8-1 does not mention SNs.



Part 2 Developing Writing Skills

Developing Cohesion and Clarity

GIVING REASONS WITH *BECAUSE, SO,* AND *THEREFORE*

When you give reasons to support your opinions, you may want to use connectors that show cause or result. Note the different punctuation and capitalization in sentences with these three connectors.

Use *because* in phrases and clauses that state a cause or reason.

Examples

Because large schools offer many different courses, students have a wide variety of subjects to choose from.

Figure 8-1 *An extract taken from a writing textbook (Pavlik and Segal, 2007, p.9)*

Therefore, both material writers and teachers of English writing should be aware of the importance of SNs in discourse and that awareness should be reflected in what is produced for the teaching of writing (writing textbooks) and what goes in classroom (teaching). Textbook writers, on the one hand, may integrate SNs within parts of textbooks where concepts such as cohesion, organisation, stance, and causality are addressed and, on the other hand, teachers of English writing may need to prepare teaching materials that include SNs in the teaching of the aforementioned discourse concepts. They may also need to consider SNs when giving feedback to students as will be discussed below in Section 8.3.2. Generally, if that awareness-raising approach towards SNs is taken on board, the English learners would be more likely to be made aware of the roles that SNs play in discourse.

The obvious question that may arise from the previous paragraph is that of what and how SNs should be taught. There is no definite answer to this question. However, suggestions may be offered here. In terms of which SNs to teach, decisions should be made on the basis of findings obtained from a learner corpus. For instance, the findings of the present study have clearly shown that Arabic learners who are learning English do not use SNs as an available lexical resource to signal stance and causality as much as native speakers of English do. Hence, more teaching focus on SNs with these two discourse functions would seem to be needed when teaching this group of learners.

In response to the second part of the question raised earlier, a number of suggestions on how the phenomenon of SNs might be approached in the classroom have been already made in the literature (Francis, 1988; Flowerdew, 2003). As rightly suggested by Francis (1988) and Flowerdew and Forest (2015, p.190), any pedagogical treatment of the phenomenon of SNs in classroom is to be considered in the context of *genre-based* approach (e.g. Christie and Martin, 1997, Hyland, 2003, 2004), an approach concerned with how language is used to make meaning in its context of use. This is because SNs can be expected to be used differently in different genres as Flowerdew and Forest (2015) have noted. Therefore, SNs cannot be taught in isolation but they must be taught in the context of teaching the genre. For example, when teaching the genres of 'exposition' and 'discussion' to learners, writing teachers might consider the possible SNs that are more likely to occur in such genres. Examples of these SNs are, to list a few, 'reason', 'cause', 'issue', 'argument', 'position' and 'issue'. Thus, teachers may teach their students how to use these nouns in their writing. The genre approach is useful to learners of English as it provides them with an explicit instruction "on how writing works to communicate" (Hyland, 2007, p.150). It "can [also] empower learners to go beyond propositional meaning, introducing them to ways to make their

intended meanings clear and signpost to their readers/listeners how to proceed through the text" (Flowerdew, 2015, p.31). Thus, the genre approach is highly relevant and is thus recommended for the teaching of SNs.

Francis (1988) has proposed four teaching activities that she used with her L2 English students. Her aim was to raise her learners' awareness of retrospective nouns (nouns referring to a previously mentioned stretch of text). Though her approach was designed to the teaching of retrospective nouns, it can be applied to the teaching of SNs in writing, however, with one caveat. This is that her approach should be modified to include SNs that are specified either cataphorically across-clause boundary or within-clause boundary. The following is a summary of the four sets of activities implemented by Francis (1988):

(1):

- The learners were provided with short texts involving effective use of retrospective nouns.
- The learners were asked to identify the specifics of the nouns.
- The learners engaged in the discussion of evaluative pre-modifiers of the nouns.
- The learners were asked about what would happen in case the retrospective nouns were replaced by "this". The learners were made aware of that if "this" was used in lieu of the retrospective nouns, the result would be a discourse that was ambiguous and less precise.

(2):

- Retrospective nouns were omitted and learners were presented with a list of nouns and asked to choose one appropriate noun from the list.

(3):

- Retrospective nouns were omitted and learners were asked to use an appropriate noun. This time, no alternative nouns were given.

(4):

- The learners were given a short text in which a retrospective noun was used and the learners were asked to provide a follow-up sentence or follow-up sentences.

The other approach to the teaching of SNs has been recommended in Flowerdew (2003). Unlike Francis' approach mentioned above, Flowerdew (2003) puts learners at centre stage in the process of learning SNs and make them *linguistic researchers* (a term coined by Johns, 2002, p.108). Flowerdew (2003) argues that learners might themselves benefit from concordancing tools to find out how SNs are used in terms of patterns and functions. The author also suggests that learners need to be trained to make full use of what concordancing tools offer— i.e. learners can be shown how to look for SNs by either using concordancing tools themselves or by using already printed out materials of SNs exported from concordancing tools. This approach to the teaching of SNs is promising for two reasons. One reason is that it encourages learners to discover how lexical items (SNs in this case) are used in terms of patterns and functions (see, Johns, 1991 for further arguments about the usefulness of concordancing tools in language learning). The second reason is that corpus tools have become at the disposal of both teachers and learners (Thurstun and Candlin, 1998; Flowerdew and Forest, 2015).

So far, I have addressed what and how SNs can be approached in the teaching of writing to second language learners in general. Now, I turn to the pedagogical implication based on the findings of the present study. As noted in the beginning of this section (Section 8.3), the comparison of SN usage between Arabic speakers who were learning English and native speakers of English showed several major differences and that these discrepancies suggested a reduced appreciation of the resource of SNs on the part of the learner group. This would seem to indicate that there are a number of areas in which this group of learners may benefit from explicit teaching of SNs. The following sections (8.3.1, 8.3.2 and 8.3.3) address these areas.

8.3.1. Refining current advice: the SN *thing(s)*

The SN *thing(s)* was found to occur significantly more in the learner corpus compared with the native speaker corpus. In view of the common teaching practice of writing (see Figure 5-2 in Chapter 5), the first response to this finding from writing teachers might well be that learners would be better taught to avoid using the noun in writing. While this recommendation is valid to a certain extent, it is only half the story. I would like to add that, based on the findings of the present study, the learners of English writing should not be warned not to use the noun per se but instead should be made aware of the patterns associated with the SN *thing(s)* that should be better avoided. The detailed analysis of the patterns that occur with *thing(s)* in the two corpora, as presented in Chapter 5, may provide the basis of a *pedagogic grammar* for the noun.

A teaching technique for the SN *thing(s)* might be to show the Arabic speakers who are learning English (as well as learners of English in general) how the noun is used in discourse. In other words, the learners should be made aware of the natural and appropriate uses of the noun as well as other uses where the noun could be avoided. The following are three sets of examples of the uses of the noun that may be good starting points in the teaching of how the noun is used in writing. The examples are all taken from the learner corpus of the present study.

(1):

- The best **thing** about that vacation was that I learned a lot from it,
- And the best **thing** about this movie was that it contains every thing fun, drama, action, adventure and everything.
- We rarely see each other outside Dubai, so we thought the best **thing** to do is to go out together.

- The first **thing** I decided to do when I arrived Oman was to visit my friends in Oman.

(2):

- In the UAE you can see many **things**, you can see how the desert is being transformed into an oasis,
- I have learned many **things**, the most important is that not to rely on anybody to do my work.
- We did many enjoyable **things** such as visiting the important place they have like parks, streets and gardens.

(3):

- it was strang **thing** that people live on the mountian with out shops , electrcity even laxurias places.
- It is a wonderful **thing** that you have a good memory about something in your life.

One teaching activity could be to ask the learners to find alternatives to the SN *thing(s)*. In the first set of examples, the learners would find it difficult to find alternative nouns that could replace the SN *thing*. In the second set of examples, the learners would possibly find it much easier to replace the SN *things* with more precise nouns. In the third set of examples, however, the learners would find it possible that the messages of the examples could be communicated without the use of the SN *thing* at all. The learners should be introduced the notion of pattern. The learners should be shown that the SN *thing* in the first set of examples is used in two patterns: (1) *the* + superlative + *thing* + post-modifier + v-link + specifics: that appositive clause and (2) *the* + superlative/ ADJ + *thing* + post-modifier + v-link + specifics: to infinitive clause. They should be made aware that these two uses of the noun sound appropriate and natural and thus are acceptable in writing.

The uses of the SN *things* in the second set of examples, however, are illustrative of cases in which the noun should be avoided and replaced with precise nouns. The noun is used in the following pattern: quantifier + (ADJ) + *things*. The instances of the noun occurring in this pattern were significantly higher in the learner corpus compared with the native speaker corpus. The Arabic speakers who are learning English should be shown how rarely such a usage of the noun was in the native speaker corpus and thus they should be encouraged to replace the noun with more precise alternatives. For example, the learners could replace the SN *things* in the first example of the second set of examples listed above with the SN 'changes'.

The uses of the SN *thing* in the third set of examples illustrate cases in which the noun could be omitted without affecting the overall meanings of the text. The noun occurs in the following pattern: *it* (impersonal) + v-link + ADJ + *thing* + specifics: that appositive clause. The use of the noun in this pattern is redundant and thus the learners should be taught to avoid such a redundant use of the noun.

8.3.2. Expanding the repertoire: SNs and causality

As noted earlier in Section 8.2.4.2, a comparison of the learner and native speaker corpora showed that SNs as a resource to signal causality in written texts was used significantly less in the learner corpus in terms of token and types and that the learners, unlike the native speakers of English, were found to make an extensive use of the conjunct *because* in lieu of the SN *reason(s)* to signal causality. The present study has demonstrated the importance of SNs as a resource to signal causality in written texts. The use of causal SNs (e.g. *reason(s)*) instead of causal conjuncts (e.g. *because*) provides writers with the opportunity to signal causality in a more sophisticated way. The sophistication of use, as discussed in Section 7.4.1, is summarised in the following two points. First, by using causal SNs and not causal conjuncts, writers would have the advantage of being able to enumerate and evaluate the offered causal

meanings (Hoey, 1993). Second, the use of nouns to express causality demonstrates a higher level of language development as it involves a degree of grammatical abstraction in the construction of causal meanings in text. These two points are illustrated with examples in Section 7.4.1. If we consider these two points in view of how causality was signalled by the learners as mentioned in the beginning of this paragraph, it is evident that the Arabic speakers who were learning English did not develop control over the resource of causal SNs in their writing. Given the fact that the control over causal SNs is essential in academic writing, this group of learners would benefit from explicit instruction that would encourage them to use causal SNs in their writing.

One teaching strategy to encourage learners to use causal SNs might be through teachers' feedback on learners' writing. To illustrate how the feedback might be given, a paragraph taken from the learner corpus of the present study is presented below and possible feedback is discussed afterwards:

> It was the perfect holiday for me **because** I was with my bother and my sister, the weather there was fantastick, the people were so kind and humble and finally the hotel where we stayed was so perfect and the serves were very good. I really enjoyed my holiday there and I hope I can go there one more time. **CEPA 5**

When responding to the learner's paragraph above, one of the criticism that might be levelled at the paragraph by writing teachers would probably be its spelling and its punctuation. However, these two concerns even if amended would not transform the paragraph into a mature piece of writing. Therefore, the learner should be explicitly made aware of the central issue of the paragraph (the organisation of the causal meanings expressed in the text). The learner in the paragraph above used the conjunct *because*, followed by a number of causal meanings compressed in one long run-on sentence. The learner should be made aware that

this would make the writing sound immature and ‘spoken-like’. It could be argued, however, that the spoken like nature of this extract is appropriate to the topic, a point that I will return to later in Section 8.5. The teacher could offer the learner alternatives and could encourage the learner to use the SN *reasons* instead of the conjunct *because*—along with the appropriate pattern to occur with the noun (e.g., for a number of reasons) and the appropriate linking items (e.g., firstly, secondly, thirdly, finally). The teacher may show how that might take place by providing the learner with a revised version of the paragraph. The following is one possible revised version of the learner’s paragraph:

> It was the perfect holiday for me **for a number of reasons. Firstly,** I was with my bother and my sister. **Secondly,** the weather there was fantastic. **Thirdly,** the people were so kind and humble. **Finally,** the hotel where we stayed was so perfect and the services were very good. I really enjoyed my holiday there and I hope I can go there one more time. **CEPA 5**

The revised paragraph may have a long way to go with regard to the development of arguments but as a first step it may work well to raise the learner’s awareness of the resource of causal SNs and of the effect of such a resource in the organisation of texts expressing causal meanings.

So far, I have shown how feedback might be given to learners as a way to encourage them to use causal SNs in their writing. Feedback, though important, is only one step that might be considered by writing teachers. Other steps might include preparing teaching activities that incorporate causal SNs in the expression of causality, asking learners to use concordancing tools to look for causal SNs such as *reason(s)*, *cause(s)*, *result(s)* and *factor(s)* with the aim to encourage learners to observe how such nouns are used in text and so on.

8.3.3. Introducing the personal: SNs and stance

As the title of this section suggests, SSNs are one of the resources by which writers can bring in their personal voice in their writing. What makes this resource so particularly important is that SSNs can enable writers to avoid sounding 'too personal' when they voice their opinions on the topics they address. This is something that is generally valued in many genres of writing, more particularly academic ones. In such genres, writers are often expected to voice their opinions on the topics they address without these opinions sounding too subjective (Recski, 2004). Based on personal experience as a writing teacher, Arabic speakers who are learning English when writing in English often have difficulty in presenting their opinions without sounding too personal. This observation is also true for other learner populations (see, e.g. Granger, 1998b; Gilquin and Paquot, 2008). The obvious way to sound too personal when giving an opinion about something in writing is for a writer to say, for example, 'I think 'something' is bad'. There are a number of ways to present the example in a way that the strong presence of the writer is avoided. One obvious way is to avoid using the phrase 'I think'. Another way might be made possible by using an SSN in phrases such as 'the problem is that' and 'this problem' with the latter phrase a description of the problem needs to be presented first and this description is then referred to as a problem. By using an instance of an SSN, the writer's stance can be signalled in text without it being marked as too personal. It is this aspect that makes SSNs vitally important in writing.

As stated earlier in Section 8.2.4.1, the essays written by the Arabic speakers who were learning English when compared to those written by native speakers of English were characterised by strikingly lower occurrences of SNs expressing stance and by less flexibility in the use these nouns in general. Given the fact that the expression of stance through SNs is prevalent in academic discourse and is of importance in the construction of arguments (e.g.

Jiang and Hyland, 2015; Charles, 2003; 2007), this group of learners (as well as other learners of English by extension) would highly benefit from explicit teaching of stance construction through SNs.

The present study has demonstrated the patterns associated with the construction of stance through SNs—along with examples of stance signalling nouns representing these patterns. Teachers might use some of these patterns along with the examples as one way to encourage learners to see how stance-signalling nouns are used. Teachers also might discuss the reasons behind the choices of stance signalling nouns observed in the examples.

8.4. Assessment of the methodology

This study has investigated the use of SNs in the written texts by Arabic speakers who are learning English and by native speakers of English. The study has broadly followed Flowerdew and Forest's (2015) methodology with certain modifications to accommodate learner language and to extend our understanding of how SNs are used in general (see Chapter 3). However, the general approach towards the phenomenon of SNs advocated by Flowerdew and Forest (2015), which emphasises that SNs are to be looked at from a discourse perspective, has been found to be effective in the present study. The benefit of this approach is that the identification of SNs is not based on pre-specified patterns. Following this approach, the present study has been productive in retrieving all SN instances in the learner and native speaker corpora, a point that has helped in giving a broader picture of how SNs were used in the two corpora.

The study has also demonstrated the effectiveness of combining quantitative and qualitative analysis as a way to uncover how SNs are used.

8.5. Concluding remarks and suggestions for further research

I have now come to the final section of the thesis. In this section, I would like to present my concluding remarks about the present thesis and also would like to spell out possible areas of future research on the phenomenon of SNs.

To begin with, the thesis has investigated the use of SNs in the writing by Arabic speakers who are learning English. The study has demonstrated the importance of SNs in the development of learner proficiency. As a teacher, I have become much more aware of the importance of SNs to students' writing.

As noted earlier in this chapter as well as in Chapter 1, SNs are frequent in academic prose (Flowerdew, 2003; Flowerdew 2006; Flowerdew and Forest, 2015) and probably in writing of a similar level of formality. The present study has demonstrated that SNs are less used in very informal genres. This brings us to the comment I would like to raise about the topics that the Arabic learners who were learning English were asked to write about. The topics that the learners were asked to write about, as discussed in Chapter 3, were best/ worst holiday, favourite possession and so on. The present study has demonstrated that such topics are not really preparing the students for writing academic type things. This raises a question about the appropriateness of these kinds of questions as to whether such questions would encourage learners to develop the set of skills they would really need at a university level. The topics set for the learners examined in the current thesis largely encourage descriptions and narratives but not arguments. This may explain the fewer instances of SNs in the learner corpus. I think if the aim is to teach learners to write argument, the teaching practice should reflect that by giving learners topics that are more likely to prepare them to write about topics that encourage arguments, which is at the heart of academia. When that is done, the learners are

encouraged to practice and thus develop a wide range of aspects present in academic language; SNs is one aspect of that type of language.

As for suggestions for further research on SNs, there are a number of directions. Given what I have said in the previous paragraph, it would be interesting to test the notion that different topics lead learners to different use of SNs. This is one possible area in which the present study might be taken further. Carrying out a genuine longitudinal study on SNs instead of the current pseudo longitudinal one could be another area of future work on Arabic learners who are learning English. This might be possible by following the same students through their university years. The present thesis has raised the question as to whether the very general SNs act differently than other SNs. This in turn suggests another direction for future work. It would be interesting to look at the uses of other very general SNs (e.g. *point*, *idea*) as well as the uses of other archetypical SNs in more detail to see whether these two types of SNs behave the same.

There has been considerable, and justified, interest in the phenomenon of SNs for a number of years. The present thesis has shown, however, that research into SNs is still exciting and that much remained to be explored. I have come to the end of my thesis. I hope that my thesis has contributed to the existing literature on SNs.

APPENDIX A:

CEPA test specifications



CEPA-English Public Test Specifications

Total Time for Test: 2 hours
Number of Questions: 110

CEPA-English consists of three sections. There are no breaks between the parts of the test. Recommended times are given for each part. Instructions for each part are provided in English and Arabic.

Part 1: Grammar and Vocabulary Recommended Time: 45 minutes

Part 1 consists of 45 four-option multiple-choice grammar items and 40 four-option multiple-choice vocabulary items. Examinees are tested on their ability to choose a word to complete a sentence. In addition, the test includes word-order items in which candidates must identify the correct sentence from a group of four.

Grammar domain

- Agreement
- Articles and determiners
- Quantifiers
- Intensifiers
- Pronouns
- Prepositions
- Verb forms
- Questions
- Modals
- Conditionals
- Coordinating conjunctions
- Subordinating conjunctions
- Comparatives and superlatives
- Word order
- Parts of speech

Vocabulary domain

CEPA samples from a list of 2500 high-frequency words derived from a corpus representing all major dialects of the English language. The list has been screened to ensure that the words are culturally appropriate and useful in an academic context.

Part 2: Reading Recommended Time: 45 minutes

Part 2 includes three reading texts. One is a non-linear prose text (website, advertisement, etc.) and the other two texts are descriptive, narrative, expository or a combination thereof. The texts are approximately 400 words each. There are 25 multiple-choice questions in this section.

Question types focus on:

- Identifying the main idea
- Simple factual details
- Finding the meaning of an unfamiliar word
- Pronoun reference
- Inference (understanding what is implied)
- Sequencing (understanding the order of events)
- Textual coherence

Part 3: Writing Recommended Time: 30 minutes

The Writing Section consists of one writing task. The expected length is 150-200 words. The prompt is given in English and Arabic. Candidates are rated on their language ability only.

The task is assessed by at least two markers using a holistic rubric with described levels from 0–6. Half-band scores are allowed. See the next page for the CEPA rubric.

For more information, visit our website: http://ws1.mohesr.ae/CEPA/

MOHESR-OD-CEPA-02 Version: 03 Issue date: 16/NOV/2014

CEPA Writing Bands

6	• Meaning is clear and unambiguous throughout. • Main and subsidiary points are well organized. • A range of cohesive devices is used appropriately. • Generally accurate use of simple and complex sentences. Errors rarely impede understanding. • Vocabulary choice is generally adequate, but may be inadequate to express a wide range of ideas with precision. • Most of the time, appropriate choice of words, idioms and register gives the text a feeling of fluency. • Occasional errors in spelling may still occur. • Punctuation is used appropriately.
5	• Meaning is generally clear and unambiguous. • Main and subsidiary points are generally well organized. • A range of cohesive devices is used, though not always appropriately. • Simple sentences are generally correct; some complex sentences may be used, but not always accurately. • Basic errors may still occur. • Vocabulary is generally appropriate for the topic. • Appropriate choice of words, idioms and register occasionally gives a sense of fluency. • Spelling errors can intrude, but do not impair meaning. • Punctuation is used appropriately, with only occasional errors.
4	• Meaning is generally clear, but can become unclear in complex communication. • Simple cohesive devices used appropriately. • Can construct simple sentences but basic errors (e.g., subject verb agreement and tense) still occur. • Attempts complex sentences. • Range of vocabulary becomes wider, but may be inappropriate. • Text is stilted. • Spelling errors can intrude, though words are mainly recognizable. • Uses capital letters and full stops almost without error; commas and apostrophes missing or misused.
3	• Meaning is clear in short, straightforward communications but becomes unclear if content is longer or more complex. • Some attempt to organize ideas, but little evidence of cohesive devices. • Produces simple sentences with some awareness of, but limited control of, basic sentence structure. • Vocabulary is related to the topic but limited to the simplest range. • Spelling of familiar words is generally accurate, but unfamiliar words may be unrecognizable. • Uses capital letters and full stops most of the time.
2	• Can convey only the simplest ideas. • Attempts to produce short sentences and phrases independently, but with little control of sentence structure. • Vocabulary is limited to common words. • Can spell a few common words accurately. • Some evidence of punctuation, but usually inaccurate.
1	• Message is frequently unclear. • Essentially unable to make sentences or multi-word messages. • Vocabulary is extremely limited. • Spelling errors are frequent, even in common words, except when copied. • May confuse upper and lower case letters. • Punctuation is mainly missing or inaccurate. • Text may be so short that it is difficult to assess meaningfully.
0	• Any of the following: • No sample available. • Whole text appears to be copied, canned or memorized. • Arabic, random typing or illegible handwriting (PBT only).

For more information, visit our website: http://ws1.mohesr.ae/CEPA/

APPENDIX B:

Candidate SNs in descending order according to their overall raw frequency (signalling and non-signalling) in the learner corpus

Rank	Singular	Overall raw frequency	Plural	Overall raw frequency	Total
1	Holiday	1625	Holidays	97	1722
2	Time	867	Times	75	942
3	Thing	352	Things	365	717
4	Place	506	Places	189	695
5	Way	165	Ways	14	179
6	Trip	157	Trips	7	164
7	Favourite/favourite	147	Favourites/favourites	3	150
8	Hope	147	Hopes	1	148
9	End	137	-	-	137
10	Story	131	-	-	131
11	Vacation	108	Vacations	5	113
12	Problem	57	Problems	46	103
13	Kind	68	Kinds	28	96
14	Use	88	Uses	5	93
15	Reason	39	Reasons	43	82
16	Rest	79	-	-	79
17	Thought	70	Thoughts	3	73
18	Plan	61	Plans	7	68
19	Moment	40	Moments	14	54
20	Part	44	Parts	10	54
21	Feeling	49	Feelings	3	52
22	Dream	40	Dreams	9	49
23	Idea	43	Ideas	6	49

Rank	Singular	Overall raw frequency	Plural	Overall raw frequency	Total
24	View	40	Views	6	46
25	Accident	43	Accidents	1	44
26	Activity	32	Activities	12	44
27	Change	39	Changes	4	43
28	Opinion	42	Opinions	1	43
29	Addition	42	-	-	42
30	Journey	41	-	-	41
31	Experience	37	Experiences	3	40
32	Group	34	Groups	6	40
33	Minute	12	Minutes	26	38
34	Mean	18	Means	19	37
35	Example	30	Examples	5	35
36	Word	16	Words	19	35
37	Fact	32	Facts	1	33
38	Nature	31	-	-	31
39	Chance	29	-	-	29
40	Information	27	Informations	2	29
41	Job	25	Jobs	3	28
42	Advice	26	Advices	1	27
43	Death	26	-	-	26
44	News	24	-	-	24
45	Conclusion	23	-	-	23
46	Event	7	Events	15	22
47	Matter	18	-	-	18
48	Point	15	-	-	15

APPENDIX C:

Non-lemmatised SNs in descending order according to raw frequency in the learner corpus

Rank	Singular	Raw frequency	Plural	Raw frequency	Total
1	Holiday	497	Holidays	8	505
2	Place	267	Places	47	314
3	Time	294	Times	11	305
4	Thing	159	Things	109	268
5	Trip	112	Trips	0	112
6	End	91	Ends	0	91
7	Story	86	Stories	0	86
8	Reason	37	Reasons	38	75
9	Problem	42	Problems	15	57
10	Vacation	54	Vacations	0	54
11	Kind	24	Kinds	26	50
12	Addition	42	Additions	0	42
13	Way	34	Ways	7	41
14	Idea	38	Ideas	2	40
15	Opinion	40	Opinions	0	40
16	Moment	29	Moments	6	35
17	Example	30	Examples	4	34
18	Part	32	Parts	2	34
19	Activity	26	Activities	4	30
20	Group	26	Groups	4	30
21	Experience	27	Experiences	2	29
22	Fact	28	Facts	0	28
23	Chance	27	Chances	0	27

Rank	Singular	Raw frequency	Plural	Raw frequency	Total
24	Journey	26	Journeys	0	26
25	Conclusion	23	Conclusions	0	23
26	View	19	Views	2	21
27	Dream	16	Dreams	2	18
28	Feeling	17	Feeling	1	18
29	News	18	-	-	18
30	Rest	17	-	-	17
31	Event	6	Events	7	13
32	Plan	12	Plans	1	13
33	Death	12	Deaths	0	12
34	Word	5	Words	7	12
35	Accident	11	Accidents	0	11
36	Job	10	Jobs	1	11
37	Point	11	Points	0	11
38	Favourite	8	Favourites	1	9
39	Minute	6	Minutes	3	9
40	Information	8	-	-	8
41	Change	5	Changes	1	6
42	Hope	6	Hopes	0	6
43	Matter	6	Matters	0	6
44	Nature	5	Natures	0	5
45	Use	3	Uses	2	5
46	Mean	1	Means	3	4
47	Advice	3	-	-	3
48	Thought	2	Thoughts	0	2

APPENDIX D:

Candidate SNs in descending order according to their overall raw frequency (signalling and non-signalling) in the native speaker corpus

Rank	Singular	Raw frequency	Plural	Raw frequency	Total
1	Time	1105	Times	28	1133
2	Job	620	Jobs	142	762
3	Part	688	Parts	9	697
4	Ban	188	Bans	62	250
5	Way	183	Ways	36	219
6	Experience	176	Experiences	15	191
7	Need	135	Needs	34	169
8	Thing	72	Things	96	168
9	Place	67	Places	84	151
10	Right	98	Rights	37	135
11	Reason	43	Reasons	84	127
12	Skill	16	Skills	88	104
13	Business	73	Businesses	14	87
14	Case	53	Cases	26	79
15	Idea	62	Ideas	13	75
16	Problem	45	Problems	28	73
17	Opinion	54	Opinions	17	71
18	Choice	55	Choices	9	64
19	Issue	51	Issues	11	62
20	Kind	48	Kinds	11	59
21	Fact	49	Facts	6	55
22	Activity	12	Activities	42	54
23	Benefit	18	Benefits	35	53

Rank	Singular	Raw frequency	Plural	Raw frequency	Total
24	Responsibility	42	Responsibilities	11	53
25	Situation	44	Situations	8	52
26	Example	48	Examples	2	50
27	Habit	36	Habits	14	50
28	Smell	41	Smells	6	47
29	Use	44	Uses	2	46
30	Matter	39	Matters	5	44
31	Effect	17	Effects	26	43
32	Statement	41	Statements	1	42
33	Position	31	Positions	10	41
34	Freedom	36	Freedoms	4	40
35	Risk	27	Risks	13	40
36	Opportunity	29	Opportunities	10	39
37	Cause	37	-	-	37
38	Point	34	Points	3	37
39	Solution	33	Solutions	4	37
40	Cost	24	Costs	12	36
41	Result	27	Results	6	33
42	Sense	32	-	-	32
43	End	27	Ends	4	31
44	Chance	26	Chances	2	28
45	Stress	24	Stresses	3	27
46	Deal	24	Deals	2	26
47	Pressure	21	Pressures	4	25
48	Thought	23	Thoughts	2	25
49	Hope	24	-	-	24
50	Culture	21	Cultures	2	23
51	Majority	22	-	-	22

Rank	Singular	Raw frequency	Plural	Raw frequency	Total
52	-	-	Means	22	22
53	Variety	21	Varieties	1	22
54	Addition	20	Additions	1	21

APPENDIX E:

Non-lemmatised SNs in descending order according to raw frequency in the native speaker corpus

Rank	Singular	Raw frequency	Plural	Raw frequency	Total
1	Time	218	Times	10	228
2	Way	98	Ways	9	107
3	Reason	28	Reasons	78	106
4	Ban	60	Bans	26	86
5	Thing	58	Things	28	86
6	Right	62	Rights	3	65
7	Idea	58	Ideas	2	60
8	Opinion	48	Opinions	6	54
9	Fact	47	Facts	2	49
10	Example	48	Examples	0	48
11	Issue	44	Issues	3	47
12	Case	36	Cases	10	46
13	Experience	38	Experiences	4	42
14	Part	38	Parts	3	41
15	Statement	40	Statements	0	40
16	Choice	36	Choices	3	39
17	Skill	7	Skills	31	38
18	Job	17	Jobs	16	33
19	Place	15	Places	18	33
20	Problem	28	Problems	5	33
21	Opportunity	26	Opportunities	6	32
22	Habit	27	Habits	1	28
23	Kind	22	Kinds	5	27
24	Result	26	Results	1	27

Rank	Singular	Raw frequency	Plural	Raw frequency	Total
25	Chance	24	Chances	2	26
26	Freedom	22	Freedoms	0	22
27	Point	19	Points	3	22
28	Addition	20	Additions	1	21
29	Solution	18	Solutions	3	21
30	Variety	20	Varieties	1	21
31	Benefit	3	Benefits	15	18
32	Sense	18	Senses	0	18
33	Majority	17	-	-	17
34	Matter	16	Matters	1	17
35	Need	14	Needs	3	17
36	Situation	14	Situations	3	17
37	End	16	Ends	0	16
38	Position	14	Positions	1	15
39	Activity	4	Activities	9	13
40	Responsibility	11	Responsibilities	2	13
41	Smell	11	Smells	2	13
42	Cost	7	Costs	5	12
43	Risk	10	Risks	2	12
44	Effect	4	Effects	6	10
45	Business	8	Businesses	1	9
46	Cause	9	-	-	9
47	Pressure	5	Pressures	1	6
48	Use	6	Uses	0	6
49	Thought	3	Thoughts	2	5
50	Deal	3	Deals	1	4
51	Stress	3	Stresses	0	3
52	Culture	2	Cultures	0	2

Rank	Singular	Raw frequency	Plural	Raw frequency	Total
53	Hope	2	-	-	2
54	-	-	Means	1	1

APPENDIX F:

Signalling and non-signalling instances of the SN *reason(s)* in the learner corpus

Signalling instances: (75)

1 the whole world . The all muslime loved to death to go there for many reasons like: Al Haj and Al _Omra. There we bouthe so thing we name it CEPA 3 200607400.txt 75 1

2 hot and solt. I liked food. I liked this town Because some reasons The building is anice. The shopping very fantastic. Finally, I hope to v CEPA 3 200611181.txt 133 1

3 nice time. The wrost holiday for us. It was last because a lot of reasons. Firstly, we stated at normal hotel. The hotel gived us food, but this f CEPA 3 200611191.txt 135 1

5 mall we watched restrand, gems and clothes. I like Al_ain mall for many reason. Firstly, it is near from my home. Moreover, I see My Friend in Al_ CEPA 3 200611351.txt 161 1

6 we returned to our home. Every body was very sad . So It was the reasons to be the worst trip I have ever had. CEPA 4 200605828.txt 299 1

7 I knew it is my room and realy I like my room for two reasons. First , because its so confortible place for me, I had anice woody desk CEPA 4 200605925.txt 316 1

8 with my family and I had a worst holiday ever for two main reasons: First. India have a very clever and good docter in addition , we went CEPA 4 200605954.txt 318 1

9 all the suitcase stolen , we w<o>a</o>s very sad to that reason , and we called the police. to solve this problem.Then my father and my CEPA 4 200607205.txt 356 1

10 CEPA 4 200607303 The main reason that people travel in holidayes is to have fun but the last time I CEPA 4 200607303.txt 363 1

11 If we don`t have holidy, the life be sad and worst. For that reason Iwill tell you about worst holidy. Me and my family went to the Jabal CEPA 4 200607409.txt 379 1

12 " but I hate this place because it was very clawedy, and there are many reasons that let me heat this place. Finally I want to say if you want CEPA 4 200607428.txt 382 1

14 but sumury in a short paragraph .. I like the movie alot For a many reason First, the name of the movie "New York" the most lovely , huge and nice CEPA 4 200607802.txt 442 1

15 computer help me to do my job on fast , easy and atractive. For these reason My computer is the best thing for me . CEPA 4 200611048.txt 535 1

16 best holiday was the latest summer It was the best one because of many reasons. Firstly, It was a last smmer holiday in the school. Secandly, My father CEPA 4 200611055.txt 537 1

17 with my father, mother and two sisters. We choosed to visit India for two reasons. First, to visit my relatives there because my mother is an Indian . Th CEPA 5 200601014.txt 582 1

18 First, to visit my relatives there because my mother is an Indian . The second reason is to enjoy our vacation there and have a really quality time there . We CEPA 5 200601014.txt 582 2

20 my last holiday was the worst holiday I had it ever ! There are many reasons that it made my last holiday the worst . Firstly, At the begaining of th CEPA 5 200601200.txt 601 1

21 . So my father canceled the picnic and we were so sad. Secondly, The main reason of why the last holiday was the worst, is that I had an accidant CEPA 5 200601200.txt 601 2

22 worst holiday I have ever Had. This was the summer vacation . There are many reasons to judge this as worst . First, my lovely friend Aisha who always goes o CEPA 5 200601245.txt 606 1

23 e fun but Unfortunatly, last year I spent worst holiday ever, because of many reasons, The first one that my father travelled a broad for bussiness, he couldn CEPA 5 200601351.txt 608 1

24 for many years to make it the best . And Tom cruise also was the reason of making the film successful. At the end I think every one should see CEPA 5 200601812.txt 622 1

25 on it . I like it very much and it is my favorite for many reasons. First,It helps me in my homework. Second, I can chat with my friends CEPA 5 200602670.txt 645 1

26 to pray and to ask allah for forgivenes. I liked this place for many reasons . The first reason is that it is the house of allah where we can CEPA 5 200603606.txt 666 1

27 to ask allah for forgivenes. I liked this place for many reasons . The first reason is that it is the house of allah where we can be ask allah CEPA 5 200603606.txt 666 2

28 country, but my best one is my school. I really like it for many reasons. First, I learn many new and useful things in the school and learning is CEPA 5 200605882.txt 715 1

29 are a big library where I can find any book I want. The second reason is that I can meet my friends and best teachers every day. I think CEPA 5 200605882.txt 715 2

30 to Muscat. It was the worst holiday I have ever had for two main reasons. First, we had a mechenical problem in the car. My father didn’t notice CEPA 5 200606060.txt 718 1

31 riqaia school it located in sharjah in Aldrari area. I like it for many reasons. First, when I enter the school I can see agreat view. Around the classe CEPA 5 200606067.txt 720 1

32 well, and encoureged me in my studies and helped me. Third, the most important reason, that I had gread friend her name is Maryam. She is funny, honest, kind, CEPA 5 200606067.txt 720 2

33 was happened, my passport was lost. I was really angry and confused For this reason we stayed at home. I tried to spend my leisure time doing anything. Also CEPA 5 200607319.txt 733 1

34 distroy it, so the peace can come back to the earth again. Fainaly, the reason why I loved the movie so much is that it has an amaizing story CEPA 5 200607675.txt 758 1

35 bring my jacket my family arrived on the top first then (came after. the reason why I where late was the poor on the stairs, at last after my CEPA 5 200607816.txt 766 1

36 the last summer because my parents were so busy all the time. For this reason, we couldn’t travel any where, in addition of this, I couldn’t talk CEPA 5 200608397.txt 810 1

37 is a rich source of information and entertainment. Entertainment is one of the reasons I use my computer. I like strategy games and I use the computer to CEPA 5 200608687.txt 818 1

38 excittment. We returned to our hometown and we were full of boreness. For that reasons it was the worst holiday in my life and that I hate it. CEPA 5 200609170.txt 825 1

39 200609914 The most beatiful place I have known is AlAin city. There are many reasons to be this place my best place and these reasons could attracted many pe CEPA 5 200609914.txt 846 1

40 city. There are many reasons to be this place my best place and these reasons could attracted many people and visitors. First, this city have a nice w CEPA 5 200609914.txt 846 2

41 ome animals, like goats,camels and hens which I enjoyed watching everyday. The reason why I called this my perfect holiday was because I got to spend great CEPA 6 200601004.txt 854 1

42 a lot of effort and spent alot of time on them. And the other reason is because they look like a black and white picture. My drawings are the CEPA 6 200607776.txt 877 1

43 beautiful place I love in Al Ain is Mubazzara. I love it for many reasons. First, because of the massive mountains and hills. Second, because of CEPA 6 200609893.txt 886 1

44 CEPA 6 200609894 I like Al-Ain very much for many reasons. There are many things to do. There is always something enteresting to d CEPA 6 200609894.txt 887 1

45 the frisbee. Finally,even though I already wrote that I don't know the reasons for loving the sea so much, ,My guess would be that I love the CEPA 6 200609997.txt 889 1

47 I know it would hold my secrets, those that other people might tell, another reason is, im sure that we, my diary and I, would never be appart because CEPA 6 200619672.txt 911 1

48 for some good websites or sometimes I play online games. One of the main reasons that I enjoy it because it has many uses and is portable and can CEPA 6 200619719.txt 915 1

49 a very special gift I got from my late grandfather. Its special for two reasons. First, it was the last gift I got from my grandfather before he passed CEPA 6 200619822.txt 929 1

50 emotions from both my cousins. It said “to one of my best sisters”. The reason I love this necklace is that it showed me that I have two more CEPA 6 200619825.txt 931 1

51 what I wanted to do and I was there for too long. My first reason is there were to many kids. What I mean by that is the people CEPA 6 200619833.txt 933 1

52 which is the root of all evil. I particularly enjoyed this trilogy for several reasons, firstly, I am a fan of Tolkiens work, secondly, the suspense created in CEPA 6 200619906.txt 939 1

53 lew to Sydney, Australia, I didn’t really see any universities there. The only reason I went there is because I wanted to see the famous Sydney Opera house. CEPA 6 200619915.txt 942 1

54 200619942 My Dream Holiday I had always dreamed of going to Japan, the reason why I love Japan so much was because ever since I was little I' CEPA 6 200619942.txt 943 1
55 awake for hours trying to remember where I left it, that's the main reason why I don't take it off. Everytime I look at it I see CEPA 6 200619945.txt 944 1
56 Italian meal we have ever had. It was a pleasant evening. There are many reasons why that specific holiday was the best. Firstly, we travelled as a big f CEPA 6 200619954.txt 947 1
57 least, Angelo and his Family who were extremely kind to us. I think these reasons altogether made this trip the best vacation I have ever had. CEPA 6 200619954.txt 947 2
58 Count of Monte Cristo" was the best movie I have ever seen for many reasons, but most importantly it deals with real life situations. "The Count of CEPA 6 200620343.txt 976 1
59 a Captain of a ship and is engaged to a pretty fiance. As a reason of his friend and one of the ships sailors envy they set him up CEPA 6 200620343.txt 976 2
60 who enjoy watching movies with great directing. In this essay I'll discuss the reasons that made me enjoy the movie. The movie talks about the life of a CEPA 6 200620357.txt 978 1
61 the 'brothers' decided to take revenge on the people who had her killed. The reason I liked it was because it showed what life in America, (which is where CEPA 6 200620391.txt 980 1
63 Greek soldiers from entering the city. I enjoyed the movie a lot for several reasons, one because it was based upon a true story, a story which was written CEPA 6 200620463.txt 998 1
64 riter "Homer" who wrote about this war in his famous epic "The Illiad" another reason for enjoying the movie was because it contained many battle scenes which CEPA 6 200620463.txt 998 2
65 een trees and flowers with great aromas. The vacation was wonderful for lots of reasons, one of them is having all those great places. The other reason is that CEPA 6 200620539.txt 1009 1
66 lots of reasons, one of them is having all those great places. The other reason is that the people in Tehran show a lot of respect towards all nationalit CEPA 6 200620539.txt 1009 2
67 . I wear it everywhere. Not a day passes without me putting it on. The reason I fancy it so much is because that it was a gift from someone CEPA 6 200620562.txt 1014 1
68 so much is because that it was a gift from someone special. And another reason is that the necklace is simple, it reminds me about peace and harmony. It CEPA 6 200620562.txt 1014 2
69 it was the best trip I had been on in my entire life. The reason I say this is because, firstly Australia is a magical country, as it has CEPA 6 200620613.txt 1018 1
70 all my childhood friends, as well as my childhood fun. But the most important reason that will keep this diary with me forever, is the fact that this diary CEPA 6 200620644.txt 1022 1
71 , but I am very proud to hold it in my hands. The most important reason why I like it is because it is a gift from my mother and CEPA 6 200620739.txt 1035 1

72 can say that it has been my worst holiday ever! There are indeed many reasons supporting this. I've been to London and the conditions their were absol CEPA 6 200620801.txt 1042 1

73 the world, I'd have to choose Boston, MA as a perfect destination. The reason why it would be Boston is because its an amazing metropolitan and industr CEPA 6 200621038.txt 1059 1

74 , especially to detox after all that time in the city. I guess the main reason why I'd choose Boston is because I grew up there, and miss it CEPA 6 200621038.txt 1059 2

75 to check the country out. We headed to Switzerland and reached that city, the reason why my mom and grandma decided to go to that country was for medical CEPA 6 200621190.txt 1064 1

77 . All my favourite actors were in the movie and that's one of the reasons that encouraged me to go watch it. I really loved the movie and I CEPA 6 200621207.txt 1069 1

78 the middle east, it faces the Arabian Gulf. I adore the UAE for many reasons, which I will try to explain in this essay. In the UAE you can CEPA 6 200621218.txt 1070 1

79 me, it provides safety for its people, free education and health care. But the reason I like it the most is because how fast it has transformed from a CEPA 6 200621218.txt 1070 2

80 ect Holiday My vaction to Austrilia was an unforgatable trip for me for many reasons. First I went with my entire family instead of being alone. I enjoyed my CEPA 6 200621470.txt 1075 1

81 car parking. Reading that article drov me to the edge of my seat for 2 reasons, first was because of me being a crazed fan of the PoP star, second CEPA 6 200622217.txt 1080 1

Non-signalling instances: (7)

4 CEPA 3 200611209 There are many possession I have. By reason like, acup, shoese, king and sunglass. My favourite possession It a birde CEPA 3 200611209.txt 137 1

13 importance and many of people they don`t have to many reasons. I advise all the people in the all countries of the CEPA 4 200607609.txt 410 1

19 problem is Zara father's foreced her to marry his friend son for work reasons And when her vionce knew that she is in love with an Indian man , CEPA 5 200601035.txt 584 1

46 prison. The plan worked and the captain ended up in the prison for no reason. While the captain was in prison he met an old man that gave him CEPA 6 200619370.txt 898 1

62 Everybody in life has something that they treasure dearly, for many personal reasons. As for me, my camera is my favorite possession. It used to belong to CEPA 6 200620393.txt 981 1

76 why my mom and grandma decided to go to that country was for medical reasons. I really imagined the country in a more exciting and Fun way, so it CEPA 6 200621190.txt 1064 2

82 never met the man himself.I wanted to,but couldn`t because of travelling reasons, so this is my dream to meet the King of PoP who dominate the CEPA 6 200622217.txt 1080 2

APPENDIX G:

Signalling and non-signalling instances of the SN *reason(s)* in the native speaker corpus

Signalling instances (106):

1 time job can get you experience that you can use after you graduate. For these reasons I think it is very important for college students to have a part time jo W_ENS_PTJ_XX1.txt 0 1

2 are looking for a real job or a real career. I think the most important reason they have a part time job is so that you can learn the value of W_ENS_PTJ_XX1.txt 0 2

3 "children" having fun and only being serious when it comes to school. The only reason anyone would really have a part time job is to make money so they can W_ENS_PTJ_XX1.txt 0 3

4 the employers to want to hire anyone who doesn't have a 4. 0 GPA. For these reasons, college student should focus on other things that part time work while W_ENS_PTJ_XX1.txt 0 4

5 g that decision since part time jobs in college are not that serious. For these reasons, a student should really decide for himself or for herself if a part tim W_ENS_PTJ_XX1.txt 0 5

8 part time job and it is also very important for them to do so. One reason I think this is because I have been working at a part time job for W_ENS_PTJ_XX1.txt 0 8

9 lot of this stuff is not cheap. Many people might think these are some stupid reasons to have a part time job, but to me they're perfectly good, and so W_ENS_PTJ_XX1.txt 0 9

11 we have left to be kids and not have to be too serious. For these reasons and many more, I think it is obvious that college students should not ha W_ENS_PTJ_XX1.txt 0 11

12 for college student to have a part time job, has some merit. There are several reasons why this is a good idea. Firstly, a college education is a significant e W_ENS_PTJ_XX1.txt 0 12

13 encouraging them to become independent, mature adults. Secondly, a significant reason for all young people, including college students, to have part time jobs W_ENS_PTJ_XX1.txt 0 13

14 ing to the future and seeing more college students with credit cards gives more reasons to be concerned. If parents of the students pay for everything the stude W_ENS_PTJ_XX1.txt 0 14

15 to get part time jobs. I agree with the above statement for a number of reasons. The most important reason however, is that I believe that anyone who go W_ENS_PTJ_XX1.txt 0 15

16 I agree with the above statement for a number of reasons. The most important reason however, is that I believe that anyone who goes to work, even part time, W_ENS_PTJ_XX1.txt 0 16

17 hink it is important for college students to work while they are in school. The reasons are 1) young people will get a sense of the real world, 2) employers loo W_ENS_PTJ_XX1.txt 0 17

18 ily home and to spread their wings and become independent. Ultimately, the main reasons are, freedom to do whatever you wanted, hang out with friends and to pro W_ENS_PTJ_XX1.txt 0 18
19 orld which is often difficult to comprehend from a textbook. Based on the above reasons, I support the opinion. This is to encourage students to not waste time W_ENS_PTJ_XX1.txt 0 19
20 it is important for college students to have a part time job. There are many reasons why. I'll give three. Firstly, the experience gained from working to ear W_ENS_PTJ_XX1.txt 0 20
21 ferences or contacts for future work. I agree with this statement for various reasons. The first reason is because studying costs a lot of money and it should W_ENS_PTJ_XX1.txt 0 21
22 s for future work. I agree with this statement for various reasons. The first reason is because studying costs a lot of money and it should not just be the W_ENS_PTJ_XX1.txt 0 22
23 get jobs as soon as they reach the correct age. There are a number of reasons why I believe college students should have a part time job. Firstly, stu W_ENS_PTJ_XX1.txt 0 23
24 ies of full time employment in the same organization upon graduation. For these reasons, I believe that having a part time job is very important for students. W_ENS_PTJ_XX1.txt 0 24
25 lives. College students should not work while they are in school for several reasons. First, it is difficult to focus, second, it takes time away from studie W_ENS_PTJ_XX1.txt 0 25
26 especially early on in one's life would choose money. It is after all, the reason we are in school; to earn a degree and ultimately get a job. College shou W_ENS_PTJ_XX1.txt 0 26
27 to have a part time job while they are in college. There are two main reasons for my opinion. They are concentration on studies and spending more time W_ENS_PTJ_XX1.txt 0 27
28 a poker tournament and ultimately win the grand prize money. Based on the above reasons, I am against the opinion. This allows students an opportunity for an in W_ENS_PTJ_XX1.txt 0 28
29 sant with the arguments for and against having a part-time job however, for the reasons mentioned above I stand by my belief that it is not important to have a W_ENS_PTJ_XX2.txt 1 1
30 a part-time job while at college. It is far more important to remember the reason students enrol in college which is to learn as much as they can, score as W_ENS_PTJ_XX2.txt 1 2
31 they may have accrued in a quick and efficient manner. For these and many other reasons, I feel that students should not be so worried about finding part-time e W_ENS_PTJ_XX2.txt 1 3
32 ses. All college students should have a part-time job. There are many excellent reasons to join the work force while still attending school. For some students, W_ENS_PTJ_XX2.txt 1 4
33 students working part-time, or even full-time for that matter for the following reasons. To begin, I believe by working it teaches students many valuable lesson W_ENS_PTJ_XX2.txt 1 5

34 ntial in order to become a responsible, contributing member of society. For the reasons stated above, I believe that a college student needs the experience of p W_ENS_PTJ_XX2.txt 1 6

35 everyone involved. Yes, I emphatically support this life style and for several reasons. Firstly, these are individuals who are entering into adulthood and if t W_ENS_PTJ_XX2.txt 1 7

36 t is indeed important for college students to have a part-time job, for several reasons. First, having a part-time job helps teach students how to use their tim W_ENS_PTJ_XX2.txt 1 8

37 ers though getting their own part-time job. I agree with this statement for two reasons. Firstly there is the issue of finance. Attending tertiary educational i W_ENS_PTJ_XX2.txt 1 9

38 it is important for college students to have a part-time job. There are three reasons for this. Firstly, the years spent at university are of utmost importanc W_ENS_PTJ_XX2.txt 1 10

39 it comes to getting a full-time job in, for example an office. For these reasons I do not think having a part-time job while studying is important. Even W_ENS_PTJ_XX2.txt 1 11

40 ly because so much time is required to fulfill our class assignments. For these reasons I am opposed to the idea of students taking on additional responsibiliti W_ENS_PTJ_XX2.txt 1 12

41 should hold a part-time job during the course of his studies. I have three reasons to support my claim. Firstly, college is an essential part of the academW_ENS_PTJ_XX2.txt 1 13

42 I believe that the college student must concentrate on their studies. For this reason I am against anybody working during their college days. This is a very sh W_ENS_PTJ_XX2.txt 1 14

43 ven be pressured into after hour meetings where drinking is involved. For these reasons I am completely against the college student taking on too much during th W_ENS_PTJ_XX2.txt 1 15

44 , it's not important for college students to have a part-time for three main reasons. First, part-time work is a distraction which robs precious hours away f W_ENS_PTJ_XX2.txt 1 16

45 decide to drop out of school to focus purely on their jobs. For these three reasons, it is my belief that part-time work is not essential for college studen W_ENS_PTJ_XX2.txt 1 17

46 . Either way, I think that all students should have a part-time job for the reasons mentioned above and it also gives them an opportunity to display whateve W_ENS_PTJ_XX2.txt 1 18

48 use I believe that it would distract too much from the students studies and the reason that students often fail or drop out of college is because of too many di W_ENS_PTJ_XX3.txt 2 1

49 tudent from allocating the majority of their time to studying which is the sole reason they are at college in the first place. If the student has no need to W_ENS_PTJ_XX3.txt 2 2

54 have a very difficult time adapting to the business world. Therefore, for these reasons and many more, having a part-time job can provide many benefits to the a W_ENS_PTJ_XX3.txt 2 7

55 would be better off spending ninety-five percent of their time on studying. The reason that I think that is because jobs are hard enough to get now as it W_ENS_PTJ_XX3.txt 2 8

56 sagree. I think that a large majority of students want to work anyway, for many reasons. Perhaps it may be to assist with repaying their parents for funding the W_ENS_PTJ_XX3.txt 2 9

57 added plus when students search for jobs after graduation and it is one of the reasons why I believe that part-time jobs are important. These days, any experie W_ENS_PTJ_XX3.txt 2 10

58 not smoke myself, most of my friends too often bring this up as a primary reason for doing so. I don't really care if all the restaurants in Japan ban W_ENS_SMK_XX1.txt 3 1

59 than in other areas of the country, and I also think this is a primary reason for the port image of southern health. Sure, we eat a lot of fried foods W_ENS_SMK_XX1.txt 3 2

60 there are the arguments of health and second hand smoke and more, but the real reason I think that smoking should be banned in restaurants is because it negati W_ENS_SMK_XX1.txt 3 3

61 and we can teach our children healthier habits and styles of living. For these reasons, I think it is a good idea to ban smoking at restaurants in Japan. I W_ENS_SMK_XX1.txt 3 4

62 to the atmosphere which even the best chef cannot deal with. For all of these reasons, banning smoking at restaurants in Japan would appear to be a good idea. W_ENS_SMK_XX1.txt 3 5

64 be a great idea to ban smoking at Japanese restaurants. I think this for three reasons. First of all, smoking is extremely dangerous for the health of everyone W_ENS_SMK_XX1.txt 3 7

65 be banned at restaurants. Smoking is a filthy and disgusting habit. For this reason, it is my sincere hope that one day smoking will become much rarer than i W_ENS_SMK_XX1.txt 3 8

66 do not like to kiss girls who smoke, so I think that is a good reason to have abstained from smoking too. Smoking should not be banned at all W_ENS_SMK_XX1.txt 3 9

67 many smokers waste quite a lot of money on cigarettes or chew. I think the reason that people keep smoking is because of the nicotine, because it is a drug W_ENS_SMK_XX1.txt 3 10

70 forever, but for now, I think it is fine and I don't see any reason that I should be prevented from smoking in restaurants after I have a bit W_ENS_SMK_XX1.txt 3 13

71 satisfying to have some dessert or a cup of coffee after a meal. One last reason is that smoking is a great way to meet new people. I have met countless W_ENS_SMK_XX1.txt 3 14

72 the idea that smoking should be banned at restaurants in Japan and I have many reasons for this opinion. First, I think that smoking is very dirty and is very W_ENS_SMK_XX1.txt 3 15

74 decide on. I strongly agree with the above statement for a number of obvious reasons. By now we all know the facts regarding second hand smoke and the health W_ENS_SMK_XX1.txt 3 17

75 I think smoking should be completely banned in all restaurants in Japan. The reasons are quite simple. First and most important smoking is dangerous to healt W_ENS_SMK_XX1.txt 3 18
76 icularly in restaurants is bad for the taste of food. To talk about the biggest reason, is almost obvious, smoking is well known to cause a variety of dangerous W_ENS_SMK_XX1.txt 3 19
77 tatement that smoking should be banned from all restaurants in the country. The reasons for my opinion are as follows: second hand smoke, distinctiveness of Jap W_ENS_SMK_XX1.txt 3 20
78 without asking if it is okay even while I'm still eating. For the above reasons, I agree 100% with the statement. I agree with this opinion. I have th W_ENS_SMK_XX1.txt 3 21
79 ons, I agree 100% with the statement. I agree with this opinion. I have three reasons. First, all non smokers are paying customers and have the right to enjoy W_ENS_SMK_XX1.txt 3 22
80 tclubs who may also benefit from lowered insurance premiums. Based on the above reasons, I support the opinion. This is fighting for the rights of non smokers a W_ENS_SMK_XX1.txt 3 23
81 nned from all the restaurants in the country. The largest and most undisputable reason is that smoking shortens people's lives. It is definitely the largest cau W_ENS_SMK_XX1.txt 3 24
82 ve people it could make eating at that restaurant impossible. Safety is another reason why smoking should be banned. Lighting fires indoor is one of the most ob W_ENS_SMK_XX1.txt 3 25
83 why smoking should be banned. Lighting fires indoor is one of the most obvious reasons for buildings to burn down, and with the potential for a cigarette butt W_ENS_SMK_XX1.txt 3 26
84 such a violation. Smoking in restaurants should be allowed for the following reasons: First, smoking is legal in Japan, second smoking and food and drinks go W_ENS_SMK_XX1.txt 3 27
85 njoy their meals. I would disagree with the above statement for the following reasons: Loss of income for restaurant owners and the rights of smokers. If smok W_ENS_SMK_XX1.txt 3 28
86 ore, there would be less liability from customer negligence. Based on the above reasons, I am against the opinion. This is supporting the business community. W_ENS_SMK_XX1.txt 3 29
87 le have their smoked dinners. I disagree with this statement. One of the main reasons is because it creates billions of revenue for the government that can be W_ENS_SMK_XX1.txt 3 30
91 h a complete ban on smoking in all restaurants throughout Japan. There are four reasons for this. Firstly, smoking is bad for health. It is now commonly accepte W_ENS_SMK_XX2.txt 4 4
92 icant loss of business since a blanket ban on smoking was introduced. For these reasons, I fully support a ban on smoking in restaurants. As for banning smoking W_ENS_SMK_XX2.txt 4 5
93 rmore, why stop with restaurants? Before I address this question I will give my reasons why I believe smoking should be banned in ALL restaurants. To begin, let W_ENS_SMK_XX2.txt 4 6

94 is caused by second hand smoke. From the worst to the least there are countless reasons smoking in restaurants should be banned. Not only should it be banned, i W_ENS_SMK_XX2.txt 4 7

95 that smoking should be banned in all restaurants throughout Japan. I have three reasons to support my position. Firstly, regarding the health of customers, more W_ENS_SMK_XX2.txt 4 8

96 t of recent shifts in medical awareness and social health. Therefore, for these reasons, smoking should be banned in all restaurants. I totally agree with this W_ENS_SMK_XX2.txt 4 9

97 y shows the connection between cigarette smoking and various ailments. For this reason, it is my belief that this dying breed should respect all non-smokers and W_ENS_SMK_XX2.txt 4 10

98 s of being around cigarette smoke. I agree with the statement above for several reasons. First, secondhand smoke harms non-smokers. Too many'non-smoking' sectio W_ENS_SMK_XX2.txt 4 11

100 welfare over personal freedom, and it's that simple. From there I derive my two reasons for banning smoking: I feel that banning smoking in restaurants will be W_ENS_SMK_XX2.txt 4 13

101 inclined behave responsibly themselves. I agree with this statement for several reasons. Smoking has been proven many times to have a large impact on the health W_ENS_SMK_XX2.txt 4 14

102 is a basic right to be able to smoke if one chooses, and for these reasons I do not agree with a blanket ban on smoking in restaurants. One of the W_ENS_SMK_XX2.txt 4 15

103 do not agree with a blanket ban on smoking in restaurants. One of the biggest reasons I believe smoking should be permitted in restaurants is because restaura W_ENS_SMK_XX2.txt 4 16

104 addition, having the freedom to enjoy a drink, or having a cigarette are added reasons why we pay to support restaurants. If restaurants took away smoking priv W_ENS_SMK_XX2.txt 4 17

106 at smoking should not be banned in restaurants in Japan for the following three reasons. Firstly, smoking is not illegal in society, so what right does an indiv W_ENS_SMK_XX2.txt 4 19

107 have the freedom to stay home. I disagree with the above statement for several reasons. First, we (thankfully) live not in a communist or fascist nation, but i W_ENS_SMK_XX2.txt 4 20

108 onally loathe smoking. I find it a foul smelling and disgusting habit. For this reason, I am a big fan of Starbucks, who ban smoking in their stores worldwide. W_ENS_SMK_XX2.txt 4 21

110 , this would not be such a problem! I disagree with this statement for a few reasons. There are currently a high percentage of people in the country who smok W_ENS_SMK_XX2.txt 4 23

112 being of millions of people, both smokers and non-smokers alike. For these four reasons, it is clear that smoking ought to be prohibited in all Japanese restaur W_ENS_SMK_XX2.txt 4 25

114 have ill effects on others or have fallen out of favor for any number of reasons. For example, when people smoke in public places or restaurants, there i W_ENS_SMK_XX3.txt 5 1

115 to back it up if they were to introduce a restaurant ban on smoking? The reason that I ask, is that there is sure to be some sort of uproar from W_ENS_SMK_XX3.txt 5 2
116 slation because I do not understand their culture. Despite this, there are many reasons why discouraging smoking especially in restaurants makes sense given the W_ENS_SMK_XX3.txt 5 3
117 they would continue to do so if they really love their families. For these two reasons, banning smoking in restaurants in Japan may be a good move for the Japa W_ENS_SMK_XX3.txt 5 4
118 end worldwide is for smoking in public places to be banned - and there are good reasons for this. Chief among these is that through second-hand smoke, smoking i W_ENS_SMK_XX3.txt 5 5
119 in Japan. It seems like a drastic measure, but I base my statement on the reasons I set out below. The first is that there are too many inconsiderate smok W_ENS_SMK_XX3.txt 5 6
120 right to enjoy my meal and actually taste it. Which brings me to my second reason: no self-respecting chef should allow their culinary creations to be ruin W_ENS_SMK_XX3.txt 5 7
121 that want their customers to enjoy the food as it should be tasted. The final reason why I think smoking should be banned is that the attempts to separate smo W_ENS_SMK_XX3.txt 5 8
122 places. I agree that smoking should be severely restricted in restaurants. The reason for this is the danger of secondhand smoke. Smoking does not just hurt th W_ENS_SMK_XX3.txt 5 9
123 able to go to restaurants without being exposed to this danger. There are other reasons that smoking should not be allowed in restaurants. The smoky atmosphere W_ENS_SMK_XX3.txt 5 10
124 xposed to the risk of suffering from an allergic reaction. Because of all these reasons, smoking should be eliminated from public restaurants in the country. I W_ENS_SMK_XX3.txt 5 11
125 many deaths and illnesses related to smoking. In itself, that should be enough reason for Japan to ban smoking in restaurants and all public places without arg W_ENS_SMK_XX3.txt 5 12
126 versa. Do it for the good of the people or for one of the very reasons that all taxpayers pay governments for and that is to protect the citize W_ENS_SMK_XX3.txt 5 13

Non-signalling instances (21):

6 are doing well in school and they are doing well as people, there is no reason they shouldn't have a part time job if they want to. It can also W_ENS_PTJ_XX1.txt 0 6
7 know that some of my friends have part time jobs and they always have some reason to smile when they come back from work. I think this means that they are W_ENS_PTJ_XX1.txt 0 7
10 plenty of time to work after they graduate from college, and thus there is no reason to jump the gun and start using all your effort to do something that you W_ENS_PTJ_XX1.txt 0 10
47 tepping stone toward bigger, better and maybe more lucrative opportunities. The reason is simple and I do not have all of the answers as to why, but W_ENS_PTJ_XX2.txt 1 19

50 different in that there will be students that need to work because of monetary reasons and if that is the case then the students will just have to work and W_ENS_PTJ_XX3.txt 2 3

51 think they need to do. If I am a millionaire doing a degree for whatever reasons that I may have, why the hell would I need to get a job at W_ENS_PTJ_XX3.txt 2 4

52 amily and manages to make it into college or university, then it only stands to reason that the kid may be a prime candidate for needing and getting a job. Noth W_ENS_PTJ_XX3.txt 2 5

53 n additional skills to supplement there college studies. For a great variety of reasons, it is very important for college students to hold a part-time job while W_ENS_PTJ_XX3.txt 2 6

63 ing any rules that would negatively affect one group. Therefore, there are some reasons why banning smoking at the restaurants in Japan would be a bad idea. S W_ENS_SMK_XX1.txt 3 6

68 had back when I was a kid. The cause of that, there really is no reason to smoke anymore. It is not cool, it is not healthy, and everyone knows t W_ENS_SMK_XX1.txt 3 11

69 was banned business has actually increased. I'm not sure I fully understand the reasons for this, but he told me that more families are visiting because our res W_ENS_SMK_XX1.txt 3 12

73 ter and better people. Smoking is kind of a curious habit that has remained for reasons I cannot fully understand, but I hope that one day soon we will not have W_ENS_SMK_XX1.txt 3 16

88 would also add that smoking should be banned in all public places in Japan. The reasons for the ban are obvious to all and should not be up for discussion. Apar W_ENS_SMK_XX2.txt 4 1

89 of the human race. Bans have been reasonably effective to date and I see no reason why they will not be as effective when applied to Japanese restaurants. I W_ENS_SMK_XX2.txt 4 2

90 hat terrible smell, it also increases litter and trash on the streets. For some reason, smokers seem to think that it is okay to throw cigarette butts anywhere W_ENS_SMK_XX2.txt 4 3

99 rotecting the health of those exposed and reducing the costs of healthcare. Two reasons why we should ban smoking Let me see I remember my last trip to Paris: W_ENS_SMK_XX2.txt 4 12

105 rsonal choice, as is the freedom to smoke. If people dislike smoke for whatever reason, they do not have to frequent restaurants. They can simply create their o W_ENS_SMK_XX2.txt 4 18

109 this position. First of all, I disagree on the grounds that there is no clear reason to pick on smokers. We smokers do not live in denial of scientific facts, W_ENS_SMK_XX2.txt 4 22

111 act this will have on the millions of children whose parents smoke. There is no reason why restaurants cannot continue the trend of separating smokers from non- W_ENS_SMK_XX2.txt 4 24

113 the same price as smokers do in terms of health issues and disease. For some reason, smokers get the impression that they are being discriminated against or W_ENS_SMK_XX2.txt 4 26

127 to us related to smoking, passive smoking and associated diseases, there is no reason in the world why anybody should be smoking these days anyway. I think tha W_ENS_SMK_XX3.txt 5 14

APPENDIX H:

Most frequent patterns of the SN *thing* in the learner corpus

(1): *The* + (superlative) + (ADJ) + *thing* + (post-modifier) + (v-link)* + specifics: that appositive clause (23) (14.46%)

- The best thing about that vacation was that I learned a CEPA6.txt
- In my opinion, the good thing about this film is that it sent a CEPA6.txt
- And the best thing about this movie was that it contains cepa 5.txt
- But the good thing is that we are safe, togather and ready to travel again. Cepa 5
- The most important thing that happened at that time , is that I made a torrible accident. Cepa 5
- The most important thing that I under stand in this film is that we shouldnt search for the friends because the friend come with days. Cepa 4
- The thing that was funny about that film,is that no one of them knew were the athoure on is working and cepa 4
- and the best thing was that I worked as a summer trainer in Dubai Municipality for about 2 months only. Cepa 5
- The last but not least important thing was that I was with the people who I love the most my family. Cepa 6
- The great thing about holiday that you have many Free time CEP4.txt
- I wasn`t happy there, Because the first thing happend that my brother called us from Dubai that my uncle died, cepa 4
- The most beautiful thing I enjoyed that I spent my whole time with my family, Cepa5
- The thing I knew that I was coved by popcorn. Cepa 4
- and the most thing I liked about it that we were all together and my grandmother was with us . cepa 5
- The worst thing in my opinion in this holiday that me and my sisters didn`t have any one to talk with because cepa 4
- Also the nice thing in this park that it has a train which takes me to any place I want and cepa 4
- But the really nice thing that after our dinner I was surprised that nobody had forgotten my birthday and cepa 5
- The worst thing that happened that I claimed an apple tree to get some apples while I was getting down I fell and brocked my leg, cepa 6
- and the most thing that I loved in Germany that the people are very friendly and kind with tourists. Cepa 5
- But the best thing that you spend your hoilday with the people you love. Cepa 4
- , so when we arrived there the first thing we did that we visited the profet mosque and cepa 5
- But the thing which I know that I love this small, cute and great machine . cepa 5
- The thing which was worstest that was no plane to take us back , cepa 4

(2): *The* + (superlative) + (ADJ) + *thing* + (post-modifier) + (v-link)* + specifics: other appositive clause (absence of that element (14) / wh-clause (6)) (20) (12.57%)

- Geneva has the best weather, kind people and the best thing is on August, Geneva has this year round festival many people around the world visit this huge annual event. Cepa 6
- and the most important thing is the people that you love are surrounding you the whole time and are always there to make the next day a day that was full of joy and happiness. Cepa 6
- but the must thing that make me love football is my father was play it before I play it cepa 3
- and the other thing that make me really angery is the my older sister took my room in the hotel so I had to stay in my little brother room and he was so noisy. Cepa 5
- and the sad thing was, she knows about her ill and cepa 4
- The Last thing I didit is went to ksa to do omra in Makah. Cepa 5
- The worst thing on the weekend was when I had news that my grandfather ahd died and left this world. Cepa 5
- The most important thing in life is How to enjoy and orgenize our holiday cepa 4
- The most thing I liked there was when walking infront of the sea shore. Cepa 6
- the first thing I did when I wake up I open the cirten and Cepa 4
- and the most important thing, I was studying my english writing exams in it. Cepa 5
- First thing in the morning I had to call my best friend to make an appointment with her and the other friends to hang on to the cinema cepa 5
- First thing in the morning, we packed our stuff and were ready to leave, cepa 6
- The last thing we do we visited my grandmother in her house and cepa 4
- Whin we went there the first thing we maked we changing oure clouthes and cepa 3
- The first thing, we went to macca to go around khaba seven times that every muslims do it, and cepa 5
- First thing you have to taked a cours in diving and get a license. Cepa 4
- The wondeful thing of our trip when we go to Museum and cepa 4
- The most funniest thing that happend to me there when my grand mother Lost there with my little sister, cepa 4
- The funniest thing whe she arrived to the hotel and cepa 4

(3): *The* + superlative + (ADJ) + *thing* + (post-modifier) + v-link + specifics: appositive NP/ nominalisation (19) (11.94%)

- The best thing about Grand Place is the chocolate museum. CEPA6.txt
- The most beautiful thing about Masirah would be its beautiful nature,CEPA6.txt
- I guess the most enjoyable thing about was the trip it self. CEPA6.txt
- The best thing and the more exciting was the introducing journey at the beginning of the summer program. cepa 5.txt
- The most thing I admired was their perfect musiem witch took me back to the old memories and cepa 5
- The most thing I liked there was the sea. Cepa 4
- and the most wonderful thing I've ever seen was the Great Pyramids. Cepa 6
- The best thing in Rome was their icecream and pizza. Cepa 6

- The most important thing in the computer is the internet. Cepa 6
- the most thing of my stuf I like is my mobile cepa 4
- The best thing of this holiday was the peaceful and lovely area. Cepa 5
- The most different thing there was the weather, really it cepa 6
- The most wonderful thing was our group so that why I liked it. Cepa 5
- The most fascinating thing was times square. Cepa 6
- and they most thing I love is chutting with my friend and contact with other people around the world.. Cepa 4
- The Japannes are really proud of thair calture and one of the most amaizing thing in them is thair kindness. Cepa 5
- The most intresting thing that I did there is climbing the mouting cepa 4
- The most important thing I use it for is doing my school work. Cepa 6
- The most beautiful thing was the sense of peace when I prayed there and when I put my forehead on the ground to announce that Allah is the greatest I felt unlimited happiness. Cepa 6

(4): *It* + v-link + a/the (superlative) + (ADJ) + *thing* (19) (11.94%)

- all people to read because it is agood thing and and cepa 5.txt
- I decided to join an art course. I went to the first class. I found that it was the most boring thing. Cepa 6
- , other people said that it's a very rude thing from the magazine to draw that cartoon and cepa 5
- I really think that it is the most favourite thing I have ever had. Cepa 5
- I love this computer because it is the bist thing I have ever got, Cepa5
- and it was the worst thing in my whole life . No one of my family imagen that will happend .What happend that is one of the famous magazines in Denimark puplished some cartoons about our leader and our guide after the holy Quraan Mohammad luaghing on him and on Islam, cepa 5
- My favourite possession is the computer. It was given to me as a gift on my fifteenth birthday by my father . He thought that it would be the most useful thing for me which is indeed . cepa 5
- It is importante thing in our lifes of people. Cepa 4
- I travlled to sudia Arbia withe the Whole family for Omra It's a one of avery important thing in our reilgtiors. Cepa 3
- So that I feel we should to quite this war and surche of the peace because it is the best thing in the life and cepa 4
- some people said that it is from their tradition to make a funny cartoons about anything so they called that it's anormal thing in their country , cepa 5
- After two weeks, we decides to go to kingdom of saudi arabia to do "Omrah" which it is a religous thing muslims do to fulfill their duty to their "Islam". Cepa 5
- At last we ate it. And it was the best thing that I have ever eaten. Cepa 5
- What makes this holiday perfect is I spent the best time ever with my family and from my opinion it’s the most important thing, cepa 5
- It's a very precious thing to me, I don't trust giving it to anybody. It's decorated with crystals and leather, I decorated it. It's my i-pod, cepa 6
- I will choose my compuer as the most favorite. You may ask me why you like it? Why it’s the favorite thing? Cepa 5
- I like my Mobile alot, it is my favorait thing . cepa 4

- But when you walk you can get ride of it and it's a good thing to do because obesity could casue you a lot of disseses. Cepa 5
- , and we witnessed how the make the Kashmirian scarf, and it is the hardest thing to do. Cepa 6

(5): Specifics: Nominalisation/NP + v-link + a/an/the + (superlative) + (ADJ) + *thing* (18)

(11.32%)

- Travilling around the world is a very exciting thing, cepa 5.txt
- that the country is the most important thing and we should always protect it. Cepa 6
- but some times the holiday can be the worst thing. Cepa 5
- I think a computer is a very important thing for all people. Cepa 5
- For these reason My computer is the best thing for me . Cepa 4
- In the 20th century, making films is one of the most entertaining thing for people to produce or to watch. Cepa 6
- In fact this computer is the bist thing I got, it means a lot to me, cepa 5
- Finally , I can say that the computer is very usfel and enjoyable thing. Cepa 5
- Holidays are the most interesting thing in my life, cepa 4
- I don't know why my laptop is an special thing in my life which I can't leave without ! cepa 5
- We all agree that the weekend is the most beautifull thing in our life. Cepa 5
- and learning is the most important thing in the life. Cepa 5
- Hence , In mine opinion the holidays after the studys is the best thing in the world. Cepa 4
- because the technolge is the important thing in these days. Cepa 4
- No one deny that.The mobile is the importante thing. Cepa 4
- What made it really wonderfull is going along with my older brother, having to spend time with my brother is not a thing. Cepa 5
- I think that traveling is the best thing to do in Holidays, especially traveling with relatives. Cepa 6
- Travilling and the nutchrl views are the best thing to me. Cepa 4

(6): *The* + (ADJ) + *thing* + (post-modifier) + v-link + specifics: appositive NP/ nominalisation (15) (9.43%)

- The first thing I did when I got there was jump with joy. Cepa6
- and the firest thing I have see it was the moinkes , cepa 3
- and my favourite thing is my computer. Cepa 4
- the second thing it was my car . I can't drived it because thier is some wrong with it and I have to take it to the graje cepa 4
- The thing that I liked the most about switzeland was the green hills and the mountains. Cepa 5
- The thing that made me like the movie very much is the actors that were in it and cepa 6
- The only thing that showed that there was life on this island was the skeletons on the ground and every where cepa 6

- And the last thing was my mother she tolde me that didn't go out ever because of my leg and cepa 4
- the first thing we did when we got there was check in at the hotel, cepa 6
- so the First thing I did is going to the mall and shop. Cepa 5
- Days passed, and the only thing I did was hope for a better holiday next time. Cepa 6
- The only thing I liked in that holiday was going to the cinema everyweek. Cepa 6
- The main thing I loved about Germany is the people's hospitality and kindness. Cepa 6
- One of the perfect thing that we did was going shopping. Cepa 5
- The first thing was I done it was chaing my clouthes to go out. Cepa 4

(7): *The* + (superlative) + ADJ + *thing* + post-modifier + v-link + specifics: to-infinitive appositive clause (5) (3.14%)

- and One of The most buatiful Thing I can do it when I go to The sea is to sail on it and go fishing.. cepa 5
- The first thing I decided to do when I arrived Oman was to visit my friends in Oman. Cepa 5
- The first thing I did it is cut my hair because I didn't like a long hair Cepa 5
- The first thing that we've done is to book a room in a 5 star hotel. Cepa 5
- We rarely see each other outside Dubai, so we thought the best thing to do is to go out together. Cepa 6

APPENDIX I:
Most frequent patterns of the SN *thing* in the native speaker corpus

(1): The + (superlative) / another/ one + (ADJ) + *thing* + (post-modifier) + v-link + specifics: that-appositive clause (12) (20.68)

- The thing is that passive smokers die from the same diseases that smokers do and
- The thing that worries me is that the next step will be to ban smoking outside
- The thing is though, is that it might be one to smoke in a bar or pug,
- Another thing that struck me was that there course averages seemed to be slightly higher not for any one particular subject but generally speaking over the course of their degrees.
- Another thing that worked out really well was that even with my degree (IT) I couldn't always get the job that I wanted when I was travelling.
- One thing that makes this difficult, however, is that most democratic societies also put a strong emphasis on freedom of speech and action.
- The one bad thing, however, is that as expected, I do sometimes have some problems bala
- The big thing for me when I was working and studying was that I learned that I
- The good thing for them is that they can find some opportunities over the summer if
- The good thing is that if you are currently a smoker and if you quit right away, your body has the chance to heal and you may even reach a point where you are in as good of health as you were before you started smoking.
- The ironic thing is that you can achieve a similar high by doing healthy activities
- the worst thing is that secondhand smokers do not have any choice in whether they will

(2): This/ that/ which + v-link + a/an/the/one + (ADJ) + *thing* (9) (15.51%)

- Smoking should definitely be banned at all the restaurants in Japan. Many of the United States have banned smoking some years ago, and I think that there are probably a few other countries in the world that have already done so as well. This seems like an obvious thing to do because
- , they said that smoking was very prevalent in Japanese restaurants. I think this is a bad thing, because
- In history class, I learned that many Japanese people are now becoming Christians, and this is a good thing.
- Smoking is not allowed in the restaurants here, and I think this is a good thing.
- The people elect the government anyway, so why not put it to the people and let them and the restaurants decide? That would be the democratic thing to do in my view
- Governments are so concerned about the environment, the ozone layer and the ecology these days, which is a good thing in my book.

- More than the actual knowledge that they teach you in college, I think that they teach you have to learn, which is a valuable thing to know how to do so that you can quickly adapt to new situations at your workplace.
- Many of them tell me that they have made a lot of new friends through their part time job, and I can say that I think that that is a bad thing.
- Smoking is still allowed at many Australian restaurants, and this is one thing that I really hate about my home country.

(3): *The* + (superlative) (ADJ) + thing + (post-modifier) + **v-link + specifics: to-infinitive appositive clause (7) (12.06%)**

- The best thing to do is not go to smoking restaurants.
- I think that the best thing to do is to just have separate restaurants for smokers and non-smoker
- book and Therefore the most important thing is for students to focus on their studies and graduate on time
- The most important thing is not to ban smoking in restaurants in Japan.
- The most important thing is to ban smoking at all public places.
- then the only intelligent thing to do is remove the cigarettes from the equation
- the last thing I want to do is chew my food combined with the flavour of someone's leftover smoke fumes.

(4): Specifics: nominalisation/ NP + v-link + an/a + ADJ + *thing* (7) (12.06%)

- I think that the bans are a good thing and that they might help people like my dad cut down at least,
- Smoking is kind of an old fashioned thing and
- I think that smoking is a bad thing.
- Of course, not smoking around children, their own or others, is a fine thing, but why not extend their thinking a little further and think about the rest of the community.
- (the rush that smokers get from feeding their addiction could hardly be considered a positive thing).
- One less cigarette has got to be a good thing in my book and
- Lastly, a part time job is a good thing to list on your resume because

(5): The/another/ one + (superlative) + (ADJ) + *thing* + (post-modifier) + v-link + specifics: nominalisation (3)/ NP (2) (5) (8.62%)

- Another thing that many people don't know about is the ability of smoking to destroy the mind.
- and one thing that really irks me is inhaling other people's filthy cigarette smoke against my will.
- The most important thing in college is doing well because it's a cutthroat world out here in the real world and we need every edge we can get.
- the only thing separating the smokers from the non-smokers is either a small, ineffective partition or

- At this level of work, the only thing that the student is replacing his own schedule is usually time that he or she has wastes by doing

(6): ***The same thing*** **(5) (8.62%)**

- We have banned a lot in Britain and most countries are all doing the same thing now and Japan should too I reckon.
- and it is really a slap in the face when you go to a party and someone is blowing smoke in your face. The same thing is true at restaurants,
- Since coming to Japan, I see nearly the same thing as far as space goes, small spaces many smokers.
- He says when he was in college he worked more than 30 hours per week at a part time job and still graduated early. He feels that he set the bar pretty high and I should be able to at least keep up with him. Honestly I have to say that I agree, but sometimes it gets tiring to hear him say the same thing over and over again.
- Three years during which time, they were still doing the same thing, releasing toxins into the air.

APPENDIX J:
The frequency of all the adjectival types occurring with the CSN *reason(s)* across the two corpora

ADJ Types	Frequency in the learner corpus	Frequency in the native speaker corpus
Above	0	5
Following	0	4
Main	7	4
Another	3	1
Biggest	0	2
Important	3	2
First	2	1
Other	2	1
Second	2	1
Only	1	1
Personal	1	0
Primary	0	2
Good	0	2
Obvious	0	2
Excellent	0	1
Real	0	1
Significant	0	1
Enough	0	1
Undisputable	0	1
Stupid	0	1
Largest	0	1
Added	0	1
Sole	0	1
Last	0	1
Final	0	1
Total frequency	21	39
Total number of ADJ types	8	24

APPENDIX K:
The frequency of all the enumerative types occurring with the CSN *reason(s)* across the two corpora

Enumerative Types	Frequency in the learner corpus	Frequency in the native speaker corpus
Many	22	4
Several	2	7
Two	4	5
One	2	8
Three	0	8
A number of	0	3
A lot of	1	0
Lots of	1	0
Four	0	2
Countless	0	1
Various	0	1
Some	1	0
Total frequency	33	39
Total number of enumerative types	7	9

References

Al-Jarf, R. S. (2001). Processing of cohesive ties by EFL Arab college students. *Foreign Language Annals, 34*(2), 141-151.

Alharbi, A. (2011). *An investigation into the English writing of two culturally and linguistically diverse student groups in the Pre-Enrolment English Program (PEP) at Adelaide University*. (Unpublished master's thesis), University of Adelaide, Adelaide, Australia.

Altenberg, B., & Tapper, Marie. (1998). The use of adverbial connectors in advanced Swedish learners' written English. In S. Granger (Ed.), *Learner English on computer* (pp. 80-93). London: Longman.

Anthony, L. (2014). AntConc (Version 3.4.4)[Computer software]. Tokyo, Japan: Waseda University.

Aston, G. (2011). Applied corpus linguistics and the learning experience. In V. Viana, S. Zyngier, & G. Barnbrook (Eds.), *Perspectives on corpus linguistics* (pp. 1-16). Amsterdam and Philadelphia: John Benjamins.

Benitez-Castro, M.-A. (2015). Coming to grips with shell-nounhood: A critical review of insights into the meaning, function and form of shell-noun phrases. *Australian Journal of Linguistics, 35*(2), 168-194. doi:10.1080/07268602.2015.1005001

Biber, D. (2006a). Stance in spoken and written university registers. *Journal of English for Academic Purposes, 5*(2), 97-116.

Biber, D. (2006b). *University language: A corpus-based study of spoken and written registers* (Vol. 23). Amsterdam: John Benjamins.

Biber, D., & Finegan, E. (1989). Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text-Interdisciplinary Journal for the Study of Discourse, 9*(1), 93-124.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of written and spoken English*. Harlow: Longman.

Birmingham City University. (2016). Study guide: Writing. Retrieved from http://library.bcu.ac.uk/learner/writingguides/1.20.htm

Charles, M. (2003). 'This mystery…': A corpus-based study of the use of nouns to construct stance in theses from two contrasting disciplines. *Journal of English for Academic Purposes, 2*(4), 313-326.

Charles, M. (2007). Argument or evidence? Disciplinary variation in the use of the noun that pattern in stance construction. *English for Specific Purposes, 26*(2), 203-218.

Christie, F., & Martin, J. R. (Eds.). (1997). *Genre and institutions: Social processes in the workplace and school*. London: Cassell.

Conrad, S., & Biber, D. (2000). Adverbial marking of stance in speech and writing. In S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 56-73). Oxford: Oxford University Press.

Dörnyei, Z. (2007). *Research methods in applied linguistics*. New York: Oxford University Press.

Ferguson, G. (2009). Issues in researching English as a lingua franca: A conceptual enquiry. *International Journal of Applied Linguistics, 19*(2), 117-135.

Flowerdew, J. (2002). A pedagogic grammar of signalling nouns in discourse. *Revista Canaria de Estudios Ingleses*, (44), 141-155.

Flowerdew, J. (2003). Signalling nouns in discourse. *English for Specific Purposes, 22*(4), 329-346.

Flowerdew, J. (2006). Use of signalling nouns in a learner corpus. *International Journal of Corpus Linguistics, 11*(3), 345-362.

Flowerdew, J. (2009). Use of signalling nouns in a learner corpus. In J. Flowerdew & M. Mahlberg (Eds.), *Lexical cohesion and corpus linguistics* (pp. 85-102). Amsterdam: John Benjamins.

Flowerdew, J. (2010). Use of signalling nouns across L1 and L2 writer corpora. *International Journal of Corpus Linguistics, 15*(1), 36-55. doi:doi 10.1075/ijcl.15.1.02flo

Flowerdew, J. (2015). Revisiting metadiscourse: Conceptual and methodological issues concerning signalling nouns. *Ibérica,* (29), 15-34.

Flowerdew, J., & Forest, R. W. (2015). *Signalling nouns in English*. Cambridge: Cambridge University Press.

Francis, G. (1986). *Anaphoric nouns. Discourse analysis monographs 11*. Birmingham: University of Birmingham, English Language Research.

Francis, G. (1988). The teaching of techniques of lexical cohesion in an ESL setting. In V. Bickley (Ed.), *Language in a bi-lingual or multi-lingual setting* (pp. 325-338). Hong Kong: Institute of Language in Education.

Francis, G. (1994). Labelling discourse: an aspect of nominal-group lexical cohesion. In M. Coulthard (Ed.), *Advances in written text analysis* (pp. 83-101). London: Routledge.

Gilquin, G., & Paquot, M. (2008). Too chatty: Learner academic writing and register variation. *English Text Construction, 1*(1), 41-61.

Granger, S. (1998a). The computer learner corpus: A versatile new source of data for SLA research. In S. Granger (Ed.), *Learner English on computer* (pp. 3-18). London: Longman.

Granger, S. (1998b). Prefabricated patterns in advanced EFL writing: Collocations and formulae. In A. P. Cowie (Ed.), *Phraseology: theory, analysis and applications* (pp. 145-160). Oxford: Oxford University Press.

Granger, S. (2003). The international corpus of learner English: A new resource for foreign language learning and teaching and second language acquisition research. *TESOL Quarterly, 37*(3), 538-546.

Granger, S. (2004). Computer learner corpus research: Current status and future prospects. In U. Connor & T. A. Upton (Eds.), *Applied Corpus Linguistics: A multidimensional perspective* (pp. 123-145). Amsterdam: Brill Rodopi.

Granger, S. (2007). The computer learner corpus: A versatile new source of data for SLA research. In W. Teubert & R. Krishnamurthy (Eds.), *Corpus linguistics: Critical concepts in linguistics* (Vol. 2, pp. 166-182). London: Routledge.

Granger, S. (2009). The contribution of learner corpora to second language acquisition and foreign language teaching. In K. Aijmer (Ed.), *Corpora and language teaching*. Amsterdam: John Benjamins.

Granger, S. (2015). Contrastive interlanguage analysis: A reappraisal. *International Journal of Learner Corpus Research, 1*(1), 7-24.

Granger, S. (Ed.) (1998c). *Learner English on computer*. London: Longman.

Guo, X. (2006). *Verbs in the written English of Chinese learners: A corpus-based comparison between non-native speakers and native speakers.* (Unpublished doctoral thesis), University of Birmingham, Birmingham, England.

Halliday, M. A. K. (1985). *An introduction to functional grammar*. London: Edward Arnold.

Halliday, M. A. K. (1994). *An introduction to functional grammar* (2nd ed.). London: Edward Arnold.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Halliday, M. A. K., & Matthiessen, C.M. (1999). *Construing experience through meaning: A language-based approach to cognition*. London: Continuum.

Halliday, M. A. K., & Matthiessen, C. M. (2004). *An introduction to functional grammar* (3rd ed.). London: Arnold.

Hasselgren, A. (1994). Lexical teddy bears and advanced learners: A study into the ways Norwegian students cope with English vocabulary. *International Journal of Applied Linguistics, 4*(2), 237-260.

Hoey, M. (1983). *On the surface of discourse*. London: George Allen and Unwin.

Hoey, M. (1993). A common signal in discourse: How the word reason is used in texts. In J. M. Sinclair, M. Hoey, & G. Fox (Eds.), *Techniques of description: Spoken and written discourse* (pp. 67-82). London: Routledge.

Hoey, M. (2005). *Lexical priming: A new theory of words and language*. London and New York: Routledge.

Holliday, A. (2009). English as a lingua franca,'non-native speakers' and cosmopolitan realities. In F. Sharifian (Ed.), *English as an international language: Perspectives and pedagogical issues* (pp. 21-33). Bristol: Multilingual Matters.

Hunston, S. (1989). *Evaluation in experimental research articles.* (Unpublished doctoral thesis), University of Birmingham, Birmingham, England.

Hunston, S. (1994). Evaluation and organization in a sample of written academic discourse. In M. Coulthard (Ed.), *Advances in written text analysis* (pp. 191-118). London: Routledge.

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.

Hunston, S. (2011). *Corpus approaches to evaluation: Phraseology and evaluative language*. New York: Routledge.

Hunston, S., & Francis, G. (1999). *Pattern grammar: A corpus-driven approach to the lexical grammar of English* (Vol. 4). Amsterdam: John Benjamins.

Hunston, S., & Sinclair, J. (2000). A local grammar of evaluation. In S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 74-101). Oxford: Oxford University Press.

Hyland, K. (1996). Writing without conviction? Hedging in science research articles. *Applied Linguistics, 17*(4), 433-454.

Hyland, K. (2003). Genre-based pedagogies: A social response to process. *Journal of Second Language Writing, 12*(1), 17-29.

Hyland, K. (2004). *Genre and second language writing*. Ann Arbor: University of Michigan Press.

Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies, 7*(2), 173-192.

Hyland, K. (2007). Genre pedagogy: Language, literacy and L2 writing instruction. *Journal of Second Language Writing, 16*(3), 148-164.

Illari, P., & Russo, F. (2014). *Causality: Philosophical theory meets scientific practice*. Oxford: Oxford University Press.

Ishikawa, S. i. (2013). The ICNALE and sophisticated contrastive interlanguage analysis of Asian learners of English. *Learner Corpus Studies in Asia and the World, 1*, 91-118.

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*(2), 93-114.

Jenkins, J. (2006a). Current perspectives on teaching world Englishes and English as a lingua franca. *TESOL Quarterly*, *40*(1), 157-181.

Jenkins, J. (2006b). Points of view and blind spots: ELF and SLA. *International Journal of Applied Linguistics, 16*(2), 137-162.

Jenkins, J. (2007). *English as a lingua franca: Attitude and identity*. Oxford: Oxford University Press.

Jenkins, J. (2011). Accommodating (to) ELF in the international university. *Journal of Pragmatics, 43*(4), 926-936.

Jiang, F., & Hyland, K. (2015). 'The fact that': Stance nouns in disciplinary writing. *Discourse Studies, 17*(5), 529-550.

John, S. (2005). *The writing process and writer identity: Investigating the influence of revision on linguistic & textual featurtes of writer identity in dissertations*. (Unpublished doctoral thesis), University of Birmingham, Birmingham, England.

Johns, T. (1991). Should you be persuaded. *ELR Journal*(4), 1-16.

Johns, T. F. (2002). Data-driven learning: The perpetual challenge. In B. Kettemann & G. Marko (Eds.), *Proceedings of the Fourth International Conference on Teaching and Language Corpora* (pp. 107-117). Amsterdam: Rodopi.

Khalil, A. (1989). A study of cohesion and coherence in Arab EFL college students' writing. *System, 17*(3), 359-371.

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor: University of Michigan Press.

Lee, D. Y., & Chen, S. X. (2009). Making a bigger deal of the smaller words: Function words and other key items in research writing by Chinese learners. *Journal of Second Language Writing, 18*(4), 149-165.

Mahlberg, M. (2005). *English general nouns: A corpus theoretical approach* (Vol. 20). Amsterdam: John Benjamins.

Martin, J. R. (1992). *English text: System and structure*. Philadelphia: John Benjamins.

Martin, J. R. (1993). Life as a noun: Arresting the universe in science and humanities. In M. A. K. Halliday & J. R. Martin (Eds.), *Writing science: Literacy and discursive power* (pp. 221-267). Pittsburgh: University of Pittsburgh Press.

Mauranen, A. (2012). *Exploring ELF: Academic English shaped by non-native speakers*. Cambridge: Cambridge University Press.

McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. London: Routledge

Mohamed-Sayidina, A. (2010). Transfer of L1 cohesive devices and transition words into L2 academic texts: The case of Arab students. *RELC Journal, 41*(3), 253-266.

Moore, S. H. & Bounchan, S. (2010). English in Cambodia: Changes and challenges. *World Englishes*, 29(1), 114–126.

Pavlik, C., & Segal, M. K. (2007). *Interactions 2 writing: Paragraph development and introduction to the essay* (Middle East, gold ed.). New York: McGraw-Hill.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. New York: Cambridge University Press.

Péry-Woodley, M.-P. (1990). Contrasting discourses: Contrastive analysis and a discourse approach to writing. *Language Teaching, 23*(3), 143-151.

Rampton, M. B. H. (1990). Displacing the 'native speaker': Expertise, affiliation, and inheritance. *ELT Journal, 44*(2), 97-101.

Rayson, P. (2017). Log-likelihood and effect size calculator. Retrieved from http://ucrel.lancs.ac.uk/llwizard.html

Rayson, P., & Garside, R. (2000). *Comparing corpora using frequency profiling.* Paper presented at the Proceedings of the Workshop on Comparing Corpora, Hong Kong.

Recski, L. J. (2004). Expressing standpoints in EFL written discourse. *Revista Virtual de Estudos da Linguagem, 2*(3), 1-16.

Rego, I. Á. M. (2001). Encapsulations and prospection in written scientific English. *Estudios ingleses de la Universidad Complutense,* (9), 81-102.

Ringbom, H. (1998). Vocabulary frequencies in advanced learner English: A cross-linguistic approach. In S. Granger (Ed.), *Learner English on computer* (pp. 41-52). London: Longman.

Russell, B. (1913). On the Notion of Cause. *Proceedings of the Aristotelian Society, 13*, 1-26.

Schmid, H.-J. (2000). *English abstract nouns as conceptual shells: From corpus to cognition*. Berlin and New York: Mouton de Gruyter.

Seidlhofer, B. (2001). Closing a conceptual gap: The case for a description of English as a lingua franca. *International Journal of Applied Linguistics, 11*(2), 133-158.

Seidlhofer, B. (2010). Giving VOICE to English as a lingua franca. In R. Facchinetti, D. Crystal, & B. Seidlhofer (Eds.), *From international to local English—and back again* (pp. 147-163). Bern: Peter Lang.

Sinclair, J. (1993). Written discourse structure. In J. M. Sinclair, M. Hoey, & G. Fox (Eds.), *Techniques of description: Spoken and written discourse* (pp. 6-31). London: Routledge.

Tajeddin, Z., Alemi, M., & Pashmforoosh, R. (2018). Idealized native-speaker linguistic and pragmatic norms in English as an international language: Exploring the perceptions of nonnative English teachers. *Language and Intercultural Communication, 18*(3), 300-314. doi:10.1080/14708477.2017.1413105

Tåqvist, M. (2016). *'Another thing' : Discourse-organising nouns in advanced learner English.* (Unpublished doctoral thesis), Karlstad University, Karlstad, Sweden.

Thompson, G., & Hunston, S. (2000). Evaluation: An introduction. In S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 1-26). Oxford: Oxford University Press.

Thurston, J., & Candlin, C. N. (1998). Concordancing and the teaching of the vocabulary of academic English. *English for Specific Purposes, 17*(3), 267-280.

Timmis, I. (2002). Native-speaker norms and international English: a classroom view. *ELT Journal, 56*(3), 240-249. doi:10.1093/elt/56.3.240

UAE Ministry of Higher Education and Scientific Research. (2014). CEPA-English public test specifications. Retrieved from http://ws1.mohesr.ae/cepa/Files/Public_CEPA_English_Specifications.pdf

Wang, C. (2017). *Patterns and meanings of intensifiers in Chinese learner corpora*. London: Routledge.

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), 1-92.

Young, T. J., & Walsh, S. (2010). Which English? Whose English? An investigation of 'non-native' teachers' beliefs about target varieties. *Language, Culture and Curriculum, 23*(2), 123-137. doi:10.1080/07908311003797627