The effect of an online collocation dictionary on advanced learners' use of collocations in L2 writing

Dung Thi Phuong Cao

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
School of Education

September 2018

The candidate confirms that the work submitted is his/her own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Acknowledgements

The completion of this thesis would have been impossible without the guidance, support and encouragement of several people to whom I would like to extend my deepest gratitude for their contribution:

First, my deep appreciation and gratitude go to my supervisors, Dr. Richard Badger and Prof. Alice Deignan. During this time, you have always provided me with invaluable support, guidance and advice. Your experience and knowledge have added immense value throughout my process of completing this thesis.

Second, I would like to thank all my colleagues and supporting staff for helping me with collecting data. Without your support I would not have accomplished the thesis.

Third, I owe thanks also to my British and American colleagues and friends, Sidsell Millerstrom, Kendra Marcus, Peter Traynor, Ben Turner, Laura Grassick, and David Smith, who helped me a lot with the data analysis process.

Finally but most importantly, special thanks to my husband for his support, sharing, and care during one of the most difficult times in my life. Thank you for giving me time to write the thesis. To my two children, thanks for being by my side and bringing me happiness and energy to overcome stressful periods.

Abstract

This doctoral research investigates the effect of the use of the Oxford Online Collocation Dictionary on Vietnamese advanced learners' collocation use in academic writing and their perceptions of the use of the dictionary as a supportive tool. The study aims to help learners improve their collocation use, especially advanced learners who are expected to enhance their store of vocabulary on their own.

This study analysed students' written texts, questionnaires, observations, and interviews. It was carried out in two phases. In phase 1, 29 participants' 350-word essays on an assigned topic were collected as baseline data. Participants were then instructed in the use of the Oxford Online Collocation Dictionary. In phase 2, a second set of essays were collected. Observations and questionnaires were also collected at this stage and 8 participants were chosen for in-depth interviews.

Results showed that the use of this online collocation dictionary as a supportive tool produced some benefits for learners, but these were mainly psychological rather than practical effects. Learners feel confident and assured that they have the tool as a resource. However, their collocation use did not show improvement overall. The study found that learners made mistakes with V-N collocation the most. The study also found that learners often used the dictionary while doing the writing, with some references to other dictionaries for collocation meaning-checking. The dictionary was highly evaluated as a tool for collocation check-up but some enhancements, mostly of content, need to be made. In particular, it needs to provide learners with the meaning of each group of collocates, and pronunciation description; more examples are also needed to illustrate how collocations are used.

Table of Contents

List of Tables

List of Figures

List of abbreviations

BNC	The British National Corpus
CALD	Cambridge Advanced Learner's Dictionary
CALL	Computer-assisted language learning
CCED	Collins COBUILD English Dictionary
CDOEL	Collins Dictionary of the English Language
CIDE	Cambridge International Dictionary of English
Cloc	Locative/Goal Complement
Co	Object Complement
COBUILD	The Collins COBUILD Dictionary for Learners of English
Cs	Subject Complement
EFL	English as a Foreign Language
ESL	English as a Second Language
FEIs	Fixed expressions and Idioms
GRE	Graduate Record Examination
IELTS	International English Language Test System
LDOCE	Longman Dictionary of Contemporary English
MOET	The Ministry of Education and Training
MWEs	Multiword expressions
OALD	Oxford Advanced Learners' Dictionary
OALDCE	Oxford Advanced Learner's Dictionary of Current English
Od	Direct Object
OHPC	Oxford Hector Pilot Corpus
Oi	Indirect Object
OOCD	The Oxford Online Collocation Dictionary
Op	Prepositional Object
RS	Recording sheets
SAT	Scholastic Assessment Test
USSH	The University of Social Sciences and Humanities

Chapter 1 : INTRODUCTION

1.1   Background and rationale of the study

The pedagogical value of the dictionary as a source of information for language learning, collocations included, has long been emphasized by lexicographers (Wright, 1998; Hornby et al., 1974; Sinclair, 1987). Nation (2001) claims that they are helpful in a variety of tasks involving the comprehension and production of a text. General dictionaries, however, as Bogaards (2003) points out, are mainly used for receptive rather than productive purposes. In particular, they are mostly used to find the meanings of unknown words in reading tasks. Learners approach dictionaries much less frequently for productive purposes and once they approach them they mostly seek help with information on spelling; collocation searches in dictionaries are much less common (Harvey and Yuill, 1997; Bogaards, 2003). Béjoint (1989) argues that dictionaries are not used as expected by the compilers. It seems that despite the efforts of dictionary compilers in recent years to improve the description and presentation of collocational information, language learners remain unaware of this potential resource. This is probably because they lack knowledge about the nature and the importance of collocations (Harvey and Yuill, 1997; Laufer 2010; Atkins and Varantola, 1998; Nesselhauf and Tschichold, 2002)

The increasing use of electronic dictionaries in recent years may be because they are more convenient than traditional paper dictionaries (Hartmann, 2001). Although there are some concerns regarding the use of electronic dictionaries for language learning, in particular that information retrieved so quickly and painlessly will be forgotten easily, electronic dictionaries have the advantage of 'providing the user with almost instant access to a database much larger than a single book' (Nesi, 1999,p. 56). Some empirical studies have shown that compared to paper-based dictionaries, electronic dictionaries on CD-ROM or dictionaries linked to World Wide Web sites are more efficient and preferable in a variety of aspects: speed and ease of consultation, quality of information supplied (Laufer and Hill, 2000; Tono, 2001) and encouraging exploratory browsing (Nesi, 2000).

It was not until around 1960 that lexicographers started to take specific users' needs into account in the design of lexicographic resources (Bogaards, 2003). Since then, different kinds of specialized dictionaries focusing either on the scope or the coverage of subject (e.g. medical or legal dictionaries) or a specific aspect of language (e.g. dictionaries of idioms and proverbs) have been compiled. Collocation dictionaries are a specialized dictionary aimed at serving learners' encoding purposes, and are addressed at learners at upper intermediate to advanced level and translators (Bogaards, 2003; Nuccorini, 2003). The Oxford Collocation Dictionary is based on a corpus of 100 million words. According to Nuccorini (2003, p. 378) it is more pedagogically-oriented than others (e.g. The BBI Dictionary of English Word Combinations, Selected English Collocations, English Adverbial Collocations). It allows language learners to freely access its online resource to seek help with collocations.

Collocations have been found to be troublesome to L2 learners from different language backgrounds, e.g. German (Bahns and Eldaw, 1993), Thai (Phoocharoensil, 2012), Japanese (Koya, 2003), and Taiwanese (Huang, 2001) as well as at different language levels (Nesselhauf, 2003; Laufer and Waldman, 2011). Previous studies using different tasks such as cloze and translation tasks (Bahns and Eldaw, 1993) or essays and reports (Granger, 1998; Nesselhauf, 2003; Laufer and Waldman, 2011) all suggest that collocations are responsible for 36% to 56% of all errors. 'The difficulties for language learners are not to understand what *weak tea* is but to actively produce *weak tea* and not *feeble tea* or *light tea*' (Herbst, 2010, p. 226). In the same vein, Laufer and Waldman (2011) point out that learners' productive knowledge of collocations is typically much worse than their receptive knowledge.

Vietnamese students are not an exception. My teaching of English major students at University of Social Sciences and Humanities has given me the opportunity to closely observe their learning process and the challenges they have to face. Reflecting on my 10-plus years in language teaching, with nearly half of that time focusing on teaching writing skills, I have seen that students are still struggling with

improving their writing. One thing easily seen in their writing was that most texts they produce are well-organized, cohesive, and with few grammatical mistakes. Careful scrutiny of some papers that received outstanding marks showed that it was language use that differentiated them from others. This suggests that for most papers, students have problems with collocational use, and it is clear that lack of knowledge of collocation greatly impedes their improvement in academic writing.

Working closely with my daughter to help her with Vietnamese homework, I recognized that the phenomenon of collocation was much the same in Vietnamese. That is, instead of combining the word *đen* (*black*) with words like *mèo* (*cat*), *ngựa* (*horse*), and *chó* (*dog*), in Vietnamese they must be accompanied by different words such as *mun* (*mèo mun: black cat*), *ô* (*ngựa ô: dark horse*), and *mực* (*chó mực: black dog*). These word combinations are clearly constrained by the conventions of the language. My daughter's question of why we do not use *đen* (*black*) in combination with all these words but *mun*, *ô*, *mực* has intrigued me greatly. She was temporarily satisfied with my answer that is what people say and learning a language means you have to learn how people of that language express ideas. Yet it is still a big question clinging to me, and if that is true of a language, then how can we, language learners of English, deal with it?

Clearly, the Vietnamese exercise that my daughter had to do on that day closely coincided with what was considered troublesome to my students. My daughter will have few difficulties with using them later since she has learned Vietnamese by immersion in the speech community from the start, and the amount of engagement with the language is large enough for her to build up her language base and thus know how to use them correctly. To my students and myself as a language learner, however, it is a matter of concern. If students miscollocate by combining words based only on syntactic features and semantic meanings of individual words regardless of conventional word combinations of the English language, they can still communicate without much interference. However, moving towards a native-like language goal, it is still a big gap to bridge. For students at around upper intermediate to advanced level, such as my students, being equipped with a tool or a strategy to

improve their collocation competence independently is much more important than teaching, since at this stage students are expected to broaden their store of vocabulary themselves.

Although numerous studies investigating the use of dictionaries and their support in learners' collocation use in writing have been carried out (Benson, 1989b; Jacobs, 1989; Laufer 2010), relatively little attention has been paid to the effects of specialized dictionary use on learners' collocation. This study, therefore, attempts to examine the effects of the Oxford Online Collocation Dictionary (OOCD) (see Appendix 1) on the collocation use of advanced language learners. Chomsky (2014) distinguishes competence and performance, stressing that the former refers to learners' ability or knowledge of rules while the latter indicates the actual use of language in concrete situations. In other words, competence involves 'knowing' the language whereas performance involves 'producing' the language. The primary aim of this study was to look at whether or not the dictionary helps learners use collocation correctly in their L2 writing rather than their actual knowledge of collocations, so their collocation performance/use was investigated. To achieve this goal, a qualitative study has been conducted. The study examines collocation use in learners' written texts without and with the support of this dictionary. Questionnaires, observation sheets and interviews were used in this study with an expectation that the study would bring about an in-depth understanding of learners' use of the dictionary and their evaluation of it as a supporting tool.

1.2   Context of the study

This section gives a brief description of the context of the study, which includes the organization of the Vietnamese education system, how English is taught, admission and the curriculum for English major students, and how dictionaries are used for language learning in Viet Nam.

1.2.1   Organization of the Vietnamese education system

The education system in Viet Nam is under the control of the Ministry of Education and Training (MOET). In the last two decades, the education system has witnessed constant changes, big and small, of the admission policy for higher education as well

as changes to meet the call for the reduction of the education curriculum, which is notorious for its bulky program and lack of practicality. There are five levels of education beginning with kindergarten for children from age three to six, to tertiary level for students aged eighteen and over. Children reaching the age of six attend primary education, which comprises five forms (Form 1-5). Education at this level is free for all children and focuses mainly on core subjects such as mathematics and literacy. English is only taught at some national standard schools and is optional. Children can approach English at this early age but with just a few periods per week. This subject is not on the standard curriculum and therefore is not free of charge.

After completing primary level, children move to secondary level, from Forms 6 to 9. At this level children start to learn a wider range of subjects, including history, geography, chemistry, physics, and English. English at this level receives greater attention with nine periods per week from Forms 6 to 8, and six periods per week in Form 9. Following this is high school level (Forms 10-12), which prepares students for tertiary level. English is studied for nine periods per week, making a total of 315 periods a year. At the end of this level, students sit a graduation exam the scores of which are used for university admission. Depending on the major that students wish to take, the scores of three subjects are considered. For instance, if students wish to apply for English teaching, their English, maths and literature scores are determining factors.

1.2.2 English in Viet Nam

English is used as a foreign language in Viet Nam. With the trend of global integration, especially after the economic Renovation and open-door policy in 1986, use of English started to grow rapidly. Although English is officially and compulsorily taught from secondary schools upwards, the mushrooming of English language centres for children at early ages shows that people are increasingly aware of its role. At secondary and high school levels, most English teachers attach great importance to materials and developing learners' reading skills. Teaching textbooks are designed and compiled by the Ministry of Education and Training. Grammar-translation methods, which allow learners limited roles and emphasize memorization and repetition, are widely used (Hoang, 2010). Learners at this level are heavily

reliant on their teachers for building their store of vocabulary. They tend not to be independent learners. In my experience, teachers are always expected to provide the meaning of every new word and their pronunciation; they therefore play a central role in the classroom. A reality that can be easily observed is that an important aspect of vocabulary teaching - teaching learners how to use vocabulary - is largely ignored. Consequently, the majority of learners' vocabulary is passive rather than active knowledge. They thus fail to recall and use vocabulary appropriately to make sentences.

At tertiary level, English is introduced as either a discipline or a subject. Students choosing English as a discipline get a Bachelor (BA) or Masters (MA) degree in English. Students choosing other disciplines than English also learn English as a compulsory subject, but it only accounts for 10% of the total credit hours of an undergraduate degree. Unlike at lower levels of education where the content of English teaching is designed and strongly observed by the Ministry of Education and Training, at tertiary level, institutions decide what to teach. Hence the content of English teaching depends on the purpose of the institution that offers the program (Banh, 2004). For students majoring in English, their English curriculum is much more intensive. New teaching methods which are more learner-centred, such as communicative approaches, are applied, but the vocabulary teaching method does not differ much from that of lower levels (Le, 2011).

1.2.3   University of Social Sciences and Humanities

The University of Social Sciences and Humanities (USSH) is currently one of the member universities of Viet Nam National University, Ho Chi Minh. Other member universities include Universities of Natural Sciences, University of Technology, University of Computer Science, University of Economics and Law, and International University. The USSH was established in 1955 and at present has 28 faculties and departments in total. English Linguistics and Literature, where I worked and conducted the study, is the faculty with the highest number of students on different training programs taken in each year, around 800 students.

The University has two campuses. The main campus is located in the city centre, and the second campus is at Linh Trung district, outside the city. The main campus is mainly used for teaching students at postgraduate level, foreign students, students of high quality training programs, and undergraduate students in their last year. First- to third-year students are assigned to study at the outer city campus. All the classrooms in both campuses are equipped with necessary teaching facilities. Internet is available at both campuses, but is not always strong enough to get access easily, especially at the campus outside the city. I had to bear this in mind when conducting the research since it involved students accessing an online dictionary.

1.2.4 Admission and English curriculum for English major students

Students majoring in English at universities in Viet Nam in general and the University of Social Sciences and Humanities, where I conducted the study, in particular, have to get, besides some basic subjects in Vietnamese, at least 140 credits of both obligatory and optional modules in English in eight semesters. To be admitted to the University, their graduation scores for the three subjects - mathematics, Vietnamese literature, and English - have to be above the threshold score set by the Ministry of Education and Training. However, this does not guarantee them a place in the university since universities will only take in applicants with the highest scores top down. The entrance score, hence, is not fixed every year, but is around 22. They all have been learning English for seven years in secondary and high schools, with language modules mostly organized around grammar syllabi with some reading and sentence-writing practice.

The eight-semester curriculum is split into three stages: the foundation stage, the intermediate stage and the specialized stage. The foundation stage comprises general knowledge and professional knowledge. Students take some modules in Vietnamese and are expected to earn 47 credits. Only general knowledge courses are taught at this stage. Professional knowledge, which comprises 98 credits, is taught from the first to the last year at the university. I carried out this research with the Academic writing module of the intermediate stage. In the first three semesters students were to learn basic language skills using *New Interaction* and *Mosaic* textbooks (Blass and Pike-Baky, 2007), which were designed with integrated listening, speaking, reading

and writing skills. Vocabulary building with new vocabulary compiled around themes occupies a separate section, or if not, as in *Mosaic*, they are highlighted in boldface. This means that vocabulary teaching is to be treated as an important part of, and hence has its own time share in, teachers' lesson plans. In contrast to these textbooks, *Writing Academic English* (Oshima and Hogue, 2000), used in the academic writing module, has no separate section on vocabulary. This does not mean that vocabulary is not taught, but at this level, around upper intermediate to advanced level, learners are expected to expand their store of vocabulary by themselves. My observation of students at this level is that they tend not to have much difficulty with grammatical structure, coherence, and essay organization in academic writing. In general, however, they tend to have difficulty in using collocations accurately in essay writing.

1.2.5   Dictionaries for language learning

There have been few studies on dictionary use and Vietnamese students' dictionary use skills so far. From my own experience as a language learner as well as a language teacher, I have observed that the roles of dictionaries in language learning are largely ignored by both language learners and educators. This is most likely because learners are not aware of the potential of dictionaries as a source of information for language learning. In most of language programs in schools, language educators have paid very little, if any, attention to giving instructions for dictionary use. Lack of encouragement from teachers has led to students overlooking dictionary use.

1.3   Structure of the thesis

Chapter 2 deals with literature that theoretically scaffolds the study. It starts with a review of how linguists define and classify collocations, from which the definition, identification and categorization of collocations in this study are drawn. Previous studies on collocational errors and causes of errors are also reviewed. This chapter then looks at issues surrounding dictionaries, comprising the role of dictionaries and vocabulary acquisition, difficulties and problems in dictionary use for production, and a comparison of the structure of general and collocation dictionaries. The research questions will be then introduced. This chapter ends with a review of literature related to methodological choices.

Chapter 3 shifts its focus to methodology and research design, addressing the decisions on research design and instruments for this study. This chapter then presents the pilot study, how it was carried out and its outcomes. It also includes a description of the process of data collection, which comprises two phases, before and after the intervention. This chapter ends with a discussion of ethical considerations and an examination of validity.

Chapter 4 provides detailed analysis of the written texts collected from the two phases. The analysis of questionnaires, recording sheets and interviews are presented in turn. Chapter 5 presents findings of the study obtained from written texts and both quantitative data (questionnaires and recording sheets) and qualitative data (recording sheets and interviews) to address the research questions. Chapter 6 is devoted to discussion of these results in relation to the research questions. The last chapter, Chapter 7, first covers a summary of the main findings from all the datasets and the contributions to knowledge that this research provides. Implications for practice, limitations, orientations for future research, and research reflections will be presented in turn in this chapter.

Chapter 2 : LITERATURE REVIEW

2.1   Introduction

In the previous chapter, I introduced issues concerning L2 learners' collocation use and the appearance of the collocation dictionary, which gave me the initial motivation to carry out this research. In order to investigate the effects of the Oxford Online Collocation Dictionary on learners' collocation performance in academic writing and learners' evaluation of the use of the dictionary as a supporting tool, in this chapter I will review the following areas of literature.

The first section of this chapter, section 2.2, begins by investigating what collocation is and explaining some other linguistic terms which shade into the concept of collocation. How collocation is defined in this study is then discussed. Next, I look at how collocations are classified in other studies, and this is the basis on which the classification of collocations in this study is developed.

Section 2.3 examines grammatical units from which combinations of some grammatical patterns are to be extracted. How grammatical units such as Clauses, Noun Phrases, Adjective Phrases, and sentence elements such as Subjects, Verbs, Objects, Complements, and Adjuncts are identified is discussed.

Sections 2.4 and 2.5 provide a review of collocational errors and reasons why collocations pose difficulties to language learners according to previous studies. An understanding of these issues is believed to help me in the process of introducing the dictionary to learners in the data collection process.

Section 2.6 shifts its focus to the roles of dictionaries in language learning – an aspect that potentially links closely to learners' collocation improvement in writing. It firstly reviews the roles of dictionaries in vocabulary acquisition, of which

collocation acquisition is assumed to be a part. An overview of difficulties of using dictionaries for production is then provided, followed by a description as well as a comparison of the structure of collocation dictionaries to that of others. On the basis of the acknowledgement of the difficulties when using dictionaries for production and the difference in structure between general and collocation dictionaries, I could give learners clear and adequate instructions on how to use the collocation dictionary effectively to support them in their writing. The discussion of this section led me to my research questions, introduced in section 2.7, focusing on the use of the online collocation dictionary as a writing aid and its impact on learners' collocation competence in L2 writing.

Section 2.8 presents a review of and an argument for the choice of research approach as well as instruments. This is the methodological foundation for the research design and methods chosen to generate the data.

2.2   Collocations

2.2.1   The concept of collocation

Early use of the term '*collocation*' was seen in Firth (1935) in his lexical studies to refer to a level of meaning which is syntagmatic-based and is not related to a conceptual approach to the meaning of words. Firth explains collocation thus: 'You shall know a word by the company it keeps' (1957, p. 182). 'One of the meanings of *night* is its collocability with *dark*, and of *dark*, of course, collocation with *night*' (1957, p. 196). However, the simplicity of the quotation is misleading as the concept of collocation is complex. Indeed, Nesselhauf states that 'word combinations are not in fact clearly delimitable and any attempt to do so involves both theoretical and practical problems' (2003 p. 224). In what follows, I will present a review of some significant approaches to collocation that usefully informed my own research.

Collocation is understood slightly differently depending whether it is viewed as a statistical or phraseological phenomenon. Sinclair (1991) sees collocation as recurrent patterns across large text collections, which implies that only combinations

of items that appear with some defined level of frequency can be treated as collocations. This statistically-based approach was followed by Stubbs (2002b, p. 29) and Clear (1993) who argued that the measurement of significant co-occurrence can be taken by comparing the observed frequency of co-occurring words in a combination against their expected co-occurrence (2004, p. 71). In Sinclair's definition, elements of a collocation can be of any part of speech while Stubbs (2002b, p. 24) only considers a combination of two or more lexical words to be a collocation. Though using the word 'word' in the definition, both Sinclair and Stubbs often use it to mean lemma. *Argue heatedly*, *heated argument*, and *the heat of the argument* are all considered instances of the same collocation (Stubbs, 2002b, p. 30). This means that collocation is the relationship between semantic units of the same lemmas.

For those scholars following a statistical approach, the syntactic relation between elements of a combination does not play a role in determining or classifying types of collocations. Instead, to decide whether or not a combination is a collocation, they consider whether those elements of a combination are within the '*span*', 'the number of word forms before and/or after the node', the element being considered (Stubbs, 2002b, p. 29). According to Sinclair (1991) and Stubbs (2002b), which of the two words in a string plays the role of the node is not fixed; the decision, in fact, depends on the focus of the study. Although there is no total agreement on the span, Sinclair's (1974) range of four either to the left or the right of the node is widely adopted when calculating frequency. Beyond this span, researchers do not usually find statistically significant relationships.

In terms of meaning, some statistically-based scholars claim that there are always 'semantic relations between node and collocates' in a collocation (Stubbs, 2002a, p. 105; Sinclair, 2004; Partington, 2004). Collocates of a word form a semantic class often characterized in respect of meaning, negative or positive (Stubbs, 2002a). These scholars call some kinds of meaning arising from the combination of a node with its typical collocates *semantic prosody*. Stubbs (2002a) analyzed concordances

for *cause* and *provide* and came to the conclusion that the former often collocates with an 'unpleasant' semantic property, e.g. *problems*, *damage*, *disease*, etc. and the latter with positive collocates, e.g. *help*, *support*, *assistance*, etc. *Happen* characteristically appears together with 'something nasty that has happened or going to happen' (Sinclair, 2004, p. 33). As such, collocations can be understood to be combinations of meaningful units.

Scholars adopting a phraseological approach, in contrast, regard collocations as phraseological units, which are used to mean a type of word combination in a particular grammatical pattern. A syntactic relation of some kind between elements of a combination, according to them, is an essential requirement (Nesselhauf, 2005; Cowie, 1994; Bahns and Eldaw, 1993; Hausmann, 1989). Most discussions of collocation from this angle involve the distinction between this term with the two key terms *free combination* and *idiom* (Benson, 1989b). Nesselhauf (2005) uses a widely accepted criterion among phraseological-based scholars for collocation identification: '*arbitrary restriction on substitutability*'. Sharing a viewpoint with Cowie (1994) she states that this description of collocation helps differentiate restricted collocations from free combinations, of which the substitution of elements depends solely on their semantic properties. *Read a newspaper* and *reach a decision* are examples of free combination and restricted collocation respectively (Nesselhauf, 2003, p. 225). As she explains, *read* can be accompanied by any nouns with semantic properties of 'containing written language', whereas *decision* in *reach a decision* can only be substituted by nouns denoting 'aim', such as *conclusion*, *verdict*, *compromise*, or *goal*.

Based on this notion of substitutability, *face her anger, face a task, face a financial crisis* might not be regarded as collocations since, as Nesselhauf argues, the choice of objects following *face* 'seems unlimited as long as it refers to some kind of difficult or unpleasant situation' (2005, p. 26). Stubbs calls the meaning arising from common semantic features of frequently occurring collocates of a given node *semantic preference* (2002a, p. 225). However, if this rule is applied to *face* with the sense of

dealing with, it can be seen that *face* can only be replaced by *cope with*; hence, they should be treated as collocations following Nesselhauf's definition. Similar to Cowie, Nesselhauf suggests that restricted commutability should be limited to synonyms only since if one simply searches for verbs that can collocate with the noun *decision*, many verbs can be found: *reach a decision, come to a decision, arrive at a decision, postpone a decision, criticise a decision, explain a decision* etc. (Nesselhauf, 2005, p. 27). Among these verbs only *reach, come to,* and *arrive at* can be used interchangeably, so according to her, if *reach a decision* is the word combination being considered, it is called a collocation.

Although Nesselhauf (2005) gave a very clear interpretation of the criterion of restricted substitutability (also called commutability by Cowie (1981), Howarth (1998), and Aisenstadt (1981)), her process of identifying restricted collocations involves a very difficult series of choices and is not free from drawbacks. In particular, in her study, in order to delimit restricted collocations from free combinations, she investigated whether or not an element of a combination is used with a restricted sense in the combination. Verbs of the verb + noun combination are often chosen for this consideration, though, as she states, either of the two elements can be taken out for examination. The decision was made based on dictionary searches of Oxford Advanced Learners' Dictionary (OALD 2000 (8th Edition) and Collins COBUILD English Dictionary (CCED 1995 (4th Edition). If a word is used with a restricted sense, the combination is a collocation (Nesselhauf, 2003, p. 225). If it is not clear from the dictionary search whether the meaning of a word is restricted or not, she chooses three synonymous words, neither too common nor too uncommon, for a substitution test. This is not very persuasive since different conclusions might be drawn depending on the choice of the three synonyms. Take, for example, the combination *perform the ceremony*. According to Nesselhauf, there are no clear indications that the verb *perform* is restricted to a few nouns from the dictionary search (of the two dictionaries she suggested above). For such a case, we need to try out a substitution test of three synonymous nouns which, according to the definition in the dictionaries, should be combinable with the verb in question. I searched the Oxford Learner's Thesaurus and found a list of 22 synonyms of

ceremony: *rite, ritual, ceremonial, observance, service, sacrament, worship, mystery, office, celebration, performance, act, practice, order, custom, tradition, convention, institution, formality, procedure, usage, form*. All of these nouns can accompany the verb *perform* except for *observance, performance* and *custom*. The combination of *perform* with these nouns is not found in the BNC and is judged unacceptable from native speaker consultation. This means that the combination in question will be identified either as a restricted collocation or a free combination depending whether *observance, perform* and *custom* are chosen for the substitution test. Nesselhauf (2003, p. 233) admits that the procedure of collocation identification is complex, and for cases that are inconclusive she has to categorize them into a 'less certain' group (RC?), which lies between collocation of little restriction and free combination on the collocational scale.

Another important criterion often used by scholars in this tradition to differentiate collocations from idioms whose meanings are frozen and do not reflect meanings of the components, e.g. *kick the bucket* (Philip, 2011; Cowie and Howarth, 1996; Benson et al., 1986; Nation, 2001), is transparency. This is interpreted slightly differently. A combination is semantically transparent, as described by Philip, when its meaning is 'clear from a compositional reading of its component words' (2011, p. 21). This view is shared by Cowie (1994), Mel'cuk (1998), and Aisenstadt (1981). However, it is required that a combination contain at least one element carrying one non-literal sense and one literal sense to be considered a collocation (Cowie, 1994). *Perform a task* was taken as a collocation example in Cowie's work. According to him, *task* is used in its literal sense while *perform* is non-literal. Nevertheless, it is true that this way of identifying collocation is problematic in that it is often difficult to decide if a word is used with a literal or non-literal sense (Howarth, 1998). For Mel'cuk (1998), a combination will be transparent if it consists of one element chosen freely based on its meaning and the other chosen depending on this freely chosen element. This means that one element of the combination will carry the primary meaning that can be found in the dictionary. Though he does not address which or whichever of the two elements of a combination will be carrying the primary meaning, it seems that this element is fixed and coincides with what

lexicographers call the base. As in verb + noun combination (i.e. *face the task*), the noun is always the base (Benson, 1989b; Lea, 2007). The base, though not equal to the node, is also used to refer to the word being considered. It can be inferred from the definition of these authors based on both approaches that collocation is restricted to content words only.

In this study, it is important for me to know how corpus-based lexicographers understand the concept of collocation, since the research examines the impact of the dictionary on learners' collocation use. Lexicographers start from the word, and therefore are more likely to take a phraseological approach. For them collocations must be restricted in some way but are transparent in meaning (Lea and Runcie, 2002, p. 819). The theory of collocability, the combinatory potential of words, is of great importance to them. Benson states that collocation should not be defined just as 'recurrent word combination' but as 'arbitrary recurrent word combinations' (1989b, p. 3). Examples given to illustrate the arbitrary nature of collocation are people saying *make an effort* but not *make an exertion*, *a running commentary* but not *a running discussion*, *warmest greetings* but not *hot greetings* (Benson, 1989b, p. 4). This understanding of 'collocation' is shared by Lea (2007), who stresses that the focus of the dictionary is on the 'medium-strength' collocations, which are elsewhere called restricted collocations (Cowie and Howarth, 1996; Nesselhauf, 2005; Mel'cuk, 1998). Frequency is used but only as a 'blunt instrument' to decide if a combination is typical and is worthy of including in a collocation dictionary (Lea and Runcie, 2002, p. 828).

Benson (1985) argues that collocation dictionaries should only provide idiosyncratic combinations (combinations with arbitrary constraints), which are unpredictable to learners. However, the dividing line between collocations and free combinations is not clear (Hottsrnonn, 1991). The decision of which collocations to be included in the dictionary, according to Lea (2007), is quite challenging. It seems hard for lexicographers to decide accurately and consistently which collocations are predictable and which are not, partly because they are from different linguistic and

cultural backgrounds to learners and partly because a collocation may be predictable to learners of a particular linguistic and cultural background but unpredictable to learners from others. *See a doctor* could be easily predicted by European learners, as Benson (1989a) claims, but not to Japanese learners (Nakamoto, 1992). It is probably not predictable to Vietnamese learners, either. Based on knowledge of grammar and vocabulary alone, it is highly likely that Vietnamese learners will construct *meet a doctor* to express the idea.

The most frequent collocations in the language fall into the less restricted or 'fairly open' categories (Lea and Runcie, 2002). Lea and Runcie (2002) compare *great importance* and *wax lyrical* and claim that though *great importance* is not as strong a collocation as *wax lyrical*, it is much more frequent and is probably more useful to learners. Given the discussion about collocations from lexicographers' standpoint, collocations in this study need to include combinations that are less restricted. As such, in this study frequency should be prioritized rather than be just a 'blunt instrument'.

2.2.2 Other linguistic terms referring to collocational phenomenon

Besides the fact that collocation is defined and used differently, scholars use different terms to describe this phenomenon. Hence, this section examines some linguistic concepts that are closely related to collocational phenomena such as *formulaic sequences* (Wray, 2002), *phrasal lexemes* (Moon, 1998) *set phrases* or *phrasemes* (Mel'cuk, 1998), and *lexical phrases* (Nattinger and DeCarrico, 1992). This understanding of closely-related terms has provided some insights into the differences among the terms and supported me in collocation recognition when analyzing learners' writing.

 Wray (2002) uses *formulaic sequences* as a general term for covering a comprehensive range of terms – as many as over fifty, including collocation – to describe aspects of formulaicity. The term 'formulaic' refers to words or strings of

words 'processed without recourse to their lowest level of composition' (Wray, 2002, p. 4). Therefore, *formulaic sequence* is used to cover the range of sequences of prefabricated words (prefabs) which are stored and retrieved whole at the time of use. In her discussion of characteristics which are typical of formulaic sequences, although not all exhibit in every lexical unit, Wray (2002) touches on some that are believed to coincide with important features of collocation in this research, as follows.

Firstly, though 'stored in mind as holistic unit', formulaic sequences need not be retrieved in an 'all-or-nothing manner' (Schmitt and Carter, 2004, p. 4). This can also be regarded as an important distinctive feature between collocations and idioms on the scale of flexibility. Idioms, the meanings of which are semantically opaque (e.g. *kick the bucket*, *spill the beans*), are at one end, and collocations are at the other end of the scale, where the base or node (Nattinger and DeCarrico, 1992; Sinclair, 1991), like *study* (as in *carry out a study*), can be combined with a restricted amount of collocators, including *do/make/conduct* (Nesselhauf, 2005, p. 18). This feature clearly overlaps with the feature about restricted variation in Cowie's definition of collocation. Another characteristic of formulaic sequences that exhibits in collocations is that they can have semantic prosody (Wray, 2002). This collocation-like characteristic is found not only between individual words but also between words with sets of words that share some common semantic features.

Besides the characteristics mentioned above, the assumption that formulaic sequences can be identified on the basis of their frequency illustrates that they are closely related to or are collocations themselves. This way of identifying formulaic strings sometimes presents difficulty since there are strings, such as *long live the King*, *all for one and one for all*, which are recognized as formulaic sequences by native speakers, but are not found frequently in corpora (Wray, 2000, p. 466). Although a formulaic sequence exhibits many important features of collocations, it is not considered an equivalent term because it covers a lot more sequences such as idioms, fixed expressions, phrasal verbs, clichés and so on.

Mel'cuk (1998) uses the terms *set phrases* or *phrasemes* as superordinates, of which collocation is a subtype. A significant property of phrasemes, which Mel'cuk uses to refer to predominant lexical units, is their 'non-compositionality' (1998). It can be understood that these phrasemes are fixed and/or ready-made for retrieving rather than being constructed from individual words. At this point, Mel'cuk's phrasemes also cover a wide range of linguistic terms, and coincide with *formulaic sequences*. Mel'cuk assigns collocations, idioms and quasi-idioms to the semantic phrasemes group whereas ready-made expressions for greetings, typical phrases used in letters, conversation formulae are accommodated to pragmatic phrasemes, which, as he describes, are pragmatically appropriate, though semantically and syntactically compositional. Calling collocations '*semi-phrasemes'* implies that one component of a collocational unit is freely chosen based on the meaning that the speaker wishes to convey while the choice of the other is constrained by the convention of the language.

Though Nattinger and DeCarrico's *lexical phrases* shade into the concept of collocations, they are distinct from collocations. In terms of form, according to Nattinger and DeCarrico (1992), lexical phrases are also collocations since they are described as multi-word chunks running on a continuum from fixed phrases (e.g *of course, for example*) to slot-and-filler frames (e.g. *the__er, the ___er*, *a ___ ago*) and can be either *canonical,* conforming to grammatical rule (e.g. *what on earth*, *at any rate*, etc.) or *non-canonical,* not conforming to grammatical rule (e.g. *by and large*, *as it were*, etc.). Collocations occupy a position somewhere near an end of the continuum, where syntagmatic substitution is possible, but restricted. Scholars from different approaches define collocation slightly differently; hence, for those adopting the frequency approach, Nattinger and Decarrico included, these strings are collocations, but for others adopting the phraseological approach, they are not. In terms of function, as Nattinger and Decarrico describe, lexical phrases perform a pragmatic function while collocations are semantic. Lexical phrases like *how do you do?* or *the__er, the ___er*, as Nattinger and DeCarrico (1992) confirm, perform

pragmatic functions, namely greetings and comparison relationships, respectively. In this way, it can be inferred that a collocation can be a lexical phrase, but a lexical phrase is not necessarily a collocation. If these scholars differentiate these terms based on their functions, lexical phrase should not be regarded as an equivalent term with collocation.

Moon's *phrasal lexeme* also covers a wide range of 'holistic units of two or more words', but embraces fixed phrases like 'frozen collocations, grammatically ill-formed collocations, proverbs, routine formulae, sayings, and similes' (1998, p. 2). In order to identify whether or not potential fixed expressions and idioms (which she terms FEIs are holistic units, she uses three criteria: lexico-grammar, pragmatics, and semantics, and calls strings with lexico-grammatical problems anomalous collocations, of which '*ill-formed*, *cranberry*, *defective*, *and phraseological collocations'* are subtypes. Moon describes strings belonging to ill-formed collocation (e.g. *by and large*, *of course*, *stay put*, etc.) as formally "ill". Conversely, based on an argument of Nattinger and DeCarrico (1992) on the grounds of pragmatic function, these strings, '*by and large'* and '*of course',* might be called lexical phrases. Examples that Moon provides as an illustration of the cranberry (e.g. *to and fro*, *in retrospect*, and *kith and kin*), defective (e.g., *at least*, *in effect*, and *beg the question*) or phraseological collocation (e.g., *in action*, *on show*, *to a --- degree*) are not considered to be collocations and are set outside the scope of this study because they only exemplify 'defective' collocations. What Moon calls phraseological collocations are elsewhere called grammatical collocations (i.e. Benson et al., 1986).

2.2.3   Identification of collocations in the study

From the above review of collocation definitions and linguistic terms referring to collocational phenomena, collocation in this study is identified using both frequency and phraseological approaches. I am only interested in collocations of some grammatical patterns, so in this study they must be of particular syntactic relations. This also means that the span of four, which is widely used among statistically-based

scholars, is not applied in the process of extracting combinations. Since this study is an examination of collocation dictionary use, the term 'base', as used by lexicographers, is used instead of 'node'. An exploration of the base-collocate relation is discussed in a later section (Section 2.6.3). Elements of collocations are considered to be lemmas rather than word forms or lexical items. This study only focuses on combinations of lexical/content words; hence, for the discussion now and throughout the study collocation is used to refer to only lexical/content word combinations, also called lexical collocations (as opposite to grammatical collocations, which will be discussed in detail in section 2.2.5). Rather than choosing a particular definition, collocation in this study was identified based on a number of criteria that the above approaches consider to be characteristics of the phenomenon.

Firstly, collocation is frequent co-occurrence of words. Corpus-based scholars agree that collocation is when words co-occur more frequently than would be expected by chance (Moon, 1998; Hunston, 2002; Clear, 1993; McIntosh and Halliday, 1969; Stubbs, 2002b). To determine this, they need to calculate how many times a word pair is expected to co-occur in a corpus of a certain size by chance. Take, for example, the co-occurrence of *strong* and *tea*. If f_x is the number of occurrences of the first word *strong* and f_y the second word *tea*, then in the British National Corpus (BNC) of 96,134,547 words, f_x equals 19,265, and f_y equals 8,357. The possibility of *strong* and *tea* randomly co-occurring in the BNC will respectively be:

$$Strong = f_x \div 96,134,547 = 19,265 \div 96,134,547 = 0.0002$$

$$Tea = f_y \div 96,134,547 = 8,357 \div 96,134,547 = 0.00008$$

The possibility the two words co-occur will be:

$$Strong\ tea\ (f_{xy}) = 0.0002 \times 0.00008 \times 96,134,547 = 1.54\ times.$$

In fact, the word pair occurs, within the span of 5, 74 times in the BNC. In 28 out of that number of co-occurring times, they stand adjacent to each other, much greater than would be expected by chance. Although most corpus-based scholars agree on greater-than-chance frequency of a word pair, there is a debate among them about the threshold of significant co-occurrence. Moon (1998, p. 57) set the significance

threshold of co-occurrence at five in her study of the Oxford Hector Pilot Corpus (OHPC); Clear (1993, p. 277) considers three occurrences as the minimum requirement while some others use different statistics. Hunston (2002, p. 71) writes that a Mutual Information score (MI score) of three upwards can be taken as significant.

'The MI score is the Observed divided by the Expected, converted to a base-2 logarithm' (with the Observed referring to instances of the co-occurring word in a designated span and the Expected to instances of the co-occurring word in a corpus as a whole) (Hunston 2002 p.70). So, the MI score of *strong tea* will be:

$$\text{MI} = \log_2 \frac{Observed\ fxy}{Expected\ fxy} = \log_2 \frac{74}{1.54} = 5.69$$

This MI score, as Hunston (2002) explains, 'indicates the strength of a collocation' and is in fact a measurement of 'the amount of non-randomness present when two words co-occur' (Hunston, 2002, p. 71). It is not, however, a very useful statistic for phraseologically-based linguists since it emphasizes rare words (Adam and Iztok, 2012; Baker, 2006; Gablasova et al., 2016). Nor is it always 'a reliable indication of meaningful association' (Hunston, 2002, p. 72). As Gablasova et al. (2016, p. 10) put it, the 'MI score is not constructed as a reliable scale for coherence or semantic unity of word combinations.' Hunston also illustrated that the number of times *baleful* and *gaze* co-occur is only 6 in the BNC but they still obtain a high MI score since in its few occurrences, *baleful* often accompanies *gaze*. The combination of the misspelling *suprising* with *hardly* has a high MI score (8.0), a figure which is higher than that of the correct combination *hardly surprising* (7.8) for this same reason (Hunston, 2002).

There are some association measures, e.g. MI3, log likelihood, Dice coefficient, but they were all evaluated as not useful either since functional words often dominate the list (Adam and Iztok, 2012; Rychlý, 2008; Baker, 2006). As well as the MI score, the T-score is a frequently used measure in recent research, but the choice of either of

these two measures is 'somewhat arbitrary' (González Fernández and Schmitt, 2015, p. 96). The T-score is calculated as follows (Gablasova et al., 2016):

$$\text{T-score} = \frac{Observed - Expected}{\sqrt{Observed}} = \frac{74 - 1.54}{\sqrt{74}} = 8.43$$

The T-score does not operate on a standardized scale (Hunston, 2002). In other words, T-scores are 'directly dependent on the corpus size' (Gablasova et al., 2016, p. 8), so it is hard to decide a cutoff point of the values of the results.

In contrast to the T-score, the Log Dice score operates on a standardized scale with a fixed value (Gablasova et al., 2016, p. 10; Rychlý, 2008). Like the MI score, it gives prominence to exclusive combinations but does not highlight rare combinations, and this makes it preferable to the MI score. As Gablasova et al. (2016, p. 11) explain, it does not 'invoke the potentially problematic shake-the-box, random distribution model of language because it does not include the expected frequency in its equation'. The formula for calculating the Log Dice score is, in fact, an improved version of the Dice score, the results of which are very small numbers (Rychlý, 2008, p. 6). The formula to work out the Log Dice score is as follows:

$$\text{Log Dice} = 14 + \log_2 D = 14 + \log_2 \frac{2fxy}{fx + fy}$$

I calculated the Log Dice score of *strong tea* in the BNC and found:

$$\textit{Strong tea} = 14 + \log_2 \frac{2 \times 74}{19,265 + 8,357} = 6.46$$

The statistic is shown to provide good results for collocation candidates (Gablasova et al., 2016) and Rychlý even calls it 'a lexicographer-friendly association score' (2008, p. 6). Different association measures prioritize different aspects and for the purpose of this study the Log Dice score was used to decide if a combination is a strong collocation. A Log Dice score of 4 or higher is taken as significant. This score was set for the study to include restricted collocations, from very strong to fairly weak. A test of the Log Dice threshold from the BNC shows that candidate collocates of the noun *chance* include all the verbs appearing in combination with that noun in some dictionaries:

- Candidate collocates found in the BNC (Log Dice of 4.0): *give, stand, miss, offer, take, have, improve, reduce, increase, seize, jump, create, lose, deny, squander, maximize, see, wait, ruin, fancy, boost, enhance, realize, win, waste, rate, leave, allow, lessen, turn, deserve, affect, jeopardize*

- Cambridge Advanced Learners' Dictionary: *get, have, give, miss, stand, improve*

- Oxford Advanced Learners' Dictionary: *have, wait, give, get, take, stand*

- Collins COBUILD English Dictionary: *have, get, stand, take*

- Macmillan Dictionary: *get, have, give, offer, provide, jump, seize*

To determine whether this Log Dice threshold has been achieved, I used the British National Corpus. Regardless of being old, built up during the period of 1980s-1993, I used it since I am not concerned with very new language uses, but with the frequent and typical core of the language. Corpus-based dictionaries, though a possible option, were not used in this study since they do not cover every possible and correct combination, and more importantly they are less likely to contain as many combinations as a corpus.

The criterion of substitutability, distinguishing collocations from free combinations, was not used in this study on the grounds that if we consider word combination as forming a continuum from very weak to very strong, then only idioms, which are specifically called pure idioms by Cowie (1988), Howarth (1998), or Moon (1998), are totally fixed (Cowie, 1988) and need to be learned as big words rather than by combining words together (Weinreich, 1969, p. 26). Hence collocation identification in this study only involved distinguishing collocations from idioms. Only the criterion of transparency, distinguishing collocations from idioms, was applied. This is taken to mean that the meaning of the combination as a whole is clear from the meanings of individual words, regardless of whether or not the base of the combination carries the literal meaning. Nevertheless, this criterion was used as an additional criterion only. In particular, if the base of a combination does not carry a primary meaning that can be found in the dictionary, the transparency of the meaning

of the combination as a whole will be considered. Take, for example, the case of *take steps* in the sentence *Where reasonably practicable, the authority must take steps to reunite the child and his family*. *Steps* here has a figurative sense but the combination as a whole is transparent, meaning to take a measure or action in order to deal with or achieve a particular thing, and therefore is considered a collocation in this study.

In this study, in order to avoid confusion, I am not using the term 'restricted collocation' since it is often used to refer to combinations of restricted commutability by Nesselhauf (2005) or Cowie (1994). Instead, combinations that met the above frequency threshold and were identified not idioms were called strong collocations. 'Casual combination' was used to refer to combinations with a frequency of co-occurrence lower than the threshold set (Log Dice < 4.0).

2.2.4 The degree of acceptability of the combinations

The process of investigating students' writing involves not only distinguishing strong collocations from other combinations, casual combinations and idioms, but also the identification of collocational errors. This means that a combination must be examined for its acceptability before a decision on its status can be made. To decide the degree of acceptability of combinations, Nesselhauf (2005) used the BNC to check for their occurrence. She then used native speakers as a back-up strategy because judgement made from corpus searches alone might result in an incorrect conclusion. She argues that though not present in the corpus, many combinations are acceptable to native speakers, such as *long live the King*. In addition, there will be cases in which the occurrence of the combination in the corpus is not enough for statements about their conventionality to be made, but are again acceptable to native speakers. In her study, a combination is considered acceptable if it is either found in a dictionary or occurs at least 5 times in the BNC. Combinations that do not meet the threshold are then judged by native speakers based on a three-point scale: acceptable, unacceptable, and questionable.

The process of determining acceptability of combinations from Nesselhauf's study was adopted in this study. However, as I have just discussed above, I used a

statistical figure (the Log Dice score) of frequency approach to identify strong collocations, rather than searching for whether a word is used with a restricted sense in a combination from dictionaries as she did. Native English speaker co-raters were also used for further judgements on acceptability in this study; detailed discussions will be presented in the next chapter.

2.2.5 Classification of collocations

Linguists from different perspectives have different views of, as well as ways of classifying, collocations. In order to categorize collocations, I needed to have an understanding of how they view and group them.

Being one of the linguists adopting the frequency-based approach, Sinclair (1991) regards collocations as a relationship between lexemes which are abstract classes of words-forms. Strings such as *a strong argument*, *he argued strongly*, *the strength of the argument*, *his argument was strengthened* reflect 'a high degree of morphological and syntactic position change' (Philip, 2011, p. 24) and are all considered collocations. The decision on whether these strings belong to *upward*, *neutral* or *downward collocation* is made based on the frequency of the *node* and its *collocate*. In particular, if, in comparison to its collocate, the node is a more frequently occurring word in a corpus, that combination will be a downward collocation; if this relation is in the opposite direction, that combination will be upward (Sinclair, 1991).

Moon (1998, p. 27) generally divides collocations into three kinds, which, according to her, 'reflect qualitatively different kinds of phenomenon'. The first kind, semantically-based, of which the co-occurrence of *toys* and *children*, *jam* and *strawberry* are examples, is said to be motivated (Hunston, 2002, p. 68). The second kind is syntactic and is specifically described as a combination of a verb, adjective, or noun with a preposition. Take, for example, *a/an __ of, too __ to, many __ of* (Moon, 1998, p. 27). The last kind is both lexico-grammatically and semantically constrained. *Face the truth/facts/problem* (Aisenstadt, 1981), *strong tea*, and *powerful car* (Halliday and Hasan, 1976, p. 73) are examples. This kind seems to be

unmotivated (Hunston, 2002) and is called restricted collocation (Aisenstadt, 1981; Cowie, Anthony Paul, 1981; Nesselhauf, 2005).

Mel'cuk (1998) breaks collocations down into four major groups:

1) collocations with support verbs such as '*to do a FAVOR*', '*to give a LOOK*', and '*to launch an APPEAL*';

2) collocations with intensifiers such as '*strong coffee*' and '*deeply moved*';

3) collocations such as '*black coffee*' and '*French window*';

4) collocations such as '*aquiline nose*' and '*rancid butter*'.

Collocations with support verbs are also called delexicalization by Stubbs (Stubbs, 2002b). According to Stubbs, nouns are the elements that carry the meanings of the whole combinations. *Take a look* is equal to *look*, and in cases like this, the verb is said to be delexicalized (Stubbs, 2002b, p. 32). As for type 2, though not being presented explicitly, it can be assumed that adjective and adverb respectively function as intensifiers in the combinations, and only some particular adjectives and adverbs can be used. So are the other types; the choice of the accompanying elements of *favor, coffee,* and *butter* are expected to be *do, strong*, and *rancid*, respectively. Though they are all included as collocations in this study, this way of classifying them will be ignored because it is not reasonable in that lexical collocation of noun + verb, such as *conditions prevail*, can hardly fit in any of these groups.

Many other researchers, however, divide collocations into two types: lexical and grammatical collocation (Aisenstadt, 1981; Benson et al., 1997; Hottsrnonn, 1991). Lexical collocation is used to refer to combinations of lexical elements, whereas grammatical collocation refers to combinations of a lexical and a grammatical element (mostly a preposition) (Nesselhauf, 2005). Though dividing collocations into two types, most of these linguists when investigating lexical collocations did not exclude other elements that are closely associated with them (e.g. prepositions in *take sth into account*, *cope with a problem*). In her study solely on verb + noun collocations, Nesselhauf (2005) claims that excluding these elements in advance

would mean that the researchers take a risk overlooking learners' problems in verb-noun combinations.

This study focuses on lexical collocations; hence only sub-classification of lexical collocations is discussed here. An attempt to explicitly sub-classify collocations on the basis of semantic characteristics of collocators is made by Cowie (1992). He divides verbs into groups of figurative meaning (*deliver a speech*), de-lexical meaning (*make a recommendation*) and technical meaning (*try a case*). However, in terms of 'internal variability' (commutability), Cowie and Howarth (1996, p. 83) divide collocations into four sub-types:

a. Invariable collocations like *break a journey, foot a bill*

b. Collocations with limited choice at one point like *give advice/a book/a chance*

c. Collocations with limited choice at two points like *get/take a lesson/a pill*

d. Overlapping collocations like *convey a point*, *communicate a view,*

As for type a), no substitution of either of the elements can be made for the collocation to maintain its specific meaning. The last type is the combination of 'apparent openness and restriction'. It can be seen that this classification and the semantically-based can only be done with lexical collocations of verb-noun.

Based on word classes of lexical elements, Hausmann (1989, p. 1010) divides collocations into six groups: adjective + noun (*heavy smoker*), noun + verb (*storm – rage*), noun + noun (*piece of advice*), adverb + adjective (*deeply disappointed*) verb + adverb (*severely criticize*) verb + (object) noun (*stand a chance*). Benson et al. (1997) makes the same sub-classification as Hausmann but includes combinations of nouns such as *soup spoon* under the noun + noun pattern.

2.2.6 Classification of collocations in this study

Given the discussion about collocation classification in the section above, I will classify collocations based on word classes. This is because the Oxford Online

Collocation Dictionary, the supporting tool that I was investigating, also classifies collocations as types of word combinations under noun entries (e.g. adjective + noun, quantifier + noun, verb + noun, noun + verb and so on), verb entries (e.g. adverb + verb) and adjective entries (e.g. adverb + adjective). Collocations are grouped into the following types:

1. V + N   *commit suicide, stand a chance*
2. N + V   *storm rage*
3. Adj + N   *heavy smoker*
4. Adv + V, V + Adv   *severely criticize, complain bitterly*
5. Adv + Adj   *deeply disappointed*
6. N + of + N   *piece of advice*
7. N + N   *soup spoon*

Informed by the definition of collocations in previous studies, strong collocations in this study would include other associated elements. This means that collocations are not restricted to just two lexical elements. Other elements such as the preposition *into* as in *take sth into account* or *take sth into consideration* would be extracted together with lexical elements as a whole for consideration.

From the above discussion of characteristics of strong collocations and the classification of strong collocations into the seven groups, in this present study I addressed strong collocations from the aspect of grammar. That is, the combinations of these seven patterns must be within grammatical units (phrases or clauses). In what follows, I am presenting a review of those grammatical units from which strong collocations of these patterns could be found.

2.3   Grammatical units

Clauses, finite and non-finite, were first reviewed and discussed since they are the grammatical units from which combinations of the V-N, N-V, and Adv-V pattern could be found. I then address the Noun phrase, of which the combinations Adj-N, N-N, N-of-N and sometimes N-V pattern may be part. Finally, I present a review of the Adjective phrase, from which combinations of Adv-Adj pattern can be extracted.

2.3.1 Clauses

At clause level, the distinction often made is between finite and non-finite clause. Their status, finite or non-finite, is specified based on the form of the verb, which, according to Downing and Locke (2006, p. 12) will signal either tense or modality in a finite clause and neither of them in a non-finite clause. (1) and (2) are examples of finite clauses of which the verbs are identified as finite since they express tense and modality, respectively, while in (3) the verb *hire* signals neither of them, and is therefore a non-finite verb of the non-finite clause:

(1) He *saw* the bottles. (past tense)

(2) He *will stay*. (modality)

(3) They want *to hire* a caravan.

Grammarians define and label elements that might occur in a clause similarly; that is, they consist of a subject, a verb, and other elements, which can be optional or obligatory depending on the verb, such as an object, a complement, or an adjunct (Downing and Locke, 2006; Carter and McCarthy, 2006; Coffin et al., 2013; Greenbaum, 1996; Burton-Roberts, 2016). However, the slight difference in the way they call or categorize some elements is worth noticing to avoid confusion at the analysis process. These differences will be pinpointed while I discuss a review of the identification of the elements of a clause.

Subjects and verbs

Most grammarians analyze sentences into two main constituents, Subject and Predicate (Downing and Locke, 2006; Aarts, 2013; Coffin et al., 2013; Burton-Roberts, 2016; Biber et al., 1999b; Greenbaum, 1996). The Predicate, whose function, as Aarts (2013, p. 9) suggests, is 'to specify what the Subject is engaged in doing' and 'is everything in the sentence except the Subject', always consists of a verb functioning as a Predicator. The Predicator determines what elements (Objects and Complements) can or must occur in the clause and is realized by either finite or non-finite verbs. In a discussion of ways to identify the Subject, Downing and Locke

(2006) and Aarts (2013) all assume that the Predicator must be known for the Subject to be specified. Therefore, before addressing how to identify the Subject, I needed to examine how the Predicator could be pinpointed.

From a semantic point of view, Predicators are the element that denotes an action or 'process' of the following kinds (Downing and Locke, 2006, p. 49):

- Material processes of 'doing' (e.g. *make, catch, go*);
- Mental processes of 'experiencing' with cognitive verbs of perception (e.g. *see, hear*) cognition (e.g. *know*), affectivity (e.g. *like*) and consideration (e.g. *hope*);
- Relational processes of 'being' (e.g. *be, belong*).

Semantically, the Subject can be identified as the constituent which tells us who does the action and who or what the clause is about (Downing and Locke, 2006; Aarts, 2013; Burton-Roberts, 2016). However, it is true, as Aarts (2013) argues, that though practical and useful, this semantic definition is problematic in cases where the Subject elements are meaningless (e.g. dummy *It,* unstressed *there*). Syntactically, Aarts (2013) proposes some tests to look at in identifying Subjects. However, some of them cannot tell whether or not a constituent is a Subject if not considered in combination with others. Take, for example, the first one which says Subjects are usually Noun Phrases. In reality, as Downing and Locke (2006) point out, besides Noun Phrases, Subjects can be Dummy *It*, unstressed *There*, a Prepositional Phrase, an Adverb Phrase, an Adjective Phrase, a Finite clause, Anticipatory *it*, and a non-finite clause. The second test which says the Subject is the first Noun Phrase we come across cannot tell if a constituent is a Subject, either. This can easily be seen in example (4) below:

4. Last night, *the teachers* were very drunk.

Last night is the first Noun Phrase but *the teachers* is the Subject instead. The other tests, position, concord and tag question, which are also proposed by Downing and Locke (2006), seem to be more effective and most usefully informed my study. In particular, the feature of position tells me that the Subject occurs before the finite

verb except in interrogative clauses, which helped me identify where the Subject could be found. The Subject determines the concord of number with the verb. If this agreement is not visible, the use of a tag question will identify the Subject since 'the subject is that element which is picked up in a question tag and referred to anaphorically by a pronoun' (Downing and Locke, 2006, p. 43). 'Anaphorically' here means referring back to a word used earlier to avoid repetition. The process of identifying the Subject and Verb of (12) would be carried out based on these tests as follows:

+ *Were* is a finite verb because it signals past tense and denotes the relational process of being.

+ *The teachers* is identified as a Subject because:

 o It is placed before the finite verb;

 o *The teachers* is plural and the verb *were* agrees with it.

However, in example (5) the tag question test needs to be applied since the agreement between Subject and Verb is not visible

5. Last night, *everyone* left early.

+ *Left* is a finite verb because it signals past tense and denotes the process of doing.

+ *Everyone* is the Subject because:

 o It occurs before the finite verb;

 o In the tagged version of (5) *they* refers back to *everyone* not to *last night*: *Last night, everyone left early, didn't they?*

Objects

Objects are also an important element after Subjects and Verbs, and their presence or absence is dependent on the verb. They are basically of two types: Direct Object (Od) and Indirect Object (Oi). From a semantic point of view, Direct Objects have the role of *Patient* and 'undergo the action or process denoted by the verb', whereas the Indirect Objects have the role of *Goal/Receiver* or *Beneficiary* (Aarts, 2013, p. 73). In (6) below, *I* is the Subject since it is the *Agent* of the action *send*. *An*

invitation undergoes a sending activity, and is therefore identified as a Direct Object. *Everyone* is the Receiver of the action *send*, and is identified as an Indirect Object.

    6. I have sent *everyone* (Oi) *an invitation* (Od).

Syntactically, similar to Subjects, Direct Objects and Indirect Objects are often Noun Phrases. Direct Objects can also be realized by Prepositional Phrases, finite clauses, and non-finite clauses while Indirect Objects can only be realized by Wh-clauses (Aarts, 2013; Downing and Locke, 2006; Burton-Roberts, 2016).

Their normal position is after the verb, and the Indirect Object always precedes Direct Objects (Downing and Locke, 2006; Carter and McCarthy, 2006; Aarts, 2013). However, identifying Objects based merely on their position is not enough since many other elements (Complements, Adjuncts, and even Subjects) can occur after the Verb. Take, for example, *a ski instructor* in *My brother is a ski instructor.* It is a Complement (see discussion of Complement in the next section) rather than an Object though it immediately follows the verb *is*. Therefore it is necessary to identify what kind of verb (copular, transitive or intransitive) it is before identifying Objects.

As Downing and Locke (2006, p. 85) point out, following copular verbs, of which *be* is the most common, is a Complement. There are other copular verbs expressing 'senses' (e.g. *look, feel, smell, sound, taste*) or referring to 'a process of becoming' (e.g. *become, get, go, grow, turn*). However, the study only focuses on lexical collocations, so combinations of copular *be* with a noun will be excluded. Intransitive verbs, called one-place verbs, do not require an Object or Complement to follow. Downing and Locke (2006) distinguish them into: 1) verbs of behavior (e.g. *laugh, smile, cry, cough, wait, stay, die, fall*); 2) verbs of weather (e.g. *rain, snow, rise*); and 3) verbs of occurrence (e.g. *appear, disappear, come, arrive, depart, happen*). Transitive verbs are of three types: 1) Mono-transitive verbs, two-place verbs, requiring an Object, Direct Object or Prepositional Object (Op) (e.g. *carry, say*); 2) Ditransitive verbs, three-place verbs, requiring either an Indirect Object and a Direct Object (e.g. *give, send, own*) or a Direct Object and a Prepositional Object

(e.g. *remind, rob*); and 3) Complex-transitive verbs requiring one Object and one Complement (e.g. *appoint, name, find*).

When the Indirect Object is paraphrased as in (7) below, there is a discrepancy in the way it is named. In particular, Downing and Locke (2006) call it a Prepositional Object while Carter and McCarthy (2006) identify it a Prepositional Complement. Though differently named, it is required by the verb *send*, which is a ditransitive verb. In this present study, this element is identified as a Prepositional Object on the grounds that it 'encodes participants' in the transitive clause (Downing and Locke, 2006).

(7) I have sent *an invitation* (Od) *to everyone* (Op).

Complements vs Adjuncts

Complements are the obligatory elements and unlike adjuncts, they complete the meanings of the verb. Complements are typically of two main kinds: Subject and Object Complement (Downing and Locke, 2006; Coffin et al., 2013; Carter and McCarthy, 2006; Greenbaum, 1996). Subject Complement (Cs) is the element following copular verbs and completes the meaning of the verbs by 'providing information about the subject with regard to its attribute or its identity', while Object Complement (Co) often follows the Direct Object and completes its meaning by providing information about a 'qualitative or substantive attribute expressing the name or status of the Object referent' (Downing and Locke, 2006, p. 67). (8) and (9) below are sentences containing a Subject Complement and Object Complement, respectively:

(8)The twins are *the same height* (Cs)

(9)You haven't made the sleeves *the same length* (Co).

Both Subject Complements and Object Complements can be identified syntactically under the form of Noun Phrases as can be seen in (8) and (9). They can also be Adjective Phrases, finite clauses, and non-finite clauses.

Aarts (2013) is different from most other grammarians analyzing clause elements in that she rejects Complement elements. In particular, she identifies *the prisoners jailed* in *We had the prisoners jailed* as a Direct Object of the verb *had.* So is *my brother a genius* in *Larry considers my brother a genius*. She argues that the answer to the question Who or What is Larry considering?, which helps identify Objects, is not *my brother* but the proposition that 'my brother is a genius.' This sounds reasonable enough, but in this study this analysis of the construction is rejected since as in this example it requires a change to the sentence structure into *Larry considers that my brother is a genius*. Also, syntactically the verb *consider* (*to believe someone or sth to be*) and *had* (*to cause sth to happen or someone to do sth*) are complex-transitive verbs and so require an Object and a Complement. Therefore, the sentence will be analyzed in the way that most grammarians (Downing and Locke, 2006; Burton-Roberts, 2016; Carter and McCarthy, 2006) do, that is, S-P-O-Co. This analysis is more convenient in that later on at the combination extraction stage both constituents functioning as Object and Complement will be extracted together with the verb, instead of choosing either *my brother* or *genius* if I treated them all as a Direct Object.

Downing and Locke (2006) and Carter and McCarthy (2006), however, besides these two main kinds of Complement, identify a third kind, Locative/Goal Complement (Cloc), which is often required by verbs of placement, direction or destination. This kind of Complement even extends to include location in time as in (11) (Downing and Locke, 2006, p. 86).

    10. She is lying *in a hammock* (Cloc).

    11. Lunch was *at one o'clock* (Cloc).

Locative Complements and Adjuncts are sometimes confusing and easily mistakenly recognized as each other. Take, for example, the Prepositional Phrase *in London* showing location in *Tom works in London*. *In London* here is an Adjunct added rather than a Locative Complement since it is not required by the verb, which,

according to Downing and Locke (2006, p. 87), is often interpreted as 'have a job' and therefore 'has sufficient weight in itself'.

Adjuncts are optional elements to add extra information to a clause, and are the most flexible element in terms of their position in a clause (Downing and Locke, 2006; Coffin et al., 2013; Carter and McCarthy, 2006; Greenbaum, 1996). In this study, an important distinction to be made is between Complements and Adjuncts rather than identifying types of Adjuncts, so that later on at the combination extraction stage I would know which elements need to be extracted out together with the verb. Therefore, when analyzing elements of a clause, I only called them generally Adjuncts.

2.3.2   Noun phrases

From the above discussion of clause elements, it can be seen that Noun Phrases can play the role of all clausal elements, except for the Predicator. Combinations of Adj-N, N-N, N-of-N and N-V patterns can be found in Noun Phrases, so the process of identifying these combinations needs to start with identifying Noun Phrases. In its minimal form, a Noun Phrase consists of a noun or a pronoun acting as the head, which might be modified by dependent elements before or after it (Carter and McCarthy, 2006, p. 323; Biber et al., 1999b, p. 242). At pre-head position, dependent elements are of two types: determiners and pre-modifiers while post-modifiers and complements are elements that can be found at post-head position. The following are examples of Noun Phrases (with Noun Head in bold) with dependent elements at the pre-head and post-head positions, respectively.

12. An / important   / **meeting**

   det.   pre-mod.   H

13. **Students** / of astronomy /at Cambridge

   H   com.   post-mod.

The Head of the Noun Phrase 'makes it clear what sort of identity is being referred to' and cannot be omitted without injuring the structure as well as the core meaning of the Noun Phrase (Biber et al., 1999b, p. 240). This is a criterion for identifying the Noun Head. Let us consider example (13), which consists of three nouns: *students, astronomy, Cambridge*. The identity being referred to here is *students*, and it is the only element that cannot be left out, and is therefore, the Head of the Noun Phrase.

Pre-modifiers in Noun Phrases are Adjectives, participial modifiers or nouns while at post-modifying position, and modifiers are a bit more complex. In particular, they can be a relative clause, a non-finite clause of either an –ing clause, to-infinitive clause or en-clause, a prepositional phrase, a noun phrase in apposition, or occasionally an adverb (Biber et al., 1999b, p. 575).

2.3.3　Adjective phrases

Adjective Phrases always have adjectives as their head and can be accompanied by modifiers or complements at pre- or post-head position (Carter and McCarthy, 2006, p. 440; Biber et al., 1999b, p. 242). Adjective Phrases can be part of a Noun Phrase (e.g. *It's a **rather unfortunate** name*) or be a Complement (e.g. *I was **pretty upset***). Besides these two syntactic roles, which Biber et al. (1999a) identify as the pre-modifier of the noun and Subject Complement, Adjective Phrases can be found as a post-modifier of the noun (e.g. *the three people **present***), or Object Complement (e.g. *he considered it **more dangerous than any horse**),*

2.4　Learners' use of collocations

Collocations are defined and identified slightly differently in different studies and the focus of these studies are on learners of different language levels as well as language backgrounds, so there are substantial discrepancies between the research's findings, or findings are indirectly comparable. However, that does not prevent a general picture which has emerged from these studies to be sketched. First, in terms of collocation density Laufer and Waldman's study (2011) on the use of English verb-noun collocations in the writing of native speakers of Hebrew reports that learners at

all three proficiency levels produced far fewer collocations than native speakers whereas both Howarth (1998) and Granger (1998) claim that there is not much difference in the number of prefabricated phrases, collocations included, produced between the two groups, natives and non-natives. A probably decisive factor leading to the discrepancy, apart from the criteria for identifying collocations, is the scope of the research, one focusing exclusively on verb + noun collocations (Laufer and Waldman, 2011) while the others all collocation patterns (Granger, 1998; Howarth, 1998). However, in a study looking at native and non-native speakers' use of strong collocations with low and high frequency, Durrant and Schmitt (2009, p. 167) have come to the conclusion that non-native speakers use significantly more collocations of high frequency (t ≥ 10, e.g., *good example, long way, hard work*) than native speakers do. As they explained, this is because non-native speakers overuse 'certain favoured collocations', which Nesselhauf (2005, p. 69) refers to as 'collocational teddy bears'. This explanation coincides with the claim made in earlier research that language learners tend to use repeatedly a small repertoire of collocations (Granger, 1998; Lorenz, 1999). In terms of diversity, in a recent study looking at Taiwanese learners' use of verb + noun collocations, Tsai (2015) found a significant difference in collocation use between native speakers and non-natives; type/token ratio was 13.63% for non-natives and 56.23% for natives. Durrant and Schmitt (2009), also found that non-natives underuse collocations with high mutual information score (MI ≥ 7, e.g., *ultimate arbiter, immortal souls, tectonic plates*) compared to their native counterparts. This seems understandable since collocations with high MI score are strongly associated but possibly less frequent, so often take longer to acquire.

The overall picture of learners' collocation use would not be complete if learners' collocation misuse is not depicted. Studies on learners' collocation use have all come to the conclusion that collocation is troublesome to L2 learners, even those at advanced levels. Initial studies using cloze and translation tasks (Bahns and Eldaw, 1993) and further studies using essays and reports (Granger, 1998; Nesselhauf, 2003; Laufer and Waldman, 2011) all confirm that collocations are responsible for a large number of learners' errors. In particular, in a study looking at German advanced learners' productive knowledge of English verb + noun collocations, Bahns and

Eldaw (1993) found that 51.9% of the collocations used to complete the 15-sentence cloze task were unacceptable and that 48.2% of errors in the translation task were relevant to the incorrect use of the collocates (the verbs). Quite similar figures were reported by Nesselhauf (2005) in her investigation of verb + noun collocations from a learners' corpus compiled from argumentative essays. Nesselhauf (2005) showed that the most frequent errors of collocations of this pattern lie in the choice of the verbs (e.g., *make one's homework*, *give a solution to*). Comparing the number of collocations and collocational errors produced by learners of three different levels, Laufer and Waldman (2011) found that though the number of collocations used by advanced learners increased, errors remained fairly high, at nearly one third of all collocations they attempted to produce (31.9%).

The results of a study by Dang (2014) on the lexical collocation errors of a group of Vietnamese learners at advanced level show that they are not an exception. Collocation errors make up over 45.6% of all attempts of collocational use. They often make mistakes with V-N and Adj-N combinations, such as, instead of saying *to obtain permission* and *to take medicine*, they often say *to take permission* and *to drink medicine*; *yellow* is often miscollocated with hair as in *yellow hair* while the expected combination is *blond hair*. The frequency of these kinds of mistakes is quite high, with V-N collocation errors accounting for 67.26%; Adj-N collocational errors following with 18.27%, and the rest for other combinations such as N-of-N, Adv-Adj, V-Adv, and N-V (Dang, 2014).

2.5 Causes of collocational errors

Causes that prevent learners from mastering collocation in second language learning are many. In what follows, I will only present causes that have been identified by most research in this area and that are likely to relate to Vietnamese learners, the research subjects of this study.

2.5.1 Lack of collocation awareness

The importance of collocation to language teaching and learning is undeniable among scholars in the field (Bahns and Eldaw, 1993; Nesselhauf, 2003; Lewis, 2000)

Though somewhat excessive in their judgement, Lewis and Conzett's (2000, p. 87) claim that 'collocations make up a huge percentage of all naturally occurring text' confirms this. Kozlowski and Seymour (2003), in a study of the importance of collocation in English language teaching, also assert that a good knowledge of collocation enables learners to express ideas clearly and accurately, and hence helps them to improve their writing. Collocations, nevertheless, are not given as much attention in language teaching and learning as they need (Nesselhauf, 2005, p. 3). It has been argued that teachers do not invest a sufficient amount of time in teaching collocations, which might be attributable to their lack of awareness of the importance of collocations to language teaching. This might lead to learners' lack of collocation awareness, as suggested by Nesselhauf, Liu (1999) and Wray (2002).

There seems to be only one empirical study conducted by Jones and Haywood (2004), suggesting that there is not a strong correlation between learners' awareness of collocations and collocational competence. In the study, raising learners' awareness of collocation in L2 learning was a strategy implemented to addressing students' lack of collocational ability. In particular, in order to raise participants' awareness of formulaic sequences, during a 9-week course the researchers first drew participants' attention to formulaic strings, particularly collocations, by re-presenting reading texts with those strings highlighted in bold italics. After this reading stage, the purpose of which was to encourage noticing, students were equipped with the strategies needed to use those strings accurately in their own writing. Those strings, which were believed to be useful for particular essays, were later reviewed before they were asked to write. Two sets of tests were carried out to measure students' awareness and production of formulaic sequences. As for the pre-test and post-test used to investigate the production of formulaic sequences, collocations included, they used the scoring scale to assess their accuracy. Results show that they succeeded in raising participants' awareness, but this did not bring about a significant increase in the usage of the prefabricated strings in participants' output. However, the results were of limited generalizability due to the limitations of the study, the most important of which was the relatively short two-week gap between the two essays. The authors acknowledged that this was a short period of time for an improvement to

be seen. One less important limitation of the study was the small number of students in the treatment group, with only seven out of ten students taking both tests.

Lack of collocation awareness together with lack of collocation knowledge, which I will discuss in a later section, are the primary causes, according to Dang (2014), of learners' false hypotheses. She explains that when seeing a combination like *make a decision,* learners would without hesitation combine this verb with some noun such as **make a conclusion* (in lieu of *reach/draw a conclusion*). The use of synonyms to create word combinations, unfortunately often resulting in incorrect combinations, is also proven to be a cause of errors in her study. **Deny an invitation*, **reinforce their knowledge*, **expand my experience* (instead of *refuse an invitation*, *advance/improve their knowledge*, and *broaden my experience*) are examples of how learners use synonyms of a word to replace it. This, I believe, also has roots in either their lack of collocation awareness or lack of collocation knowledge. So, whatever the strategy to help improve learners' collocation may be, it needs to start with raising learners' awareness of collocational phenomena.

2.5.2 Cross-linguistic influence

L1, which has a massive impact on L2 acquisition, has been found to pose difficulty to L2 learners in many studies. Phoocharoensil's (2012) study of Thai learners of different proficiency levels shows that L1 has a very strong impact on L2 acquisition, and this leads to some types of errors in collocations involving prepositions (e.g. by adding a preposition as in **left from my home*, by omitting the preposition as in **listen music*, and by using the preposition incorrectly as in **tired from studying*), and collocations involving lexical choice (e.g. verb + noun combination as in **My home stays at Nakhon Si Thammarat*, or adjective + noun as in **my working room*). Similar findings occurred in studies on German by Bahns and Eldaw (1993) and Nesselhauf (2003). Investigating the use of collocations in the free writing of advanced learners, Nesselhauf (2003, p. 235) found that the influence of L1 on V-N collocation mistakes was considerable (56% of all collocation mistakes) and that 'there was not a single type of mistake in which the L1 did not seem to play a role.' Her conclusion about L1 influence goes far beyond predictions in earlier studies.

In an empirical study based on a self-designed completion test comprising 40 items to measure learners' knowledge of free combinations, restricted collocations, figurative idioms, and pure idioms, Huang (2001) came to the conclusion that negative L1 transfer was the cause of L2 mistakes. He does not, however, demonstrate clearly what types of mistakes are associated with L1 influence. The impact of L1 on L2 acquisition is also demonstrated in Biskup's (1992) on Polish subjects. Errors reflect the lexical choice of L2 on the basis of L1 equivalents. Similar learner behavior can also be found in Nesselhauf (2005) and Laufer and Waldman (2011), in which they confirm that L1 influence is responsible for up to approximately 50% and 60% of collocational errors, respectively. Findings in the study by Dang (2014) about the errors of Vietnamese learners, which I find most relevant and useful to my study, also confirms the influence of L1 on L2. 128 out of 394 errors (equivalent to 31.92%) were attributable to their mother tongue influence.

Results from studies by Phoocharoensil (2012) on Thai, Koya (2003) on Japanese, and Laufer and Waldman (2011) on Hebrew impact on L2 learning produced unexpected findings. That is, high proficiency Thai and Hebrew learners depend greatly upon their L1. Laufer and Waldman suggest that the greater amount of deviant collocations produced by advanced learners in comparison to basic level learners is probably because they attempt to use more complex collocations. As Koya (2003) points out, Japanese learners with broader collocational knowledge rely more on L1 than those of lower level. In particular, in investigating the cross-linguistic impact of L1, Thai, to L2, Phoocharoensil (2012) concentrates on errors of both lexical collocations, including noun + verb, verb + noun, and adjective + noun, and grammatical collocations, including noun + preposition, verb + preposition, adjective + preposition, and preposition + noun. Results show that there is not a big difference in the degree of L1 influence on L2 between the two groups, the low-proficiency group (53.45% of total number of errors) and the high-proficiency group (46.55%). Koya's study focuses on only verb + noun collocations, and he divides the selected collocations into two groups: a) collocations with equivalence, and b) collocations

without equivalence in Japanese. The results of receptive and productive collocation tests on translation from three groups of participants, the highest group of which is supposed to be a little below university level, show that learners from the highest level rely more on L1. Koya explains that this is because the lower level group refrain from giving any answers to questions having no direct equivalence from L1.

In a study to investigate the use of collocations, particularly strings containing amplifiers, and lexical phrases in EFL writing, Granger (1998) compared a corpus of native English writing with a corpus of writing by advanced French-speaking learners and also arrived at the conclusion that L2 collocation use of learners was driven by their L1; however, this influence does not result in errors, but an underuse or overuse of some collocational amplifiers. Findings from the study indicate that learners overuse some amplifiers such as *completely* and *totally*, and as Granger (1998) explains, this is because learners feel safe when they use them when they have direct similarity with French collocations. Granger called them 'all-round amplifiers' since, luckily, none of the combinations with these amplifiers sound odd to native speakers. *Highly*, in contrast, is underused because its equivalent in French is 'only used in formal language and is relatively much less frequent' (Granger, 1998, p. 3).

In line with the above studies, in an investigation into the effects of congruency - that is, collocations with a word-for-word translation equivalence in L1 (e.g. *high salary*, *join a club*) and incongruency, collocations with no direct L1 translation (e.g. *strong tea*, *throw a party*) - on collocation processing of native speakers of English and Swedish-speaking learners, Wolter and Gyllstad (2011) come to the conclusion that in comparison to acquiring congruent collocations, incongruent collocations are more challenging. This is an important finding as regards making decision on which collocations to teach and learn, for it helps reduce a great number of tasks on the part of teachers as well as learners. However, Nesselhauf's (2005) claim about congruency and incongruency correlations with L2 collocational errors should be taken into consideration. That is, there are some cases where deviant expressions are

made regardless of congruence, and some others where non-congruence does not trap learners. From the above studies on cross-linguistic influence, it can be seen that L1, in some cases, plays an important role in facilitating L2 development when there is a correspondence between L1 and L2; its impact in some others, however, is more of a hindrance than a help (Kroll et al., 2002). Although these studies do not aim to prove L1 inhibition in collocation acquisition of non-native speakers, they indirectly suggest that L1 interference should be a matter of concern if learners wish to achieve collocational competence.

Though this research has nothing to do with collocation teaching, but rather the use of the OOCD as a writing aid in supporting learners' collocation use, my understanding of the influence of L1 on L2 is very useful in that it helps guide instruction on how to use the dictionary effectively. Rather than just giving learners the tool, instructions on dictionary use, a part of which involves which collocations need to be noticed and taken note of to learn, were planned. In this way, towards the longer term, learners can improve their collocation use.

2.5.3 Lack of knowledge of collocational properties

Another reason why collocation is troublesome to learners, as Pawley and Syder (1983) point out, is the selection of combinations of words that are commonly used by native speakers from a wide range of possible combinations. In other words, the 'puzzle of nativelike selection' of words from their store of vocabulary involves much effort and often hinders their production (Pawley and Syder, 1983, p. 194). This is not surprising since, as Wray (2002) explains, second language learners start with individual words and learn how to put them into grammatical patterns to make sentences. From this habit, whenever they encounter a formulaic string, they tend to separate them into the smallest segments (Huang, 2001), and unintentionally throw away a very important point, that is, the words that they often go with (Cowie and Howarth, 1996). For instance, in dealing with a string like '*major catastrophe',* instead of processing and storing it holistically to be ready for retrieval when the need arises, as native speakers probably do (Wray, 2002), L2 learners often learn the items separately. It seems true that they tend to operate relying more on the open

choice than the idiom principle, which says that language learners have 'a large number of semi-pre-constructed phrases that constitute single choices, even though they might appear to be analyzable to segments' (Sinclair, 1991, p. 110). Unfortunately, in so doing, they lay themselves a potential trap of making strings such as '*major toe'* or '*large catastrophe'*, which are not perceived as idiomatic (Wray, 2002, p. 209). Based on what I have just reviewed, the instructions on dictionary use must involve a requirement for the learners, along with using the dictionary to aid their writing, to take note of collocations that they think might be useful for their later use and to learn them as big words rather than separate them out.

2.6   The role of dictionaries in language learning

In the following sections I will provide an overview of the role of dictionaries in language learning to demonstrate their possible value as a tool for helping students with collocation.

2.6.1   Dictionary use and vocabulary acquisition

Nesselhauf (2005) conducted a study to explore the use of collocations by advanced learners with German backgrounds. She aimed to find out their difficulties in the production of collocations, factors that make collocations problematic, and material and strategies used to create collocations in L2. In the study, she examined a 150,000-word corpus compiled from argumentative essays, focusing on verb-noun combinations. Findings from the study showed that there was mostly no difference in the number of errors, nor in the number of attempts to use collocations for compositions written with and without consulting dictionaries. This suggests that learners consult dictionaries for other purposes rather than looking for the use of collocations. This kind of learner behaviour was also found in an earlier study by Atkins and Varantola (1997), which states that using a dictionary to find a collocation only accounted for one-tenth of all uses of the dictionary. Of the 910 look-ups, only 102 were seeking collocations of L2 words. A survey study by Lew (2004) of the use of monolingual, bilingual and semi-bilingual dictionaries by Polish learners also indicates that the number of learners not looking up collocations to do a translation task is 24.4%, and up to 43.8% of learners confirmed they hardly ever did. It can be inferred from these studies that many learners do not make effective

use of dictionaries to find collocations. In order to measure the effects of the online collocation dictionary on learners' collocation use in L2 writing, it is essential to ensure that learners first receive instructions on how to use the dictionary to make the best use of it.

Dictionaries provide learners with a vast amount of 'language input', which, according to Krashen (1982), is a sufficient condition for vocabulary learning. Reports on learners' vocabulary retention after consulting dictionaries vary widely, ranging from 33.3% (Laufer and Hill, 2000) to 76.2% (Lew and Doroszewska, 2009), and as can be found from these studies this is due to different look-up strategies. The results of the look-ups are also contributable to the difference in retention rate (Bruton, 2007). In particular, in an investigation of a group of intermediate students in a secondary school in Spain to seek vocabulary retention in a translation task, Bruton (2007) found that retention rate in the delayed post-test of the group of vocabulary items correctly looked up was higher than the group of the items incorrectly looked up (72% compared to 52%).

Laufer and Hill (2000) also conducted a study to investigate whether or not vocabulary was retained after being looked up and whether there was an association between dictionary consultation and vocabulary retention. They found that vocabulary retention in two groups of University students, from Hong Kong and Israel, was very different (62% for Hong Kong and 33.3% for Israel), and this seemed to be due to consulting preferences, such as looking up meanings of words in L1 only, in L2 only, in both languages, or looking up meanings and additional information. Findings also showed that the use of dictionaries to check both L1 and L2 meaning of a word resulted in greater recall (the highest rate (45%) in the Israel group and the second highest rate (67%) in the Hong Kong group). Lew and Doroszewska's (2009) study of vocabulary learning from electronic dictionaries reported that recall was 56.6%; however, results were much higher (76.2%) if learners consulted both the definition of the words in L2 and their translation in L1. If consulting dictionaries contributes to expanding learners' store of vocabulary, it is not unreasonable to expect that the same will happen with collocations since they are

composed of words. Although my study is not concerned with the relation between the use of the OOCD and retention, towards the end learners are expected to improve their collocational competence in writing, and the consultation from this dictionary is expected to some extent to help them with that. It can be inferred from the above studies that instruction on the use of the Oxford Online Collocation Dictionary should involve the encouragement of learners to use this dictionary in combination with other dictionaries to ascertain the L1 meaning of collocations, especially when collocates of a headword are unknown. In this way, the possiblility of learners recalling collocation and checking up to use it at a later point is higher.

Dziemianko's (2010) study of learners' retention of collocations after using a paper or electronic dictionary found that learners using electronic dictionaries gained much more than those using paper dictionaries (63.8% compared to 45.7%). However, the test in the study only aims at infrequent grammatical collocations (e.g. *up the creek*, *on the blink, in the offing*) which are supposed to be attention-catching, and therefore might be learnt in the same manner as single words are. The number of electronic dictionary look-ups was much higher, more than twice as many as paper dictionary look-ups (Roby, 1999). In a study to examine problems ESL learners might have using dictionaries, McAlpine and Myles (2003) also claim that the use of the online dictionary offered learners greater capability of accessing a larger amount of information on meanings and examples, and more importantly, more flexible searching paths. It seems that electronic dictionaries are more beneficial to language learners than paper dictionaries on the grounds that the more language input learners are exposed to, the more they will obtain.

2.6.2   Difficulties and problems in dictionary use for production

There are not many studies on the use of dictionaries for collocation look-ups, but those focusing on this suggest that learners did not gain much (Dziemianko, 2014; Laufer 2010). Laufer's study investigated the use of online monolingual dictionaries of ninety-five high school students, twenty of whom are native speakers of Arabic and the rest speakers of Hebrew, to seek collocations for a translation task of 15 sentences. The dictionaries she offered her learners to use in the study were English-

English-Hebrew (EEH) Longman Dictionary of Contemporary English (LDOCE) and either The Collins COBUILD Dictionary for Learners of English (COBUILD), Oxford Advanced Learner's Dictionary (OALD), or Cambridge Advanced Learner's Dictionary (CALD). She found that learners encountered difficulties with consulting dictionaries for the collocations they needed even though, as she described, the frequency of the thirty verb-noun target collocations was higher than 100 in the BNC. The participants reported that they could not find 20% of the collocations to do the task. The most correct collocations were found from LDOCE (24 out of 30 collocations), which, as Laufer explains, is because they are in bold or in different colours in the online version. In this study the examination of the number of collocations the other dictionaries contained was revealed as follows: OALD: 22; CALD: 16; COBUILD: 14; EEH: 12. These numbers somehow reflect the judgement that most of the 'College Dictionaries', dictionaries for learners at upper intermediate to advanced level, contain a very small number of collocations (Hottsrnonn, 1991, p. 231). Generally, the ineffective use of general dictionaries for collocation look-up is because either they do not contain many collocations, even those that frequently occur, or learners cannot find collocations that they want to look for since they are hidden in examples (Laufer 2010). At this point a kind of dictionary specializing in collocations might be a solution, though this is unproven.

In a study by Ard (1982) on the use of bilingual dictionaries in the writing of ESL students with Spanish, Arabic, and Japanese backgrounds, the results indicated that access to these dictionaries did not help learners to write accurately; conversely, learners seemed to make more of some types of errors, which are believed to be dictionary-based. This, according to Ard (1982), is because these dictionaries are more suitable for receptive rather than productive purposes. It seems that what learners need from dictionaries when generating text is not just L1-L2 equivalents but the differences in denotation and connotation meaning among the equivalents, and information on collocations as well. There seems to be some correlation between this conclusion and that of Summers (1988) on dictionary use, which indicated that learners are much better at using dictionaries for receptive rather than productive purposes. Though collocation was not the focus of his study, the focus was the

vocabulary items that learners were using dictionaries to look for. Bogaards' study (1996) of the use of the four dictionaries - Collins COBUILD English Dictionary, Longman Dictionary of Contemporary English (3rd ediction), Oxford Advanced Learner's Dictionary of Current English (5th ediction), and the Cambridge International Dictionary of English (CIDE) - for both receptive and productive purposes also confirms this. He found that all these dictionaries present examples which provide dictionary users with models to follow. However, dictionary users could hardly use them as models for their own production, partly because examples are rather long and partly because they often contain unfamiliar elements (Bogaards, 1996, p. 31). These studies suggest that besides training learners to use dictionaries for productive purposes, language educators' role in recommending to learners dictionaries written with consideration of their learning purpose is of great importance.

2.6.3 The structure of general and collocation dictionaries

One important distinctive feature of collocation dictionaries compared to other bilingual and monolingual dictionaries is that the presentation is explicitly all around collocations of both lexical and grammatical types (Lea, 2007). Atkins and Rundell (2008 p. 363) have studied style guides and instructions on how to handle multiword expressions (MWEs), of which collocations are an important group, and have identified a number of ways in which they are tackled.

1. Enter the MWE under the first or only lexical (not grammatical) word, i.e. *to pull someone's leg* in the *pull* entry.
2. Enter it under the least frequent lexical word, the one expected to have the shortest dictionary entry, i.e. *to open the floodgates* at *floodgates*.
3. Enter it under the first or only noun in the phrase, i.e. *big deal* in the *deal* entry.
4. Enter it under the first or only verb in the phrase, i.e. *to twist and turn* in the *twist* entry.
5. Enter it as a headword in its own right, i.e. individual main entries for *big deal*, *pull someone's leg*.

According to Hottsrnonn (1991), lexical collocations which consist of a base and a collocate should be presented at the entries of bases in collocation dictionaries. Along with Hausmann (1989), Benson (1989b) regards nouns in verb + noun (e.g. *pursue studies*), noun + verb (e.g. *storm rage*), and adjective + noun collocations

(e.g. *heavy smoker*) as bases and verbs and adjectives in those combinations of collocates. In adverb + verb (e.g. *severely criticize*) and adverb + adjective (e.g. *deeply disappointed*), the verb and adjective are respectively the bases and adverbs are the collocates in both cases. Neither Hausmann (1989) nor Benson (1989b) refer to noun + of + noun (e.g. *piece of advice*) or noun-noun collocations (e.g. *mountain bike*), but from their argument for why collocations should be presented at their bases, we can infer that as to N-of-N collocations the second noun (*advice*) is the base and the first noun (*piece*) is the collocate while the first noun in the N-N collocation is the base and the second the collocate. A random check from the OOCD reveals how these 'new' rules are applied:

Table 2.1 *Collocations and their frequency in the BNC*

Collocations		Frequency in BNC	
Collocates	bases	of collocates	of bases
Perform	*task*	3084	8917
Pay	*attention*	21314	13186
Commit	*suicide*	1318	1683

All of these collocations are presented at their base entries (*task, attention, suicide*) regardless of their frequency compared to their corresponding collocates. The collocation *perform a task* with the frequency of the base (*task*) almost three times as much as that of its collocate (*perform*) proves that rule number 2 is totally ignored. *Perform* and *pay* are, though, entered in this dictionary; they are treated as the bases (e.g. *pay + handsomely/well/gladly; perform + efficiently/ poorly/ adequately*). As for the case of the verb *commit*, it cannot be found in the OOCD, which is likely because, as can be found in BNC, it is not accompanied by any adverbs, in which case it is treated as the base. It seems that none of the above five rules are applied adequately and systematically.

To account for the reason why collocations should be presented at their bases, Hottsrnonn (1991) argues that when writing, in order to express an idea learners will think first of the base and thereafter look for collocates, which are supposed to be provided in collocation dictionaries to complete phrasal meanings. For instance, in the case of verb + noun as in *perform a task*, noun + verb as in *storm rages*, and adjective + noun as in *strong coffee*, learners will start thinking of the nouns rather than their collocates. Clearly, this principle is applied to the Oxford Online Collocation Dictionary. Hottsrnonn (1991) also argues that attempts to find collocations beginning with collocates are too difficult, if not hopeless, a task. This is because in general dictionaries many collocations are presented at the entry of collocates, but not at bases. In a study by Benson (1989b) aimed at investigating how collocations should be entered in collocation dictionaries, he provides examples to prove that learners might have to struggle to generate texts for that reason. The entry for the verb *draw* in Longman Dictionary of Contemporary English (LDOCE) presents collocations such as *draw attention*, *draw a crowd*, *draw a gun*, but they cannot be found at the entry of the noun base; collocations like *set the table*, *set a watch* can only be found at the entry of *set* but not of their according nouns in Oxford Advanced Learner's Dictionary of Current English (OALDCE) (Benson, 1989b, p. 7); the collocation *to meet someone's demands* cannot be found at the noun entry *demands* in Collins Dictionary of the English Language (CDOEL) (Hottsrnonn, 1991, p. 231), either. Nevertheless, the presentation of collocations in dictionaries should take account of users' needs regarding encoding or decoding (Hottsrnonn, 1991).

From the same perspective, in a study about the function of collocations in dictionaries, Cop (1990) even proposed different ways of presenting them based on different purposes of dictionary users; that is, for text reception collocations need to be presented at the collocate entry whereas for text production they are to be presented at the base entry. Another reason for advocating this way of presenting collocations in dictionaries is that it is likely to be more productive to learn a group of collocates accompanying a certain base since the base is semantically constant (Cop, 1990). The review of how and why collocations are compiled and presented

differently in collocation dictionaries, as confirmed by the authors, is for the sake of learners' production. Its actual effect, however, is so far untested.

There are no separate definitions of collocations in the OOCD, which is believed to help learners focus on the reference work (Coffey, 2010). Besides, as Coffey (2010) also points out, the meanings of collocates are supposed to be either known earlier or inferable from their semantic set or demonstrative instances. However, it is true that learners are not always able to infer meanings of collocates from examples. An example to illustrate this is the use of the adjective *sleek* in '*sleek design'*, though '*elegant/stylish design'* is very likely to be inferable (Coffey, 2010, p. 331). This might be a possible shortcoming of the dictionary. Understanding of the features as well as the possible strengths and weaknesses of the collocation dictionary guides me on how to instruct learners to use the dictionary to best facilitate their learning.

2.7   Introducing the research questions

Much of the discussion in the above sections details the collocation use of learners in L2 writing and the roles and difficulties of using dictionaries in supporting language learning. In order to investigate the effects of the Oxford Online Collocation Dictionary on advanced learners' collocation use in academic writing and their perceptions of the use of the dictionary as a supportive tool, I need to have an understanding of how they use collocations in academic writing and how they use the OOCD as a supportive tool. The research questions that my study will address are:

1. **How do Vietnamese advanced learners use collocations?**

2. **How do the students use the Oxford Online Collocation Dictionary to support their writing?**

3. **Does the use of the Oxford Online Collocation Dictionary aid the learners to improve collocation use in writing?**

4. **How do the learners evaluate the use of this dictionary in support of their writing?**

2.8 Previously used research methods

My concern in the present study is not only to understand the effects of this dictionary on learners' collocation performance but also to gain an in-depth understanding of how learners use this dictionary to support writing: specifically, whether they can use collocations correctly; if some mistakes persist, then what they are; whether or not the dictionary encourages more attempts to use collocations; and how learners evaluate the dictionary. In order to approach these issues, in this section I present a review of research approaches and instruments from other related studies.

2.8.1 Choice of research approach

According to Trochim and Donnelly (2008, p. 142), qualitative research would be the approach of choice 'for achieving a deep understanding of the issues'. As Bryman (1988, p. 103) puts it, only qualitative research methods could provide the intricate detail for a 'rich' and 'deep' understanding of a phenomenon, and this is shown when he contrasts the differences in the nature of data between qualitative and quantitative approaches. Specifically, in a study to investigate students' corpus use behaviour and their perceptions of the strengths and weaknesses of corpora as an assisting tool to L2 writing, Yoon and Hirvela (2004) chose a qualitative approach with both qualitative and quantitative data to gain deeper insights into how learners may benefit from corpus-based writing. As such, a multi-method, qualitative approach was selected. However, a non-experimental research project with neither a control group nor a pre-intervention measurement means researchers may 'have a hard time establishing which of the things they observed are due to the intervention rather than to other factors' (Trochim and Donnelly, 2008, p. 175). Consideration of how this problem might be addressed will be presented along with the choice of instruments.

2.8.2 Instruments

Different methodologies have been used to investigate the collocation use of learners in L2. Some researchers have used traditional error analysis of selected writing samples (Liu, 1999; Laforest, 1980); others have used collocation elicitation techniques (Biskup, 1992; Bahns and Eldaw, 1993; Wolter and Gyllstad, 2011), while some others, after the introduction of language corpora, have used corpus analysis techniques (Granger, 1998; Laufer and Waldman, 2011). Each type of data

has its own strengths, for which they were selected for these studies. In particular, corpus analysis techniques allow the investigator to tackle a large amount of data involving free writing. Calculating the frequency of specific words, extracting a word being considered for recurring pattern of use, and collecting data on underuse or overuse of specific collocations (of non-native speakers compared to native-speakers) are preeminent features of this method (Laufer and Waldman, 2011).

Elicitation techniques are, however, often used when researchers aspire to look at learners' knowledge of specific collocations. Translation tasks used in a study by Biskup (1992) and fill-in tasks used in a study by Bahns and Eldaw (1993) are examples of these techniques to test learners' productive knowledge of specific groups of collocations. This current study aims to explore learners' use of collocations, namely what kinds of collocations they can use correctly and what kinds they have trouble with in their free language production; hence, elicitation techniques are inappropriate. Of the other two methods, traditional error analysis was chosen in this research since the amount of learners' writing samples to be collected is not very big (29 350-word essays), and therefore could be handled traditionally. These writing samples, collected without any intervention, will be the data for answering the first research question regarding types of collocations learners can use correctly and types they are struggling with.

It is not easy to discover how learners use dictionaries. There are various difficulties; for example, participants might not behave normally as users when being researched, or the resources needed to observe every participants' move may not be available. Instruments used in previous studies with similar research aims include observations (Atkins and Varantola, 1997; Harvey and Yuill, 1997; Summers, 1988; Knight, 1994; Hatherall, 1984), questionnaires (Bulut and AbuSeileek, 2009; East, 2008) and interviews (Ard, 1982; East, 2008).

2.8.2.1   Observations

Observation has been used to discover learners' dictionary behaviour (Harvey and Yuill, 1997; Atkins and Varantola, 1997; Nesi and Haill, 2002) but the way it has

been conducted differs. In a study looking at what people do when they consult their dictionary, Atkins and Varantola (1997) used recording sheets to record step by step what went on when participants consulted a dictionary to solve a problem of translation and how they felt about the information they got. They assigned participants to work in pairs; one using a dictionary, and the other recording. This way of operating means the researcher can not only gather data with a limited possibility of missing a single look-up, but also record the dictionary search of a large group of participants without worrying about the matter of time. However, what is problematic with this way of recording dictionary users' searching activities is that participants doing the recording sheets would be unable to answer some of the questions on the recording sheets, even though to answer them they only need to circle the options given. In particular, they could hardly answer accurately 1) why the dictionary users need to do the search; 2) why they are moving to another dictionary; and 3) how they feel about each search.

In a study about the contribution of a monolingual dictionary to writing in a second language, Harvey and Yuill (1997) adopted self-observation with both introspective and retrospective information. They asked students to carry out a writing task to establish naturalness in the data construction. Students were free to develop their arguments and only consulted the dictionary when necessary. This meant that there was no constraint on the production of certain lexical items. Learners were asked to fill in the recording sheets for each word they looked up. The recording sheets covered what they were looking up and why. If they chose more than one reason for why they were looking the word up, they needed to complete one flow chart for each reason after using the dictionary. The flowchart then asked them further questions, such as whether they found the information easily and whether they judged the result of the search to be successful. Again, depending on the answers to these questions the flowchart led them to several other questions. This way of operating recording sheets can resolve problems pertaining to the accuracy of the answers since only the dictionary users know exactly what they want to look for, why they need to do the search, and how they feel about each search. However, it has its own problems.

Filling in the recording sheet while learners are doing the writing might distract them from the writing itself.

In a study looking at how the dictionary compiler knows what the user needs, Hatherall (1984) also asserts that direct observation is the most reliable method of data collection. Useful information could be gained from the data on user behaviour. However, it seems that this method is not without problems. Under observation, it would be very difficult for dictionary users to behave normally (Hatherall, 1984). In addition, the information that the researcher needs is not always retrievable via the visual dimension. To remedy the possible drawbacks of this method for capturing learners' behaviour, Hatherall (1984) chose to send them a printed statement asking for as normal as possible an approach to dictionary use when they performed the translation task. In so doing, the problem of learners' unnatural use of dictionaries might be resolved, or at least minimized.

From these studies, it can be seen that observation is no doubt a useful tool for data collection; however, care needs to be taken as to how it should be operated, otherwise it might only offer certain steps of a search without a detailed description of what information learners are seeking from the dictionary or the purpose of each check-up, which cannot be obtained if dictionary users do not spell it out. The use of retrospection with recording sheets in the study by Harvey and Yuill (1997) proves itself to be a remedy to this possible shortcoming since dictionary users are asked to retrieve information from their short-term memory to complete information on the recording sheets. 'Self-revelatory' introspection, being simultaneous with the processing of information, would not be appropriate here since it would influence the participants' concentration on their writing.

Besides observations, questionnaires and interviews have been used in other studies to gather learners' recollections of how they use the dictionary (e.g. Atkins and Varantola, 1998; Ard, 1982; Christianson, 1997). Nevertheless, they are believed to be unreliable in that participants might describe what a general look-up process

should be but not what they actually do with dictionaries (Hatherall, 1984). In a study of Korean freshmen's use of dictionaries, Kent (2001) also used questionnaires in order to gather information on how the students use dictionaries. However, with responses from those questions including where, when, and how they used dictionaries for translation (from English to Korean and vice versa), based on the scale of percentages he could hardly describe in detail what students actually did with the dictionaries, which is the focus of my study.

2.8.2.2   Questionnaires

Questionnaires are often used for gathering data on learners' attitudes. Bulut and AbuSeileek (2009) and Brett (1996), in their studies on learners' attitudes towards computer-assisted language learning (CALL), and Yoon and Hirvela's (2004) study of ESL students' attitude towards corpus use in L2 writing, all employed five-point Likert scale questionnaires to explore learners' general evaluations. There are a few studies on learners' attitudes towards dictionaries; the one by Kent (2001) on Korean University Freshmen's dictionary use and their perceptions regarding dictionaries shows that questionnaires are an effective instrument, especially when the number of participants is relatively large – as many as 270. In the questionnaires, the first two sections were closed questions regarding students' background in relation to dictionaries and students' use of dictionaries; the last section was open-ended questions, covering what students perceived as likeable, dislikeable, and desirable features of the dictionaries that they were using. However, the use of open-ended questions to ask about likeable, dislikeable and desirable features of a dictionary does not always bring about full answers from participants.

Similarly, questionnaires were also used to seek test takers' opinions in three case studies on the use of the *Collins German Dictionary* in writing tests by East (2008). In the first two studies with small groups of participants, six and five, the use of open-ended questions provided a wide range of opinions, both positive and negative, about the availability of the dictionary and its impact on the writing tests. However, in the third study, East (2008) used a five-point attitudinal scale questionnaire, not

only to collect a variety of qualitative perspectives from a larger group of participants (47 high school students) but also to quantify these perspectives. The instrument helped him do just that.

2.8.2.3   Interviews

According to Bell (2014), one of the major advantages of the interview is its adaptability. A response in an interview can be 'developed and clarified' while questionnaire responses can only be taken at 'face value' (Bell, 2014, p. 157). Interviews are often divided into structured, semi-structured, and unstructured interviews (Sapsford and Jupp, 2006; Trochim and Donnelly, 2008). In structured interviews participants are asked the same exactly-worded questions, whereas in semi-structured interviews the interviewer follows the guiding questions but is able to explore particular themes or responses further; unstructured interviews are the most flexible since the interviewer can adapt or change the interview questions subject to the respondents' answers (Trochim and Donnelly, 2008). As Denzin and Lincoln (2003, p. 74) point out, semi-structured and unstructured interviews help researchers generate qualitative data through the use of open questions through which they can develop a real sense of a person's assessment.

In some studies about learners' perceptions of the use of corpora (Yoon and Hirvela, 2004) or dictionaries (Koyama and Takeuchi, 2004) as a second language writing tool, questionnaires have often been used together with interviews. Yoon and Hirvela (2004) conducted semi-structured interviews for additional data, through which, they believed, the reasons underlying learners' attitudes could be unveiled. Similarly, interviews were administered in the first two studies by East (2008), in order to give participants the opportunity to expand on comments they had made earlier in the questionnaires.

Chapter 3 : RESEARCH METHODOLOGY

3.1 Introduction

In the previous chapter, I presented a review of issues concerning collocations and dictionaries comprising definitions of collocation, classification of collocations, collocational errors and causes of mistakes as well as dictionary use and their role in the development of L2 writing. Research approaches and methods used in the previous studies to address similar research questions were also discussed.

In this chapter, I will first present a detailed discussion of my research design and the administration of research instruments in this study. The pilot study carried out to test the data collection process will then be presented. Data collection procedures for the main study, involving the recruitment of participants and how data, quantitative and qualitative, were obtained, will be then discussed. Ethical considerations and validity issues will also be addressed in this chapter.

3.2 Research design of this study

Informed by the research design and research tools presented in the Literature Review chapter, in this study I used learners' essays without and with the support of the dictionary as the data to examine learners' collocation use. The first set of data, learners' essays written without consulting the dictionary, was collected at the beginning of the data collection process. After collecting the first data set, I introduced the OOCD to the learners.

When introducing the dictionary to the learners, it was important for me to take some factors into consideration. In particular, the research was primarily looking at the effects of the dictionary use and learners' perceptions of dictionary use rather than the use of the OOCD in teaching collocations, so I needed to ensure that all the activities should only be enough to help learners get familiar with and know how to make the best use of the dictionary, but not to the extent that collocations were learnt and became their language intake. Undoubtedly, it is hard to not affect to their store

of vocabulary. The impact would be less likely solely because of the dictionary use, but because of the change in their collocation knowledge through the process of interacting with the dictionary. What I could do to minimize this to the utmost was to avoid activities that required learners to interact repeatedly with the same collocations, because five times of encountering the same collocations might lead to the collocations being learnt (Webb et al., 2013), and hence language input would become language intake.

After the dictionary was introduced and learners had become familiar with it, I asked learners to produce the second piece of writing with dictionary support. How these data would be generated, what instruments would be used, how those instruments were designed, and how they were used to generate data needed to address the research questions are described in the sections below.

3.2.1 Observations by Recording Sheets

Given the discussion about research instruments in the previous chapter, in order to answer the second research question about how learners use the dictionary to support their writing, I chose what Atkins and Varantola (1997) call a 'paper approach' to record step by step what was going on when learners turned to the dictionary. The use of observation with recording sheets allowed me to gather similar information as well as every single move of quite a big group of participants, and more importantly, information that could only be obtained when spelled out by participants would be less likely to be missed.

Similar to the way in which Harvey and Yuill (1997) constructed data, in order to assess its impact, I asked learners to write a 350-word essay on a given theme with the support of this dictionary. This was part of my attempt to monitor the use of the OOCD under as natural conditions as possible. Learners were allowed to use other dictionaries if they wished to; however, they were encouraged to use this dictionary for all collocation check-ups. Once learners were allowed to use other dictionaries when necessary, plans for excluding their effects in supporting writing needed to be considered since this study was looking solely at the impact of the OOCD, not dictionaries in general. How this was done will be discussed in the data collection

section. The exploration of the contribution of the OOCD to L2 writing was done based on the assessment of learners' productive collocation use from their written work.

Recording sheets were used to generate the data, and their role is twofold: (1) to provide a detailed description of how learners consulted the OOCD, and (2) to help trace back all collocation searches with information on which and how many collocations they looked up, whether or not the use of those collocations was correct, and whether or not they used the dictionary in combination with other dictionaries. The combination of learners' written work and recording sheets helped specify the result of the intervention more accurately, and therefore they were the data for answering the third research questions.

Following Harvey and Yuill, I used a homework assignment to assess the learners' performance. Learners would do it in class instead of at home to ensure that they would do it without somebody's help. They were to write a 350-word essay with the support of the OOCD in 45 minutes. Recording sheets would be used to record what they were doing when approaching the dictionary for help. The design of the recording sheets in this study (see Appendix 2) was adapted from Atkins and Varantola (1997). Besides recording some similar information in order to portray how the process occurred, including what entries users were looking for, whether or not they were able to find what they were looking for, whether they used the dictionary being considered in combination with other dictionaries, and how they evaluated each search, I would ask some more questions to gain insights into learners' use of the dictionary, such as the purpose of each check-up, whether or not participants knew how to use the collocations that they found from the dictionary, or what types of collocations sent them to consult the dictionary. Information on the types of collocations the students were looking for is necessary because only the dictionary users knew exactly what they wanted. The table below illustrates how the recording sheet (RS) was adapted in the present study:

Table 3.1 *Questions adapted from recording sheets of Atkins and Varantola*

N	Atkins & Varantola's recording sheets	Present study's recording sheets
1	What made the dictionary user go to the dictionary?	Purpose of looking up: Checking (C)/ Finding (F)
2	What dictionary are they using?	
3	What entry are they looking up?	Words check-up
4	Why do they need it?	
5	Have they found what they were looking for?	Words found: Yes/No Instruction on using: Yes/No
6	If yes, where did they find it?	
7	What are they doing next?	
8	If they are moving to another dictionary, why?	*Using OOCD in combination with other dictionaries: Yes/No*
9	If they are ending this search, how do they feel?	Evaluation
10	Any other comment	Detailed description
11		Types of collocations

As can be seen from the above table, some questions from the recording sheet in Atkins and Varantola were not asked in the present study and vice versa. This is because Atkins and Varantola's research aim was to look at what learners were actually doing when they consulted their dictionaries, whatever dictionaries, for help with a translation task, whereas the present study focuses on the actual use of the OOCD. In particular, question 2, about what dictionaries they were using (the dictionary they used for the first look-up of each search), was not asked in this study; question 6 (If they have found what they were looking for, where they found it) and question 7 (what they did next) were not asked either, for this same reason. The research is interested solely in the use of the OOCD. Instead, questions were asked regarding the types of collocations that led learners to dictionary consultation, whether they could find target collocations from the dictionary, and whether they could get instructions on how to use them correctly.

The participants would be arranged to work in pairs, one partner using the OOCD, the other recording every check-up on the recording sheet. The idea of arranging participants to work in pairs was taken from Atkins and Varantola (1997). The purpose of so doing was to make sure that no collocation check-up would be missed, and more importantly, that those doing their writing would not be distracted. Participants in charge of recording dictionary use were to be arranged to sit behind their partners, who were supposed to do their writing at the time in order to ensure the minimum possible interference. Instead of asking the participant doing the recording sheet to complete all the information on the sheet as Atkins and Varantola did, in this research most of the information was completed by the participants doing the writing right after they finished their written work, except for column 2 (what headwords were checked up) and column 7 (whether the dictionary users used the OOCD in combination with other dictionaries). Participants were requested to exchange roles after the first 15 participants (half of the participants) had finished their writing. This swap mid-way would allow me to gather more data from all of the participants, and would enable the participants to feel fairly treated.

Each participant would be provided with a computer installed with a Vietnamese-English Dictionary and an English-English dictionary, which could be an Oxford Advanced Learner's Dictionary (OALD), Longman Dictionary of Contemporary English (LDOCE) or Cambridge Advanced Learner's Dictionary (CALD).

Every time the OOCD was used for checking up, the headword would be recorded. If participants started with the Vietnamese-English dictionary to look for an equivalent word to express an idea, or an English-English dictionary to check the spelling of a word, these steps were not recorded. As the OOCD could only be used to check for collocates of a word if users remembered its spelling, sometimes participants had to start with an English-English dictionary. If, after consulting the OOCD, participants turned to other dictionaries for the same headword, it needed to be noted down as being used in combination with others.

3.2.2   Questionnaires

To address the last research question about learners' evaluation of the OOCD as a supportive tool, I used Likert scale attitude questionnaires (see Appendix 3**).** These were adapted from Bulut and AbuSeileek (2009). The questionnaires were divided into two parts corresponding to the scales of response, and comprised a series of 17 statements. Participants were asked to tick one of the four degrees, from strongly agree to strongly disagree. Unlike a CALL program or corpora, dictionaries in general have their own features, such as the number of headwords, layout, and example sentences. The survey questions about dictionaries need to include questions about these features. Therefore, the content of the statements in the questionnaires needed to be amended appropriately. Based on participants' responses regarding the likeable and dislikeable features of dictionaries in Kent (2001) and the review of dictionary use in this study, the questionnaires were organised into 4 groups of statements, as follows:

- Group 1 (statement 1, 12, 13, 14) aimed to explore students' evaluation of the effectiveness of using the OOCD for collocation check-ups. As I have mentioned in the Literature Review chapter (section 2.6.1), learners did not make use of other dictionaries effectively since they did not use them

for collocation consultations, and if they did, they could not find them easily since they are not explicitly presented.

- Group 2 (statements 2, 3, 4, 16, 17) asked respondents about OOCD accessibility and ease of use, which are considered to be the benefits of online dictionaries compared to paper dictionaries. (McAlpine and Myles (2003)).

- Group 3 (statements 5, 6, 7, 8, 9) concerns learners' satisfaction with the use of the OOCD for writing improvement since, as Summers (1988) found, learners are better at using dictionaries, collocation dictionaries excluded, for receptive rather than productive purposes (section 2.6.2).

- Group 4 (statements 10, 11, 15) looks at learners' frequency of use as well as their commitment to future use of this collocation dictionary in order to comprehend their general evaluation of the effectiveness of the OOCD.

Likert scale questionnaires were handed out to all participants after they had completed their writing and recording sheets.

3.2.3 Interviews

Given the discussion about the potential of interviews to gain in-depth understandings, I used this instrument after the questionnaire data were collected. The interviews aimed to explore the likeable, dislikeable, and desirable features about the dictionary (see Appendix 4). Semi-structured interviews were preferable since, as Denzin and Lincoln (2003, p. 74) point out, this way of interviewing helps researchers generate qualitative data through the use of open questions through which they can develop a real sense of a person's assessment. Information from the interviews helped me not only arrive at a general understanding of the impact of the dictionary in supporting learners' use of collocation, but also to comprehend the reasons underlying their attitude towards the OOCD, through which my prediction of the potential strong points of the OOCD, as discussed in section 2.6.3, could be verified.

The semi-structured interviews would be conducted informally face-to-face. Participants were chosen for the interviews based on their choice of future use of the

OOCD (statement 10 in the questionnaires). To gain a broad range of student perspectives on the use of the OOCD to support writing, I needed to choose participants equally from the four choices of the Likert scale.

3.2.4   Assessment of students' collocations in written texts

The assessment of written work was based on learners' use of lexical collocations in form and form-meaning link. In other words, learners were expected both to use strings of collocation correctly and to understand the meaning of the strings that they used. This is because from my own experience, I have seen that when using collocation dictionaries, learners might be provided with a lot of collocates of a base to choose from, but they might not know the meanings to use appropriately. Take, for instance, *receive attention* and *attract attention*. They are both correct and high-frequency collocations in the BNC. However, instead of using *attract attention,* the use of *receive attention*, as in *They tried to receive attention by sounding their horns*, might sound odd.

Informed by the discussion of grammatical units in the previous chapter, in what follows, I am presenting how combinations of the seven patterns could be identified.

N-V

Once all elements of a clause are identified, the subject and the verb constituting the N-V pattern will be extracted if the subject is a noun (e.g. *mom + bake*). When the subject is a Noun Phrase, only the Noun Head is picked out together with the verb. When the head of a Noun Phrase is a proper noun, the combination is not extracted for further consideration since my focus is on common nouns. If the Subject is a personal pronoun which refers to a noun provided earlier in the written texts, the nouns will be extracted together with the verb. Gerunds, though formed from verbs, can act as nouns, and so need to be extracted together with the verb for consideration such as *smoking + cost* in the below example:

     (14) *Smoking costs a lot of money*.

However, when the whole –ing participial clause is the Subject of a verb, as in (15) below, the combination is disregarded since the whole clause is the Subject rather

than the gerund. Though *Roller coaster* and *scare* are adjacent and seem to form a N-V combination, they will not be extracted as a potential collocation either for this same reason.

> (15) *Riding a roller coaster / scares / my little brother*.
>   S   V   O

Likewise *city* and *take* in (16). They would not be picked out since *city* is not the Subject of *takes* but the whole to-infinite clause:

> (16) *To commute to the city / takes / three hours*.
>   S   V   O

Since this study only focuses on lexical collocations, copular be, if occurring with a noun, will not be extracted. If following a noun is *have to + verb* as in *Kids have to go to school*, the noun will be extracted together with the verb after *have to* (e.g. *kids + go*), since *have to* is a modal verb.

N-V pattern can sometimes be found from a Noun phrase when the noun head (in bold) is modified by a relative or adjectival clause, as in (17) below:

> (17) *I   like   the **table**   that stands in the kitchen*.
>   Det. H   post-mod.

Table and *stand* in such a case need to be extracted since clearly the relative pronoun *that*, the subject of the relative clause, is used to replace the noun *table* and the verb *stands* is in agreement with that noun.

In cases in which the Noun Head is post-modified by a non-finite clause, such as a to-infinitive or a present participle clause as in (18) and (19) below, the N-V combinations (e.g. *man + answer; car + come*) will also be extracted based on the argument that those Noun Heads are the doer or agent of the action.

> (18) *the **man** to answer the question*.

> (19) *the **car** coming down to the road*.

V-N

As Nesselhauf (2005) pointed out, the verb - object (VO) structure does not reflect the whole view about V-N collocations, though the structure is the most frequent of V-N collocations. She found that V-N collocations could occur in other syntactic structures such as verb - complement (VC) (e.g. *come into existence*), or verb - object - complement (VOC) (e.g. *keep sth under control*), in which the verb has a 'tighter relationship' with the noun, playing the role of a Complement (*control*) rather than the Object. In light of this, other syntactic patterns that contain Verbs and Nouns will also be included for consideration since excluding any syntactic patterns at this early stage might also exclude some learners' problems related to collocation use.

Once all components of a clause are identified, I will focus on elements of the Predicate; that is, the Verbs, Objects, Complements and Adjuncts. All the components of the Predicate will be picked out together as long as they contain nouns, except for Adjuncts, which, as I have just mentioned, are used in sentences to add more information and hence are not an obligatory element. The following sentences illustrate how it would be carried out:

(20) *He / kept / the process/ under control.*
　　　S　　V　　　O　　　　　C

(21) *They / have closed / that restaurant on Alexander Street.*
　　　S　　　V　　　　　　O

(22) *I / kept / the letter/ in my desk.*
　　　S　V　　O　　　A

In (20) the Predicate is analyzed as verb-object-complement. Both the Object and the Complement contain nouns and are therefore extracted together with the verb (e.g. *keep + process + under control*). *Process* is the Noun Head of the Noun Phrase *the process*, so it is extracted out. The noun *control* is part of the prepositional phrase *under control*, and since I do not know if the preposition *under* has some role or is required in this V-N combination or not, I will extract it out together with the noun *control*. In (21), after identifying components of the Predicate (VO), I will only extract *closed + restaurant*. *On Alexander Street* is left out, though it contains the noun *street* since it is just a post-modifier of the noun head *restaurant* of the Noun Phrase, and is therefore not an obligatory element. This also means that *closed + street* will not be extracted to be considered for their relationship. Similarly, in (22),

after identifying the structure of the Predicate as VOA, I will extract the Verb and Object and disregard the Adjunct for not being a required element.

Similarly to the N-V pattern, cases in which the whole clause, finite or non-finite, is an object of a verb as in the below examples (a non-finite clause in (23) and a finite in (24)), the V-N combinations will be dismissed.

    (23) *Daniel /avoids /using chemicals on the vegetables he grows.*
          S   V   O

    (24) *John / was enquiring/ why the injection needs repeating every year.*
          S   V   O

V-N combinations in finite and non-finite embedded clauses are extracted for consideration based on the same rule of the finite main clauses I have just mentioned. For instance, *using + chemicals* is to be extracted since it is comprised of a gerund formed from a verb with noun modifiers and/or Objects, and hence might reflect learners' collocation use. As in (25) and (26) below, verbs in gerund or infinitive forms are identified and extracted together with the nouns. As such, *riding + roller coaster* and *bake + his mother a cake* will be extracted.

    (25) (*Riding / a roller coaster) scares my little brother.*
        V   O

    (26) *Daniel helps (her friend / bake / his mother / a cake).*
               S   *V*   *Oi*   *Od*

As can be seen in (26), the non-finite bare infinitive clause functioning as an Object has a Subject of its own (*her friend*) (Downing and Locke, 2006; Aarts, 2013), so the combination of *friend + bake* needs to be picked out for consideration as well. When considering a non-finite clause playing the role of an Object in the sentence (27), besides the *buy + car* combination, I will also extract *girl + buy* since *girl* is the implicit Subject of the verb *buy* (Downing and Locke, 2006; Aarts, 2013).

    (27) *The girl / regrets / buying a sports car.*
          S   V   O

If several verbs are coordinated with a noun, they will be extracted as separate combinations. Take, for example, the sentence *The students arrived at the bus station early but waited until noon for the bus.* Both of the verbs *arrived* and *waited* are

coordinated with the subject and will therefore be extracted (*students + arrived, students + waited*).

When the noun in the V-N or N-V pattern is a Noun phrase of either N-N or N-of-N pattern rather than a single noun, the noun head will be extracted. For instance, *interpret + music* and *merit + penalty* will be picked out from the sentence *they interpret a piece of music*, and *they merit the death penalty*. As for the case of N-of-N, the head noun is not fixed. There are basically two possibilities: 1) Pre-modifier – head N (e.g. *a piece of **music***); and 2) head N – Post-modifier or complement (***the aim** of the study*). So the decision on which of the two nouns in those patterns is the head will be made based on the verb used. In fact, sometimes it is not clear which of the two nouns is the head (e.g. *the power of imagination*). In such cases, the whole noun phrase can be taken as the collocating noun (Nesselhauf, 2005). With regard to the N-N combination, it can constitute either a compound noun or modifying N-N. Whatever the case is, the whole N-N combination (e.g. *tennis shoe, beach resort*) will be extracted as the collocating noun.

Adv-V

On the basis of identifying the components of a clause, I will look for adverbs that modify verbs in a sentence. Adverbs, if occurring, will most typically be found in the adjuncts as in (28) or sometimes in the complement as in (29).

(28) *I / ate / my dinner / **very slowly***.
    S V   O   A

(29) *Could / you / put / it / **just there** / please?*
            S   V   O   C

Adverbs of time, place, and frequency will not be considered since they are supposed to pose less difficulty to language learners than adverbs of manner and degree do, and more importantly they are not included in the collocation dictionary that I was examining (McIntosh, 2009). They tell when, where, or how often rather than how or the extent to which something is done or happens. However, adverbs of manner and degree will not be considered either if they modify the whole clause rather than a certain verb. Since adverbs can appear anywhere in a sentence, at the beginning, in

the middle or at the end, identifying whether an adverb modifies verbs (and which verbs) or the whole sentence is based on prediction of the intended meaning that the student writers wanted to convey. This in reality did not cause much difficulty, since in theory when modifying the whole sentence adverbs often appear at the beginning of the sentence and express the speaker's opinion about what is being said (e.g. *Honestly, it doesn't matter*) (Carter and McCarthy, 2006). *Honestly* modifies the whole sentence rather than the verb *matter*, and will not be considered.

When extracting combinations of this pattern, attention was paid to cases where there was more than one adverb occurring in a sentence. These adverbs will be ascertained to coordinate with the same verb if they are separated by a comma or a conjunction (e.g. *and, but*) as in *he conveyed a complex message clearly, accurately.* In this way, the two combinations will be extracted separately (*convey + clearly; convey + accurately*). Cases in which two or more adverbs occur in one sentence such as *violins are played extremely rarely*, *play + extremely* will not be picked out as a potential collocation for further consideration since the adverb *extremely* modifies the other adverb *rarely* rather than the verb itself.

N-N

Some heads may consist of two nouns, and they can be either compound nouns, of which the initial noun identifies the type of entities denoted by the noun following it (e.g. *video shop*), or modifying N-N (e.g. *a mountain bike*). They both constitute the N-N pattern. There is, however, not a clear distinction between the two except for the difference in stress pattern, which cannot be exploited in this study since the data are written texts. It is, in fact, unnecessary to distinguish the modifying N-N (e.g. *a mountain bike*) and compound nouns of the kind like *bus stop*, so they will both be picked out for consideration in the analysis process.

N (quantifier) of-N

Quantifiers are defined by McIntosh (2009 p.3) as 'words used to talk about the amount of something, such as *a **drop** of water*, or *a **piece** of information'*. It can be understood from the definition that quantifiers that McIntosh refers to are not just a

type of determiner (e.g. *a lot of, few, many*) but nouns of different kinds: 1) nouns denoting container (e.g. *barrel of, pack of*); 2) nouns denoting shape (e.g. *pile of, heap of*); 3) nouns denoting measurement (e.g. *inch of, gallon of*); 4) nouns of numerals (e.g. *dozens of, scores of*); 5) nouns denoting large quantities (e.g. *a load of, a mass of*); 6) nouns ending in –ful (e.g. *handful of, spoonful of*); 7) nouns denoting pair (e.g. *pair of, couple of*); 8) collective nouns (e.g. *batch of, bunch of*); 9) unit nouns (e.g. *chip of, sheet of*).

Once the noun head of a noun phrase is identified, this N-of-N pattern can be extracted. If they are comprised of Noun Phrases rather than single nouns (e.g. *fundamental aspects* and *the social life* as in *fundamental aspects of the social life*), only the noun elements of the phrases are extracted (e.g. *aspects of life*). Though the study only focuses on combinations of which the first noun element is a quantifier, combinations of such a pattern, no matter whether the first noun is a quantifier or not, were also assessed for their conventionality. This is an attempt to avoid missing any collocational errors.

Adj-N

Once a Noun Phrase and then its Noun Head are identified, adjectives, if occurring at the pre-head position (e.g. *important meeting*), could be picked out together with the Noun Head for consideration. Other cases when adjectives are at post-head position, like *possible* as in *the shortest route possible*, are excluded since, together with the noun, they do not constitute the Adj-N pattern. When more than one adjective is coordinated with one noun (e.g. *a slim rather than substantial **plant***) they will be extracted separately (e.g. *slim + plant*; *substantial + plant*).

Problems can sometimes be the present of the "-ing" form of the verb. The present participle (Verb-ing) can act as an adjective to modify a noun, and hence will also be extracted in this group. The same form can be used as a gerund, which is also in the Verb-ing form. The following examples are to illustrate how present participles and gerund scan be identified:

(30) *Growing **plants** need a lot of **water***.

(31) *Growing **plants** is my **hobby***.

The distinction is reflected in the relationship between the -ing form and what follows and in the agreements of the Noun Phrase with its verb. So *Growing* in (30) is a present participle modifying the noun head *plants* since it presents a characteristic of the *plants*, while *growing* in (31) is a gerund since it denotes an action. Another way that might help distinguish between them is the agreement of verbs with their subjects. When a gerund is the head and subject, its verb is always in singular form, whereas a verb of a Noun Phrase subject can be singular or plural depending on its subject and *plants* is plural; the singular verb is used, so it must be the gerund. Therefore, *growing plants* in (30) will be extracted into this group.

Adv-Adj

Adverbs at the pre-head positon together with the Adjective Head constitute the Adv-Adj pattern, and will be therefore extracted as potential strong collocations. In the case of the degree adverb *enough* at post-head position, it also modifies the Adjective Head, and will be extracted together with the adjective.

Collocation errors are identified as a mismatch of collocates to a base. Minor spelling mistakes in their writing that do not misrepresent the recognition of a collocation as a whole are not regarded as incorrect. The process of identifying whether a combination is acceptable or not is far from easy because decisions about the acceptability of combinations are about the degrees of likelihood, not certainty, since "range-extension" is an important feature of language change (Carter and McCarthy, 1988). Given the discussion of the degree of acceptability in the literature, I adopted strategies for determining acceptability of combinations from Nesselhauf (2005). In doing so, I needed both British and American informants to help me make judgements on collocation use. This was also to ensure that we could accurately assess learners' use of collocations as well as the association of collocational form with meaning from context (their writing). When evaluating collocations used in the

writing, we marked combinations that had not met the threshold of five from the BNC as:

(+): acceptable

(-): unacceptable

(?): questionable

I had different native informants for the pilot and the main study since those who helped me with the pilot study could not continue to help due to time constraints. I will therefore introduce them in the later sections during which they are involved.

I feel that it is important for me to discuss how my awareness of my relationship with the participants could affect the whole research process. I will therefore discuss it before continuing with the pilot study description. I got involved in this research with the role of both the researcher and the teacher, and this brought some benefits to the study. It gave me overall control of what was going on during the research process and was possibly the reason that no participants withdrew from the study. However, this dual role and my relationship with the students would possibly impact in other ways on the whole research process. Firstly, my intrusion in the research setting with the role of the researcher rather than solely a teacher might have changed learners' behavior regarding dictionary use. In particular, when doing their writing in class with the support of the OOCD, participants might have used the dictionary more frequently than they actually needed to. This possible change in learners' behavior towards dictionary use was recognized in Hatherall's (1984) study of dictionary use. As I have mentioned in the literature, to partially address this issue he chose to send the participants a printed statement requesting them to approach of the task in a normal way as far as they can. Following this, I also announced to all the participants the purpose of the study and my expectation of their normal approach to the dictionary for help.

Not only might the relationship have affected the participants' behavior, but it might also have impacted on their response to the questionnaires as well as interviews. To ensure that the impact was reduced to a minimum, I gave their recording sheets back

to the participants chosen for the interviews. In this way, I hoped participants would recall collocations they had looked up and would be consistent with their report on dictionary use. I also reminded them at every single data collection step that it was vital that their replies were honest and that how they actually evaluated the dictionary mattered. In doing so, I believe that the data is trustworthy.

3.3 The pilot study

3.3.1 Rationale of the pilot study

The purpose of the pilot study was to examine the feasibility of the research design and data generation methods. Through the pilot study I could assess:

- whether the recording sheet adapted from Atkins and Varantola (1997) was answerable for the participants, and whether the requirement of participants to provide and categorize the types of collocations they were looking for as well as to give detailed descriptions of each look-up could be too great a requirement for them;
- whether working in pairs to do the recording sheet and the writing worked;
- whether the process of analysing and assessing participants' written work focusing on the use of collocations was feasible;
- whether questions for the interviews were answerable.

The questionnaires were the only instrument not given out to the participants at this piloting stage, but these would be proofread by the native speakers who helped me with the pilot study. The pilot study was a good opportunity for me to practice and train myself in the role of an interviewer.

3.3.2 Choice and approach to participants in the pilot study

I piloted this data collection procedure with Masters students at University of Leeds. Participants recruited for this pilot study had to be similar to the target participants of the main study. That is, they had to be non-native language learners who were around upper intermediate to advanced level of language proficiency. Therefore, postgraduate Masters students were potential participants for the pilot study since one of the criteria for university admission was 6.5 in IELTS score, with no

component below 6. Thanks to the help of the student support officer of the School of Education at University of Leeds, my volunteer recruitment message was spread around. Not long after the message had been released, I had enough participants for the pilot study: 3 MA students, 2 from China and 1 from Nigeria, contacted me to get involved.

The pilot study was planned to last for 1 month with at least 2 meetings between each participant and me. We mainly contacted each other by email and sometimes by phone to make appointments during the pilot study. Information sheets about the purpose of the pilot study and activities that they were to do were given to them at the first meeting. After reading the information sheet and having no query about the research, they were asked to sign informed consent forms. Concepts such as what collocation is as well as what the OOCD is were introduced to the participants at the first meeting. Besides giving instructions on how to use the dictionary to support their writing, I held discussions with them to ensure that they were aware of the importance of collocation in academic writing.

3.3.3   Pilot study procedure

The first phase of data collection, collecting participants' writing, was skipped since: (1) it was learners' writing without any intervention, and therefore was assumed to be collected without any difficulty; (2) this pilot study only aimed to test research instruments and data collection methods rather than the actual result; hence, the comparison of collocational use with and without intervention was unnecessary. What participants were required to do was to explore and use the dictionary to support their writing in order to get familiar with it. During this period, I made myself available for them to contact when queries regarding the use of the dictionary arose. Their involvement in the pilot study was to help me assess the feasibility of the research design and data generation methods rather than the actual result of their writing, hence we did not need to get in contact as regularly as I would do with my students in the main study. All the data were collected at the second meeting with each participant. At the second meeting, they were doing their writing with the support of the OOCD. The interviews were conducted after they had finished their writing. This face to face meeting with each participant was arranged one week in

advance. To get ready for these meetings, I asked them to bring a laptop installed with an L1-L2 dictionary and a monolingual dictionary (e.g. OALD, LDOCE, or CALD). The OOCD could be accessed online.

Due to the very limited number of participants in the pilot study and the fact that it was hard to arrange for them to do the writing in one meeting, I played the role of the participant completing the recording sheet while they did their writing. This is assumed to have no effect on the result of the pilot study since, as I have described in the research design section of the main study, participants doing the recording sheet only filled in columns 2 and 7 about what headwords were checked up and whether or not their partners used the OOCD in combination with other dictionaries. My role as a participant competing the recording sheet also brought me some hands-on experience in matters arising while doing the recording sheet, from which I could instruct participants in the main study to do this task well. Assessment of the recording sheets and the process of collecting this data was done after the meetings with all the three participants.

Under my supervisors' recommendation, I contacted two British PhD students to ask for help with the process of assessing participants' collocation use in written work a month before the pilot study started. One of these, Laura Grassick, was a final year student and the other, Ben Tuner, was in his first year. They also helped me with proofreading the questionnaires. Following some of their suggestions, I amended some of the questionnaire statements. This involved adding information or omitting some redundant elements to make the statements clearer, as follows:

Statement 3: <u>Compared to paper dictionaries,</u> this dictionary saves me time for each check-up.

Statement 8: ~~I believe~~ the OCD will help improve my writing skill.

Statement 13: When I search for information in the OOCD, I ~~usually~~ get the information that I need.

Statement 14: With this assisting tool, I ~~tend to~~ use more collocations if applicable.

As I have described in the main study, the process of analysing the written work began with my identification of combinations of the seven patterns that the study aimed to investigate. I then examined whether they were conventional based on the use of the BNC. Those combinations that were not conventional were highlighted and then were passed to the native speakers for making judgements with three-scale judgements: (+) acceptable, (-) unacceptable, or (?) questionable. More importantly, they checked whether the participants understood the meanings of the strings that they had used, and whether they were appropriate in the given contexts. The examination of this procedure was very necessary in that it helped me foresee and plan to deal with various cases that arose in the main study, such as difficulties in identifying combinations for consideration from learners' essays which contained more or fewer grammatical errors; how to treat collocations with minor mistakes in spelling; and how to use recording sheets to trace back collocations in the writing systematically.

With regard to physical, emotional, and financial aspects, the participants in this pilot study were assumed to only suffer negligible risks. In contrast, participating in this study exposed them to a writing resource that they were not familiar with and this was a benefit for them.

3.3.4   Findings from the pilot study

First of all, the pilot study was carried out to assess the effectiveness of the recording sheet. Through the pilot study I found that though being instructed, participants sometimes got confused when filling out the recording sheet, especially for column 8 which required them to describe the purposes of using other dictionaries after check-ups from the OOCD. In the last column, about evaluation, it seemed that participants misunderstood *satisfied* and *OK* given as examples with options that they could opt for. In order to avoid such confusion, some necessary adjustments needed to be made: 1) All requirements on the recording sheet would be presented in the form of questions instead of phrases; 2) No example of the evaluation would be given; participants were to choose from the Likert scale of satisfaction instead; 3) Column 8 about *detailed description* would be changed into *other comments* and *why did they*

use the other dictionary in case they did. In contrast to my worries, participants could easily complete information about the types of collocations that they were looking for, so no amendment was needed for this question (see Appendix 5).

The pilot study proves that working in pairs to do the recording sheet and the writing worked well and was beneficial for the participants doing the writing since they were not distracted. From getting involved in this study myself as the 'participant' doing the recording sheet, I recognized that it is of great importance to instruct dictionary users to use the singular form of a noun when looking for its collocates, otherwise the word that they want to look for cannot be found.

With regard to questions for the interviews, they all seemed clear and answerable to participants except for the question about desirable features of the OOCD. This question would either be asked in both languages, English and Vietnamese, for clarification in the main study, or changed to be more comprehensible.

Finally, the process of analysing students' written work was examined. The procedure for the pilot study was as follows:

1. all combinations of the seven patterns considered in this study were underlined;
2. combinations used with the support of the OOCD were highlighted by highlighters;
3. all combinations were checked for their occurrence on the BNC (threshold of 5);
4. combinations not meeting the threshold were then judged by the native informants with three-scale judgement: acceptable (+), unacceptable (-), and questionable (?).

The native informants examined all combinations not meeting the threshold, focusing on the association of form and meaning to see whether the students really understood the meanings of the combinations they had used. This process of analysing students' use of collocations could show whether or not the OOCD helped the learners to use

collocation correctly and the proportion of collocations used with OOCD support compared to all collocations in the writing, from which I could assess whether and to what extent the dictionary helped students in their writing. This process, generally speaking, worked well in this study. However, for the main study, some adjustments needed to be made to ensure that the data analysis process would bring about accurate results. These adjustments were:

- Instead of underlining combinations of the seven patterns that the research aimed to investigate and highlighting combinations that learners used the OOCD to look for, they all needed to be extracted out on excel files. This was because in the main study the number of written texts to be handled was many more than in the pilot study, which meant that the number of combinations to be considered was greater. Also, this would facilitate the judgement process of the native speakers. On separate excel files, the informants would have space to give suggestions on how the combinations should be corrected if they were unacceptable. In this way, the combinations which needed to be considered were less likely to be missed.

- The process of tracking down which collocations the learners used the OOCD to look for would be not accurate unless it was done by the learners themselves. Therefore, it was important to ask the learners to highlight collocations looked up from the dictionary in conjunction with completing the recording sheets.

As explained above, the aims of the pilot study were to examine the instruments used to collect data and the procedure of collecting and analysing data, rather than answering the research questions as in the main study; hence, the data collected in this study is not analysed here.

3.3.5 Conclusion

All in all, the examining of the recording sheet, interview questions, and the process of analysing participants' written work in the pilot study have brought me some hands-on experience and have helped me arrive at a general assessment of the instruments as well as the whole research design. Following the result of the pilot study, some adjustments were made to the main study. These adjustments were made

to avoid confusing the participants; hence it was hoped that the study would generate more useful and trustworthy data for an in-depth understanding of the impacts as well as the reasons underlying learners' perceptions of the use of the OOCD as a writing aid.

3.4 Participants

The research took place at the University of Social Sciences and Humanities in Ho Chi Minh City, where I worked from September 2002 to November 2014, so I had relatively easy access to the research site. Participants targeted for this study were students at the Faculty of English Linguistics and Literature. The majority are female. They include 29 second-year English major students at, on average, upper intermediate to advanced level. They all had completed integrated language skills, reading with writing and listening with speaking, in the first three semesters and had passed an English proficiency test designed at around upper intermediate to advanced level.

Students doing the Academic writing module were selected for this research since, as discussed in the context section, this is the last module on writing and at this level students are expected to learn to expand their store of vocabulary themselves rather than waiting to be taught. This three-credit module took place in one session per week over 9 weeks. There was one session, normally in the middle of the course, for the mid-term exam.

In preparation for recruitment, I contacted the teaching support staff via email around three months before the start of the new semester regarding the plan for conducting the research and the recruitment of students. Typically, one month before the start of the new semester, around September and January each year, students register for the modules they wish to take that semester, and they are allowed to choose modules, lecturers, and sessions suitable to their own timetable. I asked a member of the support staff responsible for the timetable to inform students of the possibility of getting involved in the study if they enrolled on the module that I was in charge of. I was aware that lack of knowledge of my project might discourage students from

enrolling on the module, but it would not have been practicable to speak to them in person. Fortunately, one of my colleagues helped me to explain to the students what it would involve to take part in this research. In particular, they were informed that there would be no change in the module content, no further tasks or assignments were required except for those obligatory in the curriculum, and that for those who neither knew about nor had used the OOCD before, taking part in the course might be beneficial for them.

Early on in the design phase of the study, the number of participants I had planned to take was thirty. The actual number of students registering for the module was far more than that, forty one. However, not all were recruited as participants for the research. Those redoing the module as a result of failing the exam in the previous semester needed to be excluded since they were deemed to be below the level that I wanted to investigate. Those redoing the module for the purpose of getting a higher score were also rejected for they were not required to attend all the learning sessions nor to do the assignments, some of which I would use as the data for the study.

The recruitment of participants for the study did not officially start until the first day of the course when students were fully introduced to why the research was being done and what it involved. In total, I had thirty-five participants whose written texts could be used as data for the study, excluding six students redoing the module. In my study, one of the research questions looked at the effects of the use of the OOCD as a supportive tool to aid learners in writing, and as I explained in the Research Methodology chapter, this involved a comparison of the two written texts, without and with the use of the OOCD. This means that for those who only did the first or the second written text, their writings could not be used as data for the study. As such, the actual number of students from whom I collected both written texts was twenty-nine, one less than planned.

3.5   Data collection procedures

After making adjustments to research design and recruiting participants, the data collection procedure started and lasted for three months. What follows is a

description of the detailed activities undertaken to collect data needed to address the research questions.

3.5.1   First set of written texts

At the first class meeting, after the introduction to the module objectives and requirements, I explained to the students what my research was about, why it was being done, and what it involved. The concepts of collocation and its importance in improving writing competence were also raised at the beginning of the course to arouse learners' awareness because, as discussed in the Literature Review chapter, one of the reasons why collocation is problematic to language learners is their lack of collocation awareness (section 2.5.1 studies by Liu (1999) and Wray (2002)), which impedes their improvement in academic writing. In response to my invitation to take part in the study, all thirty-five students present on the day happily agreed to join and signed the Participant Consent Form. Having signed the forms, they were assigned to write a 350-word essay on either of the two topics given within forty-five minutes without the support of any dictionaries (see Appendix 6). The writing texts were the data to address the research question on what kinds of collocation learners could use correctly and what kinds they were struggling with. To ensure that students' names remained anonymous, each student was assigned a number, which they would later on put on their second written work, the recording sheets, and the questionnaires.

I had anticipated that the internet link in the school campus was sometimes not good. Also, for students living in rented houses, internet access was quite limited, and this would to some extent obstruct them in using the dictionary to support their writing assignments. To deal with this, besides offering to install the electronic Oxford Collocation Dictionary on their laptops, I encouraged them to make use of the computers at the English Resource Centre to get access to the internet, stressing that for the final writing they would only be allowed to use the online rather than the installed version of the dictionary, so it would be beneficial for them to get used to the online one. This was the most that I could do to deal with the drawback.

3.5.2   Introducing the dictionary

After the first written texts had been collected, I introduced the OOCD to students. Though dictionaries in general are not a new tool, but rather a 'companion' to

learners right from the very first steps of the language learning path, the review of learners' use of the dictionary from the Literature Review chapter shows (section 2.6.1) that learners are not in the habit of looking up collocations which they can exploit from general dictionaries. With that awareness in mind, I set myself the task of instructing the students to make effective use of the OOCD.

The instruction on dictionary use first involved the introduction of the general structure of the dictionary, what collocational patterns the dictionary displayed, and how those collocational patterns were presented. Learners were then instructed in how to navigate the dictionary to locate particular entries and how to obtain information about usage from examples. They were then required to do some hands-on practice to look for collocations to complete exercises taken from the dictionary. These exercises were to help students familiarize themselves with the dictionary (see Appendix 7). Instructions on note-taking for further vocabulary learning were also given. In particular, I instructed them, as Hill et al. (2000) suggest, to extend what they already know by adding knowledge of collocation restrictions to known words and store them by keeping a notebook or making their own corpus (see Appendix 8). *Medicine* was taken as an example to illustrate how they could extend their collocations. This is supposed to be a known word to students at this level, but the combination of *medicine* when it means *substance taken to treat an illness* with adjectives such as *powerful, cough* or with verbs like *take, swallow, or treat sb with*, deemed known to them, were not necessarily known, and therefore they needed to be noted down in their notebook to learn. *Take medicine* or *swallow medicine* should be paid greater attention since they are susceptible to errors by Vietnamese learners due to L1 influence. The instruction was given with an expectation that towards the end language input would become language output and learners could use those collocations independently without referring to the dictionary. However, this was outside of the scope of this study.

The presentation of collocations in the dictionary is, as discussed in the Literature Review chapter, at the entries of bases. The introduction of which element in the combination is the base and which is the collocate was nevertheless deemed unimportant since, as I have discussed in the Literature Review (section 2.5.3), when

thinking of an idea learners think first of the base. These would be the headwords that learners look up in the dictionary, and what they are looking for is the collocator to complete the phrasal meaning. The base-collocator relationship of N + N, however, seems to be the odd one out. In particular, the first noun in the N + N pattern is the base and the appearance of collocations of this pattern is at the entry of the base, while the second noun in this pattern is the head noun, and learners, specifically Vietnamese learners, tend to think first of the noun head. For instance, the combination *office hours* can only be found at the entry of the noun *office*, not *hours*. Hence, I needed to draw learners' special attention to this collocation pattern to ensure that they would find what they want to look for.

Learners were also instructed to utilize the OOCD in writing practice. For instance, *pay attention* is assumed to be known to learners and instead of using this collocation repeatedly, if they wish to express this same idea again in written work, they could use *devote attention* as found in the OOCD; or *attract someone's attention* could be replaced by *draw/grab/capture/catch someone's attention*. I emphasized that the repeated use of certain words, if not intentionally for some particular effect, would be considered bad style in writing and that though they could be used interchangeably they were not completely the same. They could also maximize the OOCD to write more beautifully and flexibly by adding appropriate collocates to nouns, verbs or adjectives. To illustrate this, I extracted these sentences from the BNC and the OOCD:

a. He merely sipped at his ***strong*** *coffee* and ran a hand through his hair.

b. Birds *sang **cheerfully*** in the trees.

The use of *strong* and *cheerfully* make the sentences sound more vivid, though without them the core meanings might not be lost. Students were also made aware that mis-collocating might cause awkward language, and appropriate choice of collocations could replace this. The following examples were taken as illustration (extracted from TOEFL essays):

a. 'People need jobs to <u>get money for their wives and children</u>.'
   'People need jobs to *earn a living for/support* their families.'

b. 'Factories are helpful because they <u>are places for people looking for a job</u>.'

'Factories are helpful because they *provide job opportunities*.'

Students were assigned one essay every two weeks around three types of essays, as designed in the textbook. This was done with the expectation that they would get used to using this dictionary to look up collocations for their writing task. All their writings were collected and given comments. Before returning their papers, I spent approximately 20 minutes at an appropriate time during the lessons to give feedback on their writing. Besides some feedback on organization, content, coherence and cohesion, collocation errors were pointed out. However, with the awareness that this activity might lead to some collocations being learnt, I limited it to pinpointing collocation errors rather than giving them any suggestions for correction (see Appendix 9).

Table 3.2 below summarizes what was done each week and how the data was scheduled to be collected:

Table 3.2 *Working schedules*

Week	What was done	What was gained
Week 1 (29/02/2016)	- Introducing module objectives and requirements - Introducing the research - Introducing what collocation is - Inviting students to take part in the research and getting the Consent Forms signed - Students doing the first writing within 45 minutes - Giving instructing on how to use the OOCD to support their writing - Giving exercises to help students get familiar with the dictionary	- 35 students signed the Consent Forms - 35 first-written texts (without using dictionary) were collected - Students getting familiar with the dictionary
Week 2 (7/3/2016)	Chapter 1: *Process essays* - Teaching chapter 1 as planned - (Introducing and inviting students absent from the first class meeting to take part in the research at break time) - Homework: exercises to practice using the OOCD to look for collocations	- The other six students signed the consent forms and did the first writing at the end of the session.
Week 3 (14/3/2016)	Chapter 2: *Cause/Effect essays* - Teaching chapter 2 as planned - Homework: write a cause/effect essay of 350 words	

Week	What was done	What was gained
	(using the OOCD as a writing aid)	- Students getting used to the dictionary
Week 4 (21/4/2016)	Chapter 3: *Comparison/Contrast essays* - Teaching chapter 3 as planned - Spending 20 minutes on homework correction	
Week 5 (28/3/2016)	**Mid-term test**	
Week 6 (4/4/2016)	Chapter 3: *Comparison/Contrast essays (cont.)* - Teaching chapter 3 as planned - Homework: write a comparison/contrast essay of 350 words (using the OOCD as a writing aid)	- Students getting used to the dictionary
Week 7 (11/4/2016)	Chapter 4: *Argumentative essays* - Teaching chapter 4 as planned - Spending 20 minutes on homework correction - Homework: write an argumentative essay of 350 words (using the OOCD as a writing aid)	- Students getting used to the dictionary

Week	What was done	What was gained
Week 8 (18/4/2016)	Chapter 4: *Argumentative essays (cont.)* - Teaching chapter 4 as planned - Spending 20 minutes on homework correction	
Week 9 (25/4/2016)	*Review* - Students doing the second writing (using the OOCD as a writing aid) - Giving instructions on how to work in pairs (one doing the writing, the other doing the recording sheet) - Giving instructions on how to fill in the recording sheets - Getting students to answer the questionnaires - Inviting volunteers for the interviews	- 33 written texts (written with the support of the OOCD) were collected. 29 could be used as data for the research. - 33 questionnaires handed out were collected. - 33 recording sheets were collected
Week 9 (26/4/2016)	- 8 students taking part in semi-structured interviews	- All 8 interviews were recorded

3.5.3   Recording sheets and the second written texts

As planned, in the last session students were asked to do the second in-class writing (see Appendix 10). At this stage students were supposed to be familiar with the dictionary. The second 350-word essays were also written in 45 minutes about one of the two topics given, but with the support of the OOCD. Thirty-six participants present on the day were divided into two groups. As described in the research design section, the two groups took turns to do the writing and the recording sheets. They were working in pairs, one partner using the OOCD, the other recording every check-up on the recording sheet. Students were requested to use the specific numbers assigned to them on the first meeting to put on the recording sheets and their written work instead of their names, in order to remain anonymous.

In preparation for the second writing, I had asked all the students to bring their laptops on the day. Fourteen out of thirty-six students present on that day brought with them their own laptops. We borrowed four laptops from the University facility to have enough for half of the students. All these laptops were checked for, and had installed on them if necessary, a Vietnamese-English Dictionary and an English-English dictionary, which could be an Oxford Advanced Learner's Dictionary (OALD), Longman Dictionary of Contemporary English (LDOCE) or Cambridge Advanced Learner's Dictionary (CALD).

Before the start of the writing activity, I gave instructions on how to record the recording sheets carefully to all the students. To resolve the possible drawbacks of this method of observing students' use of the dictionary (mentioned in section 3.2.1), I asked them to approach to OOCD as naturally as possible and emphasized its importance to the results of my study. This, I believe, had a good effect to some extent. For as accurate an assessment as possible of the impact of the dictionary use on students' collocation competence, I attempted to minimize the possibility of collocation improvement due to natural improvement after taking the course, and especially due to the process of using the OOCD to support their homework essays. I did this by encouraging them to redo the dictionary search of words they thought

they had learned from searching the OOCD when doing homework if they needed to use them for this second writing. It was hoped that the recording sheets would reflect the whole picture of how, and in particular how many, collocations had been looked up in the OOCD.

Students completing the recording sheets were asked to record information about headword checkups (column 1) and if their partner used the OOCD in combination with other dictionaries (column 6) on the recording sheets amended after the pilot study. The rest of the information on this sheet was completed by the students doing the writing immediately after they had finished their writing. After completing the recording sheets, students were asked to highlight collocations (in their writing) looked up from the OOCD based on the records. It took around thirty minutes for them to complete both the recording sheets and highlight collocations in their writing. The essays and the recording sheets were then collected together before students swapped their roles. Thirty-six written texts and the same number of recording sheets were collected in this final week.

3.5.4   Questionnaires

After all the participants had completed their writing and recording sheets, they were provided with Likert scale questionnaires, which were designed to measure their attitudes to the use of the OOCD as a tool to assist writing. Paper-based questionnaires, distributed in person, seemed to be the most direct and effective way to collect immediate responses from the participants. I made myself available for answering any questions regarding the content of the questionnaires. It did not take more than 30 minutes for all the participants to complete. All 33 questionnaires distributed were collected. To ensure that I could appeal enough participants for the interviews, I asked them to volunteer to take part and made arrangements for the next meetings immediately after gathering all the questionnaires.

3.5.5   Interviews

In this research, the interviews were conducted after the writing task at the end of the course, in order to make sure that participants had had some time to become familiar with how to look up collocations in the dictionary. Based on their choice of future use of the OOCD on the Likert scale questionnaires (statement 10), I invited eight students for interview. Early in the design phase, I had planned to take 2 students choosing strongly agree (number 1 of the Likert scale), 2 students choosing strongly disagree (number 4), 2 who agreed (number 2) and 2 who disagreed (number 3). This plan to invite students for interview was to gain a broad range of student perspectives on the use of the OOCD to support writing. However, in reality, no student chose 'strongly disagree' and only one chose 'disagree'. As such, I decided to invite one student choosing disagree, four students who agreed, and three students who strongly agreed for the interviews, which were scheduled the following day at the English Resources Centre on the university campus.

After making arrangement for the interviews, I emailed the participants some guiding questions beforehand. In this way, they had time to prepare answers or could make notes on what they wanted to share in the interviews. The interviews were face-to-face and were recorded with the approval of all participants. The aims of the interviews were explained beforehand. One of the most important techniques in successful interviews is to 'build the confidence in the interviewee and establish some trust' (Rubin and Rubin, 2011, p. 114). This was deemed to be easily done since the relationship between the interviewer, myself, and all the interviewees, my students, was defined and well established due to the process of working together. However, if at the previous stages of the data collection process playing the role of a teacher alongside the main role of researcher brought me more advantages than disadvantages, it was at this stage that this could turn out to be a disadvantage. Participants could want to please me, so they might not share with me what they really thought if they did not appreciate or perceive the OOCD as a useful and supportive tool. Once again, in order to gain insight and trustworthy information on how the participants evaluated the use of the OOCD to support their writing, I reiterated and emphasized the importance of their sincerity in answering for the results of my study.

The questions used in the interviews were in English. However, students were encouraged to use either English or Vietnamese, whichever they felt most comfortable and confident with. In five out of eight interviews, students used English, but at some points when they found it hard to express their ideas fully and accurately, they code-switched with Vietnamese. In those instances, I also code-switched between the two languages accordingly to comfort them in sharing information. Double-barreled questions were made use of where possible, in the main or follow-up question list (Rubin and Rubin, 2011 p.113), as an attempt to glean in-depth answers.

At the interviewing venue, in order to help participants recall how they had used the OOCD for checking up as well as what they thought about this tool, I gave them back their recording sheets and prepared a laptop in case they wanted to check up to illustrate what they wanted to convey. As well as recording, I made some notes while interviewing to ensure that all the key information was captured in case the recording was not in good quality or some technical problems were to occur. After the interviews, these recording sheets were re-collected for storage.

Reflecting on the interviews, I saw that although I had preparing carefully and gained training in interviewing skills in the pilot study, I was still inconsistent and unskilled as an interviewer. For instance, though I was fully aware of the principle 'don't finish sentences' since it is a hint to the interviewees that what they have said is obvious (Trochim and Donnelly, 2008, p. 117), there were at least two times when I failed to comply with the rule. This could be an indicator of the spontaneous nature of interviewing. Table 3.3 below presents the interviewing records.

Table 3.3 *Interviewing records*

Respondents	Recoding time
Participant A	28 mins
Participant B	26 mins

Participant C	30 mins
Participant D	27 mins
Participant E	25 mins
Participant F	28 mins
Participant G	27 mins
Participant H	25 mins

3.6   Examining validity

Reliability and validity are tools for measuring the quality of quantitative research. In quantitative paradigms, the former means the replicability of the process and results while the latter refers to how truthful the research results are (Joppe, 2010). In qualitative paradigms, however, they are defined differently. Golafshani (2003), for instance, conceptualizes reliability and validity as trustworthiness, rigor, and quality. Different terms have been developed and used as essential criteria for quality such as credibility, consistency, and transferability (Lincoln and Guba, 1985) since, as Stenbacka (2001, p. 552) puts it, 'the concept of reliability is even misleading in qualitative research'. In this multi-method qualitative-based project, I justify my research methodology using the criteria of validity and transferability. It is true that, regardless of the difference in terminologies used, it is essential for qualitative researchers to be able show that their studies are sound and rigorous (Clarke et al., 2015). Burke (1997) and Clarke et al. (2015) identify a number of validation strategies, four of which have been used in this study:

- *Reflexivity*, one of the key strategies to minimize researcher bias, means that 'the researcher actively engages in critical self-reflection about his or her potential biases and predispositions' (Burke, 1997, p. 283). In awareness of this, I have discussed my personal background and role in the research, how it might affect my research, and what strategies I used to address the potential problems elsewhere in my research report from the beginning to the chapter discussing the findings.

- *Negative case sampling* is also a strategy to minimize the effect of researcher bias. It requires me to carefully and purposefully search for ideas in the transcripts that disconfirm my expectation when coding and identifying themes. Reporting disconfirming cases reassures the reader that I have taken into account and presented all the data rather than the parts that fit my viewpoint. For example, when coding I found that participants used many positive words (e.g. *very helpful, convenient, really good, very useful, time-saving, reliable, beneficial*) to express their satisfaction towards the use of the OOCD as a supportive tool. I was aware that I needed to keep myself vigilant by looking for any words to express the negative feeling and indeed I also found quite a lot of these (e.g. *hinder, not very necessary, limited number, not provide* etc.).

- *Triangulation*, specifically the triangulation of methods, means using several kinds of methods to improve the validity of the research findings. In particular, in this study questionnaires and interviews have been used to explore learners' perception of the use of the OOCD as an assisting tool, how many look-ups they made and how they used the OOCD to search for collocations. Findings obtained from these methods were triangulated with those obtained from the recording sheets recording specifically which collocations were searched for, how they were searched for and whether those look-ups were successful.

- *Low inference descriptors*, of which verbatim is the lowest, require the researcher to stay as close as possible to participants' accounts. This strategy enhances interpretive validity and ensures that the reader can accurately understand and 'experience the participants' actual language, dialect and personal meaning' (Burke, 1997, p. 285). Given this, where appropriate in the writing-up process, I reported findings from the interviews by using participants' exact words in direct quotations.

For many qualitative studies, generalizability and transferability could be a drawback since findings are usually obtained from a small amount of cases. By combining two types of data, quantitative and qualitative, on a group of 33 participants, it is hoped that this qualitative-based research might overcome some of these limitations.

Findings from the study can help me to come up with some generalizations regarding what collocations Vietnamese learners can use correctly, what collocations they are struggling with, and whether or not the dictionary helps learners improve their collocation use. Accordingly, the matter here was how to enhance the quality of data.

With regard to the questionnaires, the interviews and the recording sheets, I attempted to reduce to the minimum my influence as a tutor on their responses. This involved my informing and reminding them several times how valuable their true responses and their normal approach to the dictionary were to the success of my research. This role seemed to enable me to establish a good rapport with the participants for in-depth interviews. Two of the participants called me a day after the interviews to add some information that they had forgotten to share in the interviews. The information was important in helping me make sense of why they perceived the dictionary the way they did. Regarding the written texts (written with and without intervention), their scores were added up to the final scores of the module, so participants were assumed to have tried their best in completing them.

3.7 Ethical considerations

This study involved human participants and made use of voice recording during individual interviews. It therefore required official ethical approval. An ethical application was submitted to the University of Leeds Social Sciences Research Ethics Committee, and the study was approved. Information sheets with details of research aims, activities, and interviews that were carried out in a specific class and with some students had been given to the head of the university, the dean of the department, and the students taking part in the module. Participants were informed clearly what treatment they would receive from the beginning and were made aware of their right to withdraw at any point of the research.

Audio recording with the participants chosen for the interviews was done with their permission. They were invited to review the transcript of the interviews. This was an opportunity for the participants to ensure that they represented what they wanted to say, and to add to or change the interview information if they wished to.

With regard to anonymity issues, the participants remained anonymous. They were replaced by alphabetical letters in the interview transcripts. All of the recording sheets, questionnaires, and writing were scanned and stored on the University of Leeds M-drive. They could be accessed through my university account externally via Desktop Anywhere.

Chapter 4 : DATA ANALYSIS

4.1   Introduction

In the previous chapter I presented the research design and instruments of the study, how the pilot study was carried out, and the process of collecting different data sets. Validity was examined and ethical considerations were discussed.

In this chapter, I will first present the procedure for extracting and analysing collocations extracted from the two sets of written texts. In the following section, section 4.3, I will discuss how I analysed the variety of collocation use. How I analysed quantitative and qualitative data collected through questionnaires, recording sheets and interviews will be then addressed correspondingly in the last three sections.

4.2   Procedure for extracting and analyzing collocations

In order to answer the research question about how advanced learners use collocations, I analyzed the first pieces of writing written without the support of the OOCD and assessed the appropriateness and variety of collocations. All the written texts of this first set were labeled A, and the second set B (e.g. T1A, T1B). The process of analyzing collocations in students' writing involved four main steps: extracting all lexical combinations of the grammatical patterns discussed in the Literature Review, assessing the conventionality of the combinations by using the BNC, identifying which of these are strong collocations, and evaluating those collocations. Table 4.1 below reminds the reader of the grammatical patterns of lexical combinations that I was investigating. The first column presents the bases of combinations, the second the grammatical patterns of lexical collocations, and the last column is the correspondent examples of the patterns.

Table 4.1 *Grammatical patterns of collocations*

The base of the combinations	Grammatical patterns	Examples

Nouns	Adj + N	*Strong light*
	N (quantifier) + of + N	*A beam of light*
	V + N	*To shed light*
	N + V	*Light gleams*
	N + N	*A light source*
Verbs	Adv + V/ V + Adv	*Choose carefully*
Adjectives	Adv + Adj	*Entirely safe*

In what follows, I will first outline the procedure of extracting, identifying and evaluating combinations of the above patterns. I will then describe in detail, and give examples to illustrate, what was done at each step. Following this, I will present the analysis of two written texts chosen randomly (T12A, T22A).

Procedure

1. Following what I have discussed in the Literature Review chapter, this extracting step of the lexical combinations of these patterns was done as follows:

   a. Identify components of finite and non-finite clauses (Subject, Verb, Object, Complement, and Adjunct) of a sentence
      - Identify and extract combinations of V-N, N-V, and Adv-V pattern.
   b. Identify Noun Phrases in each sentence
      - Look for the noun head of each Noun Phrase, the base of all lexical combinations that contain nouns.
      - Look for adjectives that together with the nouns constitute the Adj-N pattern
      - Identify components of the N-N pattern and the N-of-N pattern
      - When the noun head is post-modified by a clause, finite or non-finite, N-V (the noun head and the verb in the post-modifying clause) will be extracted if the noun is the doer or agent of the action.
   c. Identify Adjective Phrases in the sentence

- Look for adjective head of each adjective phrase, the base of the Adv-Adj combination
- Look for adverbs modifying the adjective head

2. Assess the conventionality of the above combinations by using the BNC

- Identify the contextual meaning of the base (headword) in the context (the text)

a. If the combinations met the frequency threshold of five, they were considered to be conventional and would be processed in the next step.

b. If the combinations did not meet the frequency threshold, I asked native speakers for their judgements on their acceptability

- If the combinations were judged to be acceptable, the combinations would be considered further in the next step.

- If the combinations were judged to be unacceptable, the combinations would be treated as collocation errors.

- If the combinations were questionable, they were also considered collocation errors.

3. Distinguish strong collocations from casual combinations and idioms

- Based on the meaning of the base identified from the previous step and the transparency criterion, distinguishing strong collocations from idioms.

- Look for Log Dice score of the combinations from the BNC to identify strong collocations.

4. Evaluate the use of collocations

I only used descriptive statistics to describe what was going on in my data but not inferential statistics since the sample population is quite small (29 participants) and they are not a random sample representing the general population due to some restrictions on the recruitment of the participants for the research. Below is a detailed description of what was done at each step.

Step 1: Extracting lexical combinations of the grammatical patterns

Before extracting combinations of these patterns, I needed to deal with grammatically odd or misspelling cases, and this was done in conjunction with the process of identifying and analyzing components of the sentence. The focus of the

study is on whether or not the combination of words sounds native-like and the appropriate use of those combinations, so inflectional morphological errors (e.g. *the student face* instead of *the student faces*), article errors (e.g. *have chance* instead of *have a/the chance*) were disregarded. Minor spelling mistakes were also disregarded if it was clear which word the students intended to use. Take, for instance, the misspelled adverb *comfortbly* in the sentence *For those who are not independent, they cannot live comfortbly* (T13A). From its orthography and the context, there can be little or no doubt that *comfortably* is the word intended. These errors were corrected before combinations were extracted for examination for two reasons: 1) the search of misspelled words against the BNC would not bring an accurate result of frequency of occurrence of those words with collocates being considered; 2) this is an attempt to avoid making native informants confused when they had to make a decision on whether the combinations were acceptable, unacceptable, or questionable.

Cases in which the syntactic relation between elements of the patterns was unclear were still extracted. This was an attempt to avoid missing any errors due to failure of collocation use. As for cases which were syntactically wrong and apparently did not fit into any of the above categories (e.g. *abroad students*), I treated them as collocational errors if the intended meaning (inferred from the context) could be best expressed by an established collocation (e.g. *overseas students*). Any other minor mistakes would be considered in detail while I was dealing with the sentence.

All lexical combinations appearing in the seven patterns that are the focus of this study were then extracted. This process of extracting combinations was carried out sentence by sentence; that is, I examined and extracted all combinations of the above patterns occurring in one sentence before moving to the next. The following is an illustration of how this first step was carried out on written text T12A. The written text was typed and numbered for ease of presentation and analysis.

1. *To many students, / having chances to study abroad / is / the chance of life.*
    A   S   V   C
   - *having / chances to study abroad*

V   O

- to study / abroad

 V Adv

+ Sentence (1) has no errors in spelling and is syntactically correct.

+ *Is* is the copular finite verb because it signals tense.

+ Non-finite clause *Having ... abroad* is the subject because it occurs before the verb and the verb *is* agrees with the head of the Noun Phrase (*having*).

+ The Noun Phrase *the chance of life* is Subject Complement required by the copula.

+ The non-finite clause Subject contains an embedded to-infinitive clause, and was analysed separately.

+ *To many students* is an adjunct since this constituent adds further information and the sentence is still grammatical if it is omitted.

Noun Phrases and Adjective Phrases were then identified and presented in the first column of the table below. The heads are in bold. Combinations extracted from these phrases and clauses were presented in the second column of the table.

Table 4.2 *Text 12A sentence 1*

Noun & Adj Phrases (head in bold)	Combinations extracted
Many **students** **Chances** *to study abroad* *The* **chance** *of life*	 *The chance of life* *(N-of-N)*
	Combinations extracted (from clauses)
	Having chances *(N-V)* *Study abroad* *(V-Adv)*

Based on the same procedure as for the first sentence, I extracted combinations of the remaining sentences.

2. There / are / many benefits that make them want to study in foreign countries.

> S   V   Cs
> - *That / make / them / want to study in foreign countries*
>
>   *S   V   Od   Co*
>
> - *to study / in foreign countries*
>
>   *V   Cloc*

The relative pronoun *That* replaces the noun *benefits*, so *benefits + make* (N-V) was extracted. *Them* is a pronoun anaphorically referring to the noun *students* in sentence (1) and is the implicit Subject of the bare-infinitive clause *want to study in foreign countries*.

Table 4.3 *Text 12A sentence 2*

Noun & Adj Phrases (head in bold)	Combinations extracted	
Many **benefits** *that make them want to study in foreign countries* *Foreign* **countries** **Foreign**	*Benefits (that) make* *Foreign countries*	*(N-V)* *(Adj-N)*
	Combinations extracted (from clauses)	
	Make them (students) *Them (students) want* *Study in (foreign) countries*	*(V-N)* *(N-V)* *(V-N)*

3. However, / in my opinion, / they / also / have to deal with / a numerous problems
      A   A   S   A   V   Op

  such as lack of family support, culture shock and financial issues.

They is also an anaphoric referent of students in sentence (1).

Table 4.4 *Text 12A sentence 3*

N & Adj Phrases (head in bold)	Combinations extracted	
My **opinion** *A numerous* **problems** *…issues*	 *Numerous problems*	 *(Adj-N)*

Lack *of family support*	*Lack of (family) support* *(N-of-N)*
Family **support**	*Family support* *(N-N)*
Culture **shock**	*Culture shock* *(N-N)*
Financial **issues**	*Financial issues* *(Adj-N)*
Numerous	
financial	
	Combinations extracted (from clauses)
	They (students) (have to) deal with (N-V) *Deal with (a numerous) problems (V-N)*

4. *First, / being far away from home / means / lack of family support.*
   A   *S*   *V*   *Cs*

   - *being / far away from home*
    V   *Cloc*

Table 4.5 *Text 12A sentence 4*

N & Adj Phrases (head in bold)	Combinations extracted
Home **Lack** *of family support* *Family* **support**	 *Lack of (family) support* *(N-of-N)* *Family support* *(N-N)*
	Combinations extracted (from clauses)
	Means lack of support *(V-N)*

5. *It / is / easy to realize that when a student goes to another country to pursuit his*
   S *V*   *Cs*

 learning career, he has to be ready to work as chief cook to bottle washer.

   - *To realize / that when a student goes to another country to pursuit*
     his...washer.
     V   *O*

- *When / a student / goes / to another country / to pursuit his learning career*
  A   S   V   Cloc   A

- *to pursuit / his learning career*
    V   O

- *he / has to be / ready to work as chief cook to bottle washer*
   S   V   Cs

- *to work / as chief cook to bottle washer*
    V   A

In this sentence *pursuit* is a noun and was misused as a verb (*pursue*). However, this is considered a minor error and is therefore ignored. *A student* is an implicit Subject of the to-infinitive verb *to pursue*. The combination *student + pursue* was not extracted for the reason that *to pursue* is part of an Adjunct, which is not a required element. The pronoun *he* refers to *a student* occurring before it.

Table 4.6 *Text 12A sentence 5*

N & Adj Phrases (head in bold)	Combinations extracted	
A **student**		
Another **country**		
His learning **career**	*Learning career*	*(N-N)*
Chief **cook**	*Chief cook*	*(N-N)*
Bottle **washer**	*Bottle washer*	*(N-N)*
Easy *to … bottle washer*		
Ready *to work … washer*		
	Combinations extracted (from clauses)	
	Student goes	*(N-V)*
	Goes to (another) country	*(V-N)*
	Pursue (his learning) career	*(V-N)*

6. *His family / can / only / support / him / mainly by encouraging.*
      S   mod   A   V   O   A

- *Encouraging (V)*

+ Pronoun *him* refers to *a student* in sentence (5)

Table 4.7 *Text 12A sentence 6*

N & Adj Phrases (head in bold)	Combinations extracted
*His **family***	
	Combinations extracted (from clauses)
	Family ... support (N-V) *Support him (student)* (V-N)

7. *As a result,/ it / would be / a hard time for a student / when having health*

    A   S   V   Cs   A

 problems and being alone in a foreign country.

 -when / having / health problems // and /being / alone / in a foreign country

   A   V   O   +   V   Cs   A

Table 4.8 *Text 12A sentence 7*

N & Adj Phrases (head in bold)	Combinations extracted
*A **result*** *A hard **time** ... country* *A **student*** *Health **problems*** *A foreign **country***	 *Hard time* (Adj-N) *Health problems* (N-N) *Foreign country* (Adj-N)
	Combinations extracted (from clauses)
	Having health problems (V-N)

8. *Second, / culture shock / is / always / a nightmare / to every foreign students.*

A	*S*	*V*	*A*	*Cs*	*A*

The use of the plural form *students* after *every* is incorrect but was ignored.

Table 4.9 *Text 12A sentence 8*

Noun & Adj Phrases (head in bold)	Combinations extracted
*Culture **shock*** *A **nightmare*** *Every foreign **students*** *Foreign **students***	*Culture shock* (N-N) *Foreign students* (Adj-N)
	Combinations extracted (from clauses)

9. *They / have to face / a lot of differences in various fields such as lifestyle,*

   S   *V*   *O*

  education system and health care system.

Pronoun *they* refers to *students* in the previous sentence.

Table 4.10 *Text 12A sentence 9*

Noun & Adj Phrases (head in bold)	Combinations extracted
*A lot of **differences** … system* *Various **field*** *Education **system*** *Health care **system***	 *Various field* (Adj-N) *Education system* (Adj-N) *Health care system* (N-N)
	Combinations extracted (from clauses)
	They (students) … face (N-V) *Face … differences* (V-N)

10.   *It / takes / time and efforts to integrate.*

       S   *V*   *O + O*

- *To integrate (V)*

It is the dummy Subject and does not refer to any noun occurring before.

Table 4.11 *Text 12A sentence 10*

Noun & Adj Phrases (head in bold)	Combinations extracted
Time *efforts*	
	Combinations extracted (from clauses)
	Takes time (V-N) *Takes efforts (V-N)*

11. *Moreover, / students / do not / only / suffer / culture shock // when / they / are*
                S   Vaux   A   V   O   A   S   V

      studying / in a foreign country // but also / after / they / come back / to their
      hometown.
               A   +   A   S   V   A

Both of the pronoun *they* refer to the subject *students* in the main clause of the sentence.

Table 4.12 *Text 12A sentence 11*

Noun & Adj Phrases (head in bold)	Combinations extracted
Students *Culture **shock*** *A foreign **country*** *Their **hometown*** *foreign*	*Culture shock (N-N)* *Foreign country (Adj-N)*
	Combinations extracted (from clauses)

	Students ... suffer (N-V)
	Suffer culture shock (V-N)
	They (students) studying (N-V)
	They (students) come back (N-V)

12. *Those / create / a big challenge / for a <u>student</u> / to apply what they have*
    S   V   Od   Op   A
    learned abroad to their countries, reality.

  - *To apply / what they have learned abroad / to their countries, reality*
    V   O   Co

    - what / they / have learned / abroad
       O   S   V   A

The use of the singular noun *a student* as a Prepositional Object is not appropriate since in this sentence it refers to students in general rather than a particular student, and the pronouns *they* and *their* refer to these students. However, such an error does not affect to the extraction of combinations. *Those* refers to all difficulties that students have to face as mentioned in the previous sentences and is not a personal pronoun, so it was not extracted.

Table 4.13 *Text 12A sentence 12*

Noun & Adj Phrases (head in bold)	Combinations extracted
A big **challenge** *A* **student** *Their* **countries** **Reality** **big**	*Big challenge* (Adj-N)
	Combinations extracted (from clauses)
	Create a challenge for a student (V-N) *They (students) learned* (N-V) *Learn abroad* (V-Adv)

13.   *Lastly, / financial issues / usually / get on / foreign students' nerves.*

    A   *S*   *A*   *V*   *O*

Table 4.14 *Text 12A sentence 13*

Noun & Adj Phrases (head in bold)	Combinations extracted	
*Financial **issues***	*Financial issues*	*(Adj-N)*
*Foreign students' **nerves***	*Students' nerves*	*(N-N)*
*Foreign **students***	*Foreign students*	*(Adj-N)*
Financial		
foreign		
	Combinations extracted (from clauses)	
	Issues get on	*(N-V)*
	Get on (students') nerves	*(V-N)*

14.   *Many students / may have / scholarship / to support them.*

      S   *V*   *O*   *A*

   -to support / them

    V   *O*

The pronoun *them* refers to the Subject *many students*. *Many students* is also an implicit Subject of the to-infinitive verb *to support*; however, as discussed above the combination was not extracted when the verb is part of the Adjunct.

Table 4.15 *Text 12A sentence 14*

N & Adj Phrases (head in bold)	Combinations extracted	
*Many **students***		
Scholarship		
	Combinations extracted (from clauses)	
	Students have	*(N-V)*
	Have scholarship	*(V-N)*
	Support them (students)	*(V-N)*

15.   However, / daily expense / <u>cost /</u> a great deal of money, / especially in the
        Conj   S   V   O   A

      case of a student from a developing country <u>come</u> to a developed one.

  -   Come to / a developed one
        V   Cloc

The use of the verb *come* as a bare infinitive verb and *cost* not in agreement with the Subject here is incorrect; however, as I have discussed earlier in this section, those grammatical errors would be corrected before combinations were extracted.

Table 4.16 *Text 12A sentence 15*

N & Adj Phrases (head in bold)	Combinations extracted	
Daily **expense**	Daily expense	(Adj-N)
A great deal of **money**	(a great) deal of money	(N-of-N)
A great **deal**	A great deal	(Adj-N)
In the **case** of ... one		
A **student** from ... one		
A developing **country**	Developing country	(Adj-N)
A developed **one** (**country**)	Developed country	(Adj-N)
Daily		
Great		
Developing		
developed		
	Combinations extracted (from clauses)	
	(daily) expense cost	(N-V)
	Costs ... money	(V-N)
	Coming to (a developed) country	(V-N)

16.   As a result,/ foreign students / have to find / part-time jobs / in order to earn their living.
            A   S   V   O   A

- to earn / their living

      V   *O*

Table 4.17 *Text 12A sentence 16*

N & Adj Phrases (head in bold)	Combinations extracted	
*A **result*** *Foreign **students*** *Part-time **jobs*** *Their **living*** ***Foreign*** ***Part-time***	*Foreign students* *(Adj-N)* *Part-time jobs* *(Adj-N)*	
	Combinations extracted (from clauses)	
	Students … find *(N-V)* *Find … jobs* *(V-N)* *Earn … living* *(V-N)*	

17.   *It*   */ requires / great efforts / to balance studying and working.*

    Dummy S   *V*   *O*   *S*

 - to balance / studying and working

    V   *O*   *+ O*

The pronoun *It* is a dummy Subject, so the combination of *it* + *requires* was not considered.

Table 4.18 *Text 12A sentence 17*

N & Adj Phrases (head in bold)	Combinations extracted	
*Great **efforts*** ***Studying*** ***working***	*Great efforts* *(Adj-N)*	
	Combinations extracted (from clauses)	

	It requires	(pro-V)
	Require ... efforts	(V-N)
	Balance studying	(V-N)
	Balance ... working	(V-N)

18. In conclusion, / studying abroad / is not / as easy as a piece of cake.

    A   *S*   *V*   *Cs*

    - *studying / abroad*

      V   *Adv*

Table 4.19 *Text 12A sentence 18*

N & Adj Phrases (head in bold)	Combinations extracted	
Conclusion *studying* abroad a piece of *cake* *easy*	*Piece of cake*	*(N-of-N)*
	Combinations extracted (from clauses)	
	Studying abroad	*(V-Adv)*

19.   *It*   / takes / time, efforts and money.

    Dummy S   *V*   *O*

Table 4.20 *Text 12A sentence 19*

N & Adj Phrases (head in bold)	Combinations extracted	
Time *Efforts* *money*		
	Combinations extracted (from clauses)	
	Takes time	*(V-N)*
	Takes efforts	*(V-N)*
	Takes money	*(V-N)*

20.   *However, / it   / is not / an impossible task.*

      A   dummy S   V   Cs

Table 4.21 *Text 12A sentence 20*

N & Adj Phrases (head in bold)	Combinations extracted
An impossible **task** **impossible**	*Impossible task (Adj-N)*

21.   *If / students / have / enough discipline, skills, and patience, // studying abroad*

      A   S   V   O   S

      / would be / an interesting experience.

      V   O

  -   *Studying / abroad*

      V   Adv

Table 4.22 *Text 12A sentence 21*

N & Adj Phrases (head in bold)	Combinations extracted
Students **Discipline** **Skills** **Patience** *An interesting* **experience**	*Interesting experience (Adj-N)*
	Combinations extracted (from clauses)
	Students have (N-V) *Have discipline (V-N)* *Have skills (V-N)* *Have patience (V-N)* *Studying abroad (V-Adv)*

After combinations of the seven patterns had been extracted, I transferred them into an Excel file. I also kept a record of the number of times the same combinations occurred for later comparison, in order to view the variety of collocation use in students' texts. All combinations extracted from the text continued to be processed according to the following steps.

Step 2: Determining the conventionality of the combinations

This second step involved searching the BNC to decide whether or not a combination occurs. In other words, all the lexical combinations were checked against the BNC to determine their degree of conventionality. The five times threshold from Nesselhauf (2003) was adopted in this study. A window of five tokens to the left and five tokens to the right of the base was applied. Inflections of the same word were all counted (e.g. *stand/stands/stood/standing a chance*) since the core meaning of the combination (*stand + a chance*) did not change. Before searching for the occurrence of the combinations in the BNC, I looked for the specific meanings of the bases of the combinations because, as mentioned in the Literature Review chapter, these meanings were needed to differentiate collocations from idioms. Combinations checked against the BNC were then processed as follows.

Lexical combinations met the frequency threshold (from the BNC search)

Combinations which were found in the BNC search (threshold of five) continued to be processed in the next step (step 3), which aimed to distinguish strong collocations from casual combinations and idioms.

Lexical combinations did not meet the frequency threshold (from the BNC search)

Combinations that did not occur or occurred fewer than five times in the BNC were put through another test for their acceptability, since though the BNC corpus, in general, includes a wide range of texts from different registers and language domains, it evidently did not include all acceptable combinations. As I discussed earlier (section 2.2.4), the approach to judging the acceptability of the combinations was also taken from Nesselhauf's study. In particular, I presented these combinations to native speakers for further judgements on their acceptability. There were two pairs of native speakers involved in the study, one British and one American in each pair. The British informants that helped me with the judgement of these combinations

were a Masters student at the of School of Education and a British lawyer who had just finished his PhD thesis from the School of Law. The American informants included a lecturer at University of California at Berkeley and an English teacher working at a language centre in Viet Nam.

The first pair of informants judged these combinations on a three point scale: acceptable (+), questionable (?), unacceptable (-). Their judgements constituted the final judgements if they were the same. If their judgements were not the same, the other pair of native speakers was asked to make judgements. Table 4.23 below, taken from Nesselhauf (2005, p. 52), presents how I made decisions on the acceptability of combinations.

Table 4.23 *Acceptability judgements*

Judgement 1	Judgement 2	Judgement 3	Judgement 4	Ultimate judgement
+	+			+
+	+	+	?	+
+	+	+	*	(+) largely acceptable
+	+	?	?	(+)
+	+	?	*	(+)
+	+	*	*	?
+	?	?	?	?
+	?	?	*	?
?	?			?
?	?	?	?	?
?	?	?	*	?
+	?	*	*	(*)largely unacceptable

+	*	*	*	(*)
?	?	*	*	(*)
?	*	*	*	*
*	*			*
*	*	*	*	*

All combinations presented to native speakers were presented in particular contexts; that is, they were sentences in which the combinations occurred. Sometimes one previous or following sentence also needed to be given. In this way, the native speakers would have a clearer understanding about the intended meaning that the student writer had wanted to convey before making a judgement or suggesting corrections. If it was judged unacceptable or largely unacceptable, I decided that the combination was non-native and concluded that it was an error in collocation use. All acceptable and largely acceptable combinations (judged so by native speakers) were examined further in the next step. Combinations judged questionable were also determined to be collocational errors since they were not conventional and to native speakers they sounded slightly awkward. For example, *advanced way* as in (T27A), *Student who study abroad can approach to an advanced way of learning*, was judged questionable by both informants, who suggested it be changed to *advanced form.* Towards native-like language use, combinations like this should also be avoided.

Step 3: Distinguishing strong collocations from casual combinations and idioms
After assessing the conventionality of the combinations extracted from the writing texts, I sought to separate strong collocations from other combinations, casual combinations and idioms. As discussed in the Literature Review chapter (section 2.2.1), if the meaning of the base was a primary meaning found in the dictionary, the combination was judged not to be an idiom. The transparency criterion was only applied to see if the combination as a whole is transparent in meaning when the base does not carry a particular meaning in a general English-English dictionary. For all those combinations identified as not-idioms, I looked for their Log Dice score in the BNC. The Log Dice score differs from the standard frequency search because the meaning and syntactic relations between elements are not accounted for. However,

since the search for Log Dice score between elements with some particular meaning and syntactic relation was not doable from the BNC search and more importantly the purpose of this search was to look for the association measure between these elements rather than the conventionality of the combination, the Log Dice score between these elements regardless of their meaning and syntactic relations was still used. If the Log Dice score was above the threshold of 4.0 (discussed in section 2.2.3), the combination was a strong collocation. If it was below this threshold, it was a casual combination.

Step 4: Evaluating the use of strong collocations
After identifying strong collocations from the written texts, I examined them closely to see whether they were correctly used, both syntactically and semantically.

Combinations extracted from the two texts
The search for the co-occurrence of elements of these combinations was done by using the filtering function of Sketch Engine (Kilgarriff, 2004). However, as discussed in the procedure section, before searching for the 'existence' of the combinations, I needed to check the meaning of the base of the combination. Once this was found, I started by searching the base of the combinations in the BNC. It did not make any difference to start searching from the base or the collocate when filtering. Take, for example, the combination *have chance* (V-N) of which the noun *chance* is the base. OALD search for the meaning of *chance* showed that *chance* means *a possibility of something happening especially something that you want* in the context. In Sketch Engine the word *Chance* occurred more than 15,800 times. I then filtered the number of times *chance* co-occurs with *have* within the span of 5 to the left and the right of the base. This showed more than 3,900 co-occurrences. This number, however, did not tell much since *have* and *chance* do not have the rigid syntactic relation (verb-noun) that I was considering every time they co-occur. What I needed to do next was to scan and count manually to see if their valid co-occurrence was above the threshold of five. Given the flaw of the BNC in presenting results of the same file types close to each other, I requested the results of 500 random samples of the combination being considered. Below is an example of a concordance page of *have* and *chance* (Figure 4.1). Of the 20 times they co-occur, only 10 times were counted as a valid co-occurrence; the others (as in lines 2, 4, 7, 11, 14, 15, 16, 18, 19, 20) were not taken into account. This is because, as in lines 2

and 4, *have* is not a main verb but an auxiliary verb, and *chance* is not an object of *have*. In other words, they do not constitute the verb-object relation as *have a chance* does. *Have* and *chance* meet the co-occurrence threshold, so it was then identified as conventional.

Figure 4.1 *A random concordance page of have and chance*

```
      I gave my first recital, and also had a chance  of playing the Beethoven Concerto No. 4
       : 'Children have affected life in taking chances of my own away. Working full-time and their
      up clean and quick and before you had a chance  to believe it was happening the barrel
  something from his desk. That gave me the  chance  to tell what Braden had done, sounding
  McLaren girl, staring right at her, so I had a chance  to look at Mrs. Favor. Audra was her name
  would mean that marine life would have no chance  of escape, while the natural dispersion
              sold openly. This, they say, is the only chance  the rhino has of surviving in the wild </p>
words vary enormously, and we shall have a chance  to hear some of the tunes, certainly. Sometimes
     what would they do to us if they have the  chance  ? There are suggestions talking and writing
  them. Thanks Okay, hopefully you 've had a chance  to make, to make notes on that. I was I
      quality as well. You have that choice, that chance  to do that for my patients and I tell you
    commercial farmers even when it had the  chance  to do so. (Mostly it had not, since it
          outside the city. If they rarely had much  chance  in the more skilled manual jobs, unless
  etc. would be resolved and so increase the chances of having the complete story. The information
  well, it all went to show we 'd stumbled by chance  on an opium-den. Only the front man kept
      in Brighton, and didn't being to take the chance  that had been offered to him. A great shame
      aware of what's going on, doesn't have a chance  to meet them. Is it possible when we get
  'm gonna ask some questions but there's a chance  to have a look through Oh spent a bit of
          voice? Yes. If it is you have n't got a chance  in hell. I'm a bit hoarse. Mm. Yeah. I
          that if somebody says, Look is there any chance  I can have colour, that's the only place
```

With regard to combinations of three or more elements, such as *keep stress under control*, I needed to look for the frequency of co-occurrence of the collocate *keep* with individual nouns accompanying it, since the search for the combinations as a whole often bring up a very small amount of co-occurrence in the BNC. *Keep + under control* occurs 139 times and *keep + stress* occurs 29 times, while the search of *keep + stress under control* does not appear in the BNC.

Potential strong collocations from the two written texts (T12A, T22A), written without the support of the OOCD, were extracted and presented in the tables below. The last two columns in the tables present the amount of occurrence of the combinations in the texts and the results of the BNC check for their conventionality, respectively. '>5' means the combination occurs more than five times in the BNC

and is therefore treated as conventional. Specific numbers indicate the amount of occurrence of the combination in the BNC, and they are lower than the threshold set.

Table 4.24 *Potential strong collocations of text 12A*

N	Combinations	Occurrence in the text	BNC check
I	V-N		
1	having chance	1	>5
2	to deal with (numerous) problems	1	>5
3	goes to (another) country	1	>5
4	pursue (his learning) career	1	>5
5	having … problems	1	>5
6	face (a lot of) differences in	1	0
7	takes time	2	>5
8	takes … effort	2	>5
9	suffer (culture) shock	1	>5
10	studying in (a foreign) country	1	>5
11	create (a big) challenge for a student	1	>5
12	get on … (students') nerves	1	>5
13	have scholarship	1	>5
14	cost (a great deal of) money	1	>5
15	earn (their) living	1	>5
16	require (great) effort	1	>5
17	takes money	1	>5
18	have discipline	1	>5
19	Have … skills	1	>5
20	have … patience	1	>5
21	balance work	1	>5

N	Combinations	Occurrence in the text	BNC check
22	balance studying	1	0
23	come to (a developed) country	1	>5
24	support him/them (student/s)	1	>5
25	find … jobs	1	>5
26	make them (student)	1	>5
27	means lack of support	1	0
28	come back to (their) hometown	1	0
II	**N-V**		
29	benefits (that) make	1	>5
30	a student goes (to another country)	1	>5
31	a student … to pursue	1	>5
32	his family … support	1	>5
33	students … suffer	1	>5
34	students (may) have	2	>5
35	a student … come (to a developed one)	1	>5
36	issues get on	1	0
37	they (students) … deal with	1	>5
38	they (students) … face	1	>5
39	they (students) … come back	1	>5
40	they (students) learned	1	>5
41	they (students) ... studying	1	>5
42	students (have to) find (jobs)	1	>5
III	**Adj-N**		
43	numerous problems	1	>5
44	financial issues	2	>5

N	Combinations	Occurrence in the text	BNC check
45	a hard time	1	>5
46	foreign students	3	>5
47	various field	1	>5
48	a big challenge	1	>5
49	daily expense	1	2
50	a great deal		>5
51	developing country		>5
52	part-time job		>5
53	great effort		>5
54	foreign countries (used in the question)		>5
55	developed countries (used in the question)		>5
56	impossible task		>5
57	interesting experience		>5
IV	**N-N**		
58	family support	2	>5
59	culture shock	3	>5
60	learning career	1	1
61	chief cook	1	3
62	bottle washer	1	2
63	health problem	1	>5
64	education system	1	>5
65	health care system	1	>5
66	students' nerves	1	0
67	expense cost	1	0
V	**N-of-N**		

N	Combinations	Occurrence in the text	BNC check
68	the chance of life	1	>5
69	lack of … support	2	>5
70	a great deal of money	1	>5
71	the case of a student	1	1
72	a piece of cake	idiom	>5
VI	**V-Adv**		
73	learned abroad	1	2
74	studying abroad (used in the question)	1	>5
VII	**Adv-Adj**		

Table 4.25 *Potential strong collocations of text 22A*

N	Combinations	Occurrence in the text	BNC check
I	**V-N**		
1	take courses	1	>5
2	give students some disadvantages	2	0
3	face problem	1	>5
4	starting a (new) life	1	>5
5	communicating with (local) people	1	>5
6	have difficulty	1	>5
7	adapt to the (new) culture	1	3
8	put students under pressure	1	>5
9	cause depression	1	>5
10	commit suicide	1	>5
11	put up with (their negative) feelings	1	0
12	have (many financial) issues	1	0

N	Combinations	Occurrence in the text	BNC check
13	save money	1	>5
14	have no choice	1	>5
15	pay (their high) tuition fee	1	>5
16	affect (their) study	1	>5
17	earn money	1	>5
18	neglect (their) study	1	1
19	provide students (many) benefits	1	>5
20	have (a good) preparation	1	>5
21	avoid disadvantages	1	>5
22	eat something (pro)	1	>5
24	benefit from studying	1	>5
II	**N-V**		
25	people face	1	>5
26	these students (may) have	4	>5
27	students (will) feel	1	>5
28	students commit	1	2
29	students tend (to have)	1	>5
30	students (must) save	1	>5
31	students … to work	1	>5
32	students … to pay	1	>5
33	tuition fee (which may) affect	1	0
34	students (just) focus	1	>5
35	Students … take	1	>5
36	they (people) live	1	>5
37	they (people) study	1	>5

N	Combinations	Occurrence in the text	BNC check
38	they (students) … adapt to	1	>5
39	they (students) … pay	1	>5
40	they (students) … put up with	1	>5
41	they (students) eat	1	>5
42	they (students) force	1	>5
43	they (students) ... study	1	>5
44	they (students) … avoid	1	>5
45	they (students) benefit	1	>5
46	students (are) faced with	1	>5
47	students … neglect	1	0
48	marks get	1	1
III	**Adj-N**		
48	foreign countries (used in the question)	1	
49	the high cost	3	>5
50	the common problem	1	>5
51	a new life	1	>5
52	a new country	1	>5
53	real challenges	1	>5
54	local people	1	>5
55	a (completely) strange country	1	>5
56	the new culture	1	>5
57	developing countries	1	>5
58	financial issues	1	>5
59	developed countries (used in the question)	1	>5
60	daily expenses	1	1

N	Combinations	Occurrence in the text	BNC check
61	high tuition fee	1	1
62	the worst case	1	>5
63	negative feelings	1	>5
64	good preparation	1	>5
IV	**N-N**		
65	culture shock	3	>5
66	language barrier	1	>5
67	tuition fee	1	>5
V	**N-of-N**		
68	in terms of education	1	>5
69	cost of living	3	>5
70	cost of tuition fee	1	2
71	cost of accommodation	1	>5
72	cost of food	1	>5
VI	**Adv-adj**		
73	much easier	1	>5
74	completely strange	1	2
VII	**Adv-V**		
75	live far (away) (from houses)	1	>5
76	live abroad	1	>5
77	studying abroad (used in the question)	1	>5
78	work part time	1	>5

As can be seen in the above tables, the number of combinations extracted from the two texts was 74 and 78, including 3 combinations used in the essay questions in each text. After the subtraction of the 3 combinations used in the question from the

total, for learners did not use these on their own, the number of combinations from the two texts needing to be checked against BNC for their conventionality was 71 and 75. After the BNC check, the number of combinations that did not reach the threshold was equally 13 from text T12A and text T22A. The number of combinations not meeting the threshold was distributed over the seven groups as in Table 4.26:

Table 4.26 *Number of combinations not meeting the threshold*

Texts	V-N	N-V	Adj-N	N-N	N-of-N	Adv-Adj	V-Adv	Total
T12A	4	1	1	5	1	0	1	13
T22A	5	4	2	0	1	1	0	13

All of these combinations were then judged by native speakers. Before sending them Excel files of combinations for making judgements, I arranged a face-to-face meeting to give them detailed instructions on what exactly they needed to do on the British side and a meeting on Skype on the American side. The important message to be communicated was what was meant by 'acceptable', 'unacceptable' or 'questionable'. The case of *receive attention* was taken to demonstrate here that the combination of *receive* and *attention* does sound natural and therefore needed to be judged 'acceptable' even though their use in the sentence *They tried to receive attention by sounding their horns* sounds awkward. Such a case would be regarded as inappropriate use of the combination in a particular context and the informants were required to give suggestions for correction. Appendix 11 presents detailed information on the judgements and suggested corrections for the combinations from the two written texts.

After the first pair of informants had judged the acceptability of the combinations, they were passed back to me for filtering and checking before passing them on to the second pair of informants. The same judgement from the two informants would constitute the final judgement, so only combinations having different judgement results would be examined further.

8 out of 13 combinations in T12A and 10 out of 12 in T22A were judged the same. These similar judgements constituted the final judgements. Combinations that were finally judged acceptable (e.g. *daily expense*) were then processed to identify whether or not they were strong collocations. Combinations judged unacceptable, such as *face differences* as in *They have to face a lot of differences in various fields such as lifestyle*, were considered to be an error in collocational use. A suggested correction for this combination was *they have to confront a lot of differences* by the American informants and *they find differences in various life aspects such as…* by the British informant. Table 4.27 presents combinations that were judged differently, and which were passed to the second pair of native speakers for judgement.

Table 4.27 *Combinations needing further judgement*

N	Combinations	Acceptable (+)	unacceptable (-)	Questionable (?)	context
I	**TEXT 12A**				
1	means lack of support				being far away from home means lack of family support
2	come back to (their) hometown				...after they come back to their hometown.
3	chief cook				…, he has to be ready to work as chief cook to bottle washer.
4	expense cost				daily expense cost a great deal of money
5	balance studying				It requires great effort to balance studying and working.
	TEXT 22A				
6	marks get				their marks at school get worse and worse
7	cost of tuition fee				

The second pair of informants, in turn, was requested to make judgements. Table 4.28 below presents the results of the final judgements for all the combinations in the two texts T12A and T22A that did not reach five occurrences in the BNC.

130

Table 4.28 *Final judgement results*

	Combinations	(+)	(-)	(?)	(+)	(-)	(?)	(+)	(-)	(?)	(+)	(-)	(?)	Final
	TEXT 12A													
1	face (a lot of) differences		(-)			(-)								(-)
2	means lack (of support)		(-)		(+)				(-)		(+)			?
3	come back to (their) hometown		(-)		(+)				(-)		(+)			?
4	issues get on		(-)			(-)								(-)
5	daily expense	(+)			(+)									(+)
6	learning career		(-)			(-)								(-)
7	chief cook	(+)					(?)			?	(+)			(+)
8	bottle washer	(+)			(+)									(+)
9	students' nerves		(-)			(-)								(-)
10	expense cost		(-)		(+)				(-)		(+)			?
11	the case of a student	(+)			(+)									(+)

	Combinations	(+)	(-)	(?)	(+)	(-)	(?)	(+)	(-)	(?)	(+)	(-)	(?)	Final
12	learned abroad	(+)			(+)									(+)
13	balance studying							(+)			(+)			(+)
	TEXT 22A													
1	give students some disadvantages		(-)			(-)								(-)
2	adapt to the (new) culture	(+)			(+)									(+)
3	put up with (their negative) feelings		(-)			(-)								(-)
4	have (many financial) issues	(+)			(+)									(+)
5	neglect (their) study	(+)			(+)									(+)
6	students commit	(+)			(+)									(+)
7	tuition fee (which may) affect	(+)			(+)									(+)
8	students … neglect	(+)			(+)									(+)
9	marks get		(-)		(+)			(+)			(+)			(+)

	Combinations	(+)	(-)	(?)	(+)	(-)	(?)	(+)	(-)	(?)	(+)	(-)	(?)	Final
10	high tuition fee	(+)			(+)									(+)
11	cost of tuition fee		(-)		(+)			(+)			(+)			(+)
12	daily expense	(+)			(+)									(+)
13	completely strange	(+)			(+)									(+)

I was looking for the Log Dice score of combinations that were conventional from the BNC search. *Have + chance* was the first combination to be processed, as follows:

(1) *chance* (the base) was the first word of the combination to be searched for in the BNC.

(2) I then used the collocation sorting function from Sketch Engine (Kilgarriff, 2004) to look for collocation candidates from the BNC. Figure 4.2 below is the screenshot of the statistics and functions chosen to look for collocation candidates for the noun *chance*. A lemma was chosen for the attribute function, and collocation candidates had to be in the range of five to the left and the right of the base to be considered.

Figure 4.2 *Statistics and functions to look for collocation candidates*

- The Log Dice score of *have + chance* was identified as 6.4, above the threshold of 4.0. Thus the combination was identified as a strong collocation.

As for combinations occurring fewer than five times in the BNC, but judged acceptable or largely acceptable by native speakers, I processed them as follows:

- Combinations occurring three or four times in the BNC: I also looked for their Log Dice score and this process was carried out in the same way as above.

- Combinations not occurring or only one to two times in the BNC: these were all considered casual combinations. This is because the BNC only presents the Log Dice score of combinations occurring at least three times. Their co-occurrences were too few to be considered strong collocations. Details of the Log Dice score of all combinations extracted can be found in Appendix A (on the CD-ROM).

4.2.1 Collocation analysis from the first set of written texts

Following the procedure presented in the section above, I processed the first set of written texts, 29 essays written without the OOCD support. A total of 1,982 combinations were extracted from those writings (see Appendix A on the CD-ROM). As discussed in the Literature Review chapter, combinations which had appeared in the question rubric were not counted since they were not produced by students. After the subtraction of those combinations (157 in total), the number of combinations that needed to be processed was 1,825. After searching the BNC using Sketch Engine software, 1,435 combinations occurring above the threshold of 5 were identified as conventional. The remaining 390 combinations not reaching the threshold were sent to native informants for judgement (see Appendix B on the CD-ROM). The distribution of judgement results is shown in Table 4.29 below:

Table 4.29 *Distribution of judgement results*

+ Acceptable	(+) Largely acceptable	? Questionable	(*) Largely unacceptable	* Unacceptable	Total
260	31	34	4	61	390

The judgement process was not always straightforward, however. There were some cases where the native informants asked for more information regarding the context

to understand what the students wanted to express. They judged some combinations 'questionable' when they were not clear what the students wanted to convey, rather than because combinations did not sound natural. For such cases, they often did not give suggestions on how to correct them. Examples for such cases are *the pace of relationship* and *brand new environment* in these sentences:

> (T25B) *The generation gap in family,* **the space of relationship** *has become larger*.
>
> (T7A) *Many people have difficulty in making friends in the new country because they do not know how to make small talks or what is considered normal behaviour in this* **brand new environment***.*

As presented in the above section, combinations that were acceptable and largely acceptable (291) were processed in exactly the same way as conventional combinations. From the analysis of combinations extracted I found that students made mistakes with collocations, but perhaps more often they produced language that was not wrong but sounded awkward or slightly awkward. They sometimes might not clearly express what they want to say. Hence, unacceptable, largely unacceptable, and questionable combinations should be considered odd collocations rather than errors. Take, for example, *to build their independence* as in *There is a chance for students to build their independence while they are studying in foreign countries* (T14A). This was judged questionable and one judge suggested it be changed to *to increase their independence*. If native-like language use is the goal, combinations such as this should be avoided, but it is inappropriate to call them errors.

It should be noted that there were quite a lot of conventional combinations (56 cases) that were identified as awkward. This is because they fell into one of the following two cases:

> (1) The combination was not used appropriately in meaning. Take *making ... conversation* as in *Having dinner and making a close conversation with our friends or families can help us clear away the stress* (T16B), for instance.

The Sketch Engine search for *make* and *conversation* showed that they co-occur 196 times, of which 113 are counted as validly co-occurring. That is, they are combinations whose element relationship is verb-object. They lie scattered across all of the 10 pages in which the two words co-occur in the BNC. Hence, the combination is conventional. Its Log Dice score is 4.8 - above the threshold. The collocation is in fact fairly fixed; it does not allow for versions such as **make a conversation* or **make a close conversation*. It was therefore identified as not being used appropriately in the '*a close conversation*' context. In this particular context the informants suggested *have a close conversation* as the collocation to be used.

(2). The mistake lies in other elements rather than the base or the collocate. *Approach to a … language* as in *Students can approach to a new culture, new language and study in varied environment* (T1A) is an example to illustrate the point. *Approach* and *language* co-occur more than five times in the BNC and their Log Dice score is quite high at 6.8, but they were identified as 'wrong' owing to the preposition '*to*' being added.

After extracting and identifying odd collocations, I noticed that many mistakes could possibly be counted twice. Take, for instance, the combinations *people + take advantage on* (N-V) and *take advantage on + students* (V-N). These were extracted from *Some bad people can take advantage on students and lead them to social evils and crimes such as drugs and prostitutes* (T13A). The same token *take advantage on* was extracted into two combinations, and the error lies in the use of the preposition *on* of the verb group. There is no principled way to decide whether the error belongs to N-V or V-N combination. For such a case, I have applied a particular rule; that is, the error was counted as a V-N error (*take advantage *on students*) since it belongs more to the V-N side than the other side. This phenomenon was deemed to occur only with collocation patterns containing verbs of which there might be more than one element. In an attempt to avoid counting one error twice, I needed to get back to the written texts to scrutinize whether they were the same tokens extracted out before deciding. 11 cases like this were found from the first set of essays.

All of the combinations were subsequently identified as strong collocations, casual combinations, idioms and odd collocations. Their distribution is shown in Table 4.30 below:

Table 4.30 *Distribution of the processed combinations*

Strong collocations	Casual combinations	Idioms	Odd collocations
1,290	415	2	117

Only two idioms were found: *a piece of cake, get on students' nerves* (T12A). These were identified as idioms because firstly the noun bases do not carry a literal meaning in these contexts. Their meanings according to the Oxford dictionary are:

> + *cake*: *An item of soft sweet food made from a mixture of flour, fat, eggs, sugar, and other ingredients, baked and sometimes iced or decorated*;
>
> + *nerves*: *A whitish fibre or bundle of fibres in the body that transmits impulses of sensation to the brain or spinal cord, and impulses from these to the muscles and organs*.

Nor are the meanings of the combinations as a whole transparent. In other words, the meaning of the combinations is not the combination of meanings of the individual elements. *A piece of cake* means *something easily achieved*; *get on someone's nerves* means *irritate someone*. I will not make any comment on the use of these idioms in these contexts because it is beyond the scope of this study.

The odd collocations here comprise combinations identified as 'unacceptable'/ 'largely unacceptable' (-), 'questionable' (?), and even those which were identified as acceptable or largely acceptable (+) but were semantically inappropriately used in the contexts. The native speakers made suggestions as to more appropriate collocations for these (109 out of 117 odd collocations). In an effort to understand clearly what the mistakes are, and from which pedagogical implications could be drawn, I divided these odd collocations into different types. They fall into the following three groups:

(1) *Wrong choice of collocating word (coll)*: These are the combinations that were not conventional and were later judged unacceptable by the native informants. Odd collocations of this type can be subdivided into three categories:

a. *Errors in the collocate (coll)*: The suggested correction involved the use of a different collocate. Take, for example, **face a lot of differences* (T12A), which was corrected to *confront a lot of differences* by both informants of the first pair.

b. *Errors in the base (coll-base)*: the category involved the inappropriate use of a base in a particular context, such as *demerits* in *Living far from home, taking a lot of money, and being shock-cultured are considered as its three main **demerits***. *Drawbacks* was deemed to be a better word to express the idea.

c. *Errors in the combination as a whole (coll-whole)*: These are cases where the use of the combination as a whole sounds awkward rather than the base or the collocate individually. An example to illustrate the point is the combination **Productivity ... take (the responsibility)* as in *Productivity at work takes the responsibility for the increasing rate of stress among workers*. Combinations were also classified into this group of errors if it was not clear what the combination meant in the context or no specific suggestion on how to correct it could be given out. **Rewarding trend* in (T11A) *These days, studying abroad is becoming a rewarding trend for students all over the world* is one of the cases.

(2) *Errors not due to wrong choice of collocating word (~~coll~~)*: These are combinations with errors occurring neither at the base nor at the collocate. *Take advantage on students* (T13A) was an example in which the error lies in the use of the preposition *of* instead of *on*, not at the base *advantage* or the collocate *take*. The identification of errors of this type was based on a Sketch Engine search and was carried out in conjunction with the process of identifying whether or not a combination was conventional.

(3) *Inappropriate meaning (meaning)*: Based on the Log Dice score, this is a group of strong collocations; however, their meanings in certain contexts were inappropriate. Semantically, *have some drawbacks* was inappropriately used in the sentence *Students studying abroad have some drawbacks* (T36A). It was

suggested that this be changed into *face/experience some drawbacks* by the informants.

Table 4.31 provides full details of odd collocations identified from these written texts. The second column is odd collocations, the third column presents error types and the last column the suggested correction by native informants.

Table 4.31 *Odd collocations from the first set of essays*

N	Odd collocations	Error types	Suggested correction
I	**V-N**		
1	face (a lot of) differences in	coll	to confront (a lot of) differences
2	shorten the gap	coll	reduce/minimize the gap
3	come back to (their) hometown	meaning	return to (their) hometown
4	give students some disadvantages	coll-whole	students studying abroad present some disadvantages
5	put up with (their negative) feelings	coll	live up to (their negative) feelings
6	approach to (a new) language	~~coll~~	approach a new language
7	approach to (a new) culture	~~coll~~	approach a new culture
8	face with paying (for)	~~coll~~	be faced with/face sth
9	bring students (new) environment	~~coll~~	bring students to a new environment/ place students in a (new) environment
10	suffer (the) homesickness	~~coll~~	suffer from homesickness
11	support their living	coll	to offset/support their cost of living
12	have a (part-time) job	meaning	get a (part-time) job
13	enlarge (their) knowledge	meaning	Increase/improve (their) knowledge
14	know about (their) tradition	meaning	learn about the tradition
15	brand new environment	coll-whole	No suggestion
16	prepare (more) money	coll	have enough money
17	afford for eating	~~coll~~	afford to eat/afford meals
18	meet robberies	meaning	be victim of crime
19	dominate drawbacks	coll	outweigh the drawbacks
20	perceive knowledge	coll	increase/ acquire knowledge
21	brand knowledge	coll	acquire/ gain knowledge
22	cope under pressure	~~coll~~	cope with pressure
23	raise students' independence	meaning	increase students' independence
24	take advantage on students	~~coll~~	take advantage of students

N	Odd collocations	Error types	Suggested correction
25	build (their) independence	coll	become/increase (their) independence
26	face with (many) problems	~~coll~~	be faced with/ face problems
27	promoting for (senior) posts	~~coll~~	be promoted to (senior) post
28	appoint to (a higher) level	coll-whole	No suggestion
29	disagree to the idea	~~coll~~	disagree with the idea
30	means lack of support	meaning	means loss of support
31	learn about communication	~~coll~~	learn communication
32	live with roommates	coll	to get along with roommates
33	lead to mind-set broadening	coll-base	lead to a broaden of their mind-set
34	go outside (their) home town	meaning	leave their hometown
35	suffer from … eating habit	meaning	adopt eating habit
36	prepare for (their) knowledge	~~coll~~	prepare their knowledge/learn
37	go back hometown	coll	return/back to their hometown
38	be led into stress	coll	experience/endure stress
39	affect to (studying) result	~~coll~~	affect (studying) result
40	adapt with the environment	~~coll~~	adapt to the environment
41	get familiar with culture	meaning	to become familiar with culture
42	draw problem	coll	create problem
43	bring (bad) attitudes	~~coll~~	bring about (bad) attitudes
44	face with drawbacks	~~coll~~	face drawbacks
45	help language barrier	coll	minimize language barrier
46	intriguing a (great) concern	coll	engendering a (great) concern/ create … intrigue
47	put students under disorders	coll-whole	put students under the threat of serious disorders
48	have some drawback	meaning	face/experience (some) drawback
49	considered … demerits	coll-base	considered …drawbacks
50	has (their own) cooking	meaning	has (their own) cuisine/cooking process
51	face with many difficulties	~~coll~~	be faced with/face many difficulties
52	face with many drawbacks	~~coll~~	face sth/be faced many drawbacks
53	listen a language	~~coll~~	listen to a language
54	subordinate the beneficial	coll	are subordinate the benefits
55	disturb (their) studying	meaning	impair their study
56	show (their) perspective/opinion	meaning	share (their) perspective/opinion
57	live independent way	~~coll~~	live independently
58	suffer stuffs	coll-base	suffer problems
59	come back hometown	meaning	return hometown
60	pay attention another thing	~~coll~~	pay attention to
61	digress students from studying	coll	distract students from studying

N	Odd collocations	Error types	Suggested correction
II	**N-V**		
1	issues get on	coll	No suggestion
2	the loop (will) keep on	coll-whole	the cycle will go on
3	benefits were commented	coll	benefits were noted
4	a chance (of work) is waiting for	coll-base	a strong likelihood of finding work is waiting
5	gender equality presents	coll	gender equality is present
6	students concern	coll	students be concerned about
7	the barrier cause	coll	the barrier result in/create
8	this trend has been intriguing	coll	this trend has been engendering
9	students … unaccepted	coll	students … not accepted
III	**Adj-N**		
1	abroad students	coll	overseas students
2	rewarding (popular) trend	coll-whole	No suggestion
3	countless experience	meaning	priceless experience
4	old-fashion minds	coll	old-fashioned mind
5	studying progress	meaning	academic progress
6	studying result	coll	study result/result of studying
7	usual aspects	coll-whole	other professionals
8	complicate tasks	coll-whole	No suggestion
9	malnutrious (eating) habit	coll	poor eating habits
10	advanced way	Coll-base	advanced form of
11	normal students	meaning	traditional/domestic students
12	unpredicted issues	coll	unpredictable/unanticipated issues
13	learning career	coll	academic career
14	studying responsibility	coll	responsibility of the studies
15	staying time	coll-whole	duration of their stay
16	psycholic disorder	coll	psychological disorder
17	main demerits	coll-base	main drawbacks
18	temporary world	coll-base	temporary situation
19	well-equipment environment	coll	well-equipped classroom
20	first-stage time	coll-whole	period of adaptation
21	unappropriate action	coll	inappropriate action
22	good certificate	coll	good grade/worthwhile certificate
23	working treatment	coll-whole	treatment of workers
III	**N-N**		
1	students' nerves	coll-whole	No suggestion
2	expense cost	coll-whole	No suggestion
3	environment's danger	Coll-whole	dangerous environment

N	Odd collocations	Error types	Suggested correction
4	study career	coll-base	academic/student career
5	education quality	coll-whole	the quality of education
6	study fee	coll-base	course fee, tuition fee
7	electricity fee	coll	utility bill/electricity costs
8	mind-set broadening	coll-whole	to broaden mind-set
9	food cuisine	coll-whole	either food or cuisine
10	country cuisine	Coll-base	country's cuisine/national cuisine
11	youth life	Coll-whole	youths' life
12	shock culture	coll-whole	culture shock
III	N-of-N		
1	port of knowledge	coll	a source of knowledge
2	their need of studying	coll-whole	their study requirements
3	large number of money	coll	large amount of money
III	Adv-V		
1	study safely	coll-whole	No suggestion
2	live normally	meaning	has difficulty coexisting/living peacefully with
3	strictly affect	coll	seriously affect
4	abroad studying	coll-whole	study abroad
5	upgrade regularly	meaning	consistently improve
6	learn … seriously	coll	learn effectively
VI	Adv-Adj		
1	definitely valuable	coll-whole	No suggestion
2	far difficult	~~coll~~	far more difficult/far too difficult
3	more opened	coll-base	more open

4.2.2 Collocation analysis from the second set of written texts

After processing combinations extracted out from the first set of written texts, I continued analysing the second set of written texts. 1,919 combinations were extracted from these texts (see the CD-ROM). Again, after eliminating the combinations used in the question rubric (88 combinations), I processed 1,831 combinations. 1,374 combinations were identified as conventional based on their frequency of co-occurrence on the BNC. I then searched for their Log Dice score to separate strong collocations from casual combinations. The remaining 457 combinations not meeting the conventional test were sent to the native speakers for judgements. Table 4.32 provides the results of the judgements:

Table 4.32 *Distribution of judgement results of the second set of essays*

+ Acceptable	(+) Largely acceptable	? Questionable	(*) Largely Unacceptable	* Unacceptable	Total
314	25	39	3	76	457

They were subsequently identified as strong collocations, casual combinations, idioms and odd collocations, and their distribution is shown in Table 4.33 below:

Table 4.33 *Distribution of the processed combinations*

Strong collocations	Casual combinations	Idioms	Odd collocations
1,184	513	0	134

Results showed that odd collocations with V-N pattern were the highest with 62 collocations. Adj-N pattern was the second highest with 31 odd collocations. N-of-N and Adv-Adj pattern had the lowest numbers, 3 and 2, respectively. Figure 4.3 below summarises the distribution of odd collocations by their patterns.

Figure 4.3 *Distribution of odd collocations by patterns of the second set of written texts*
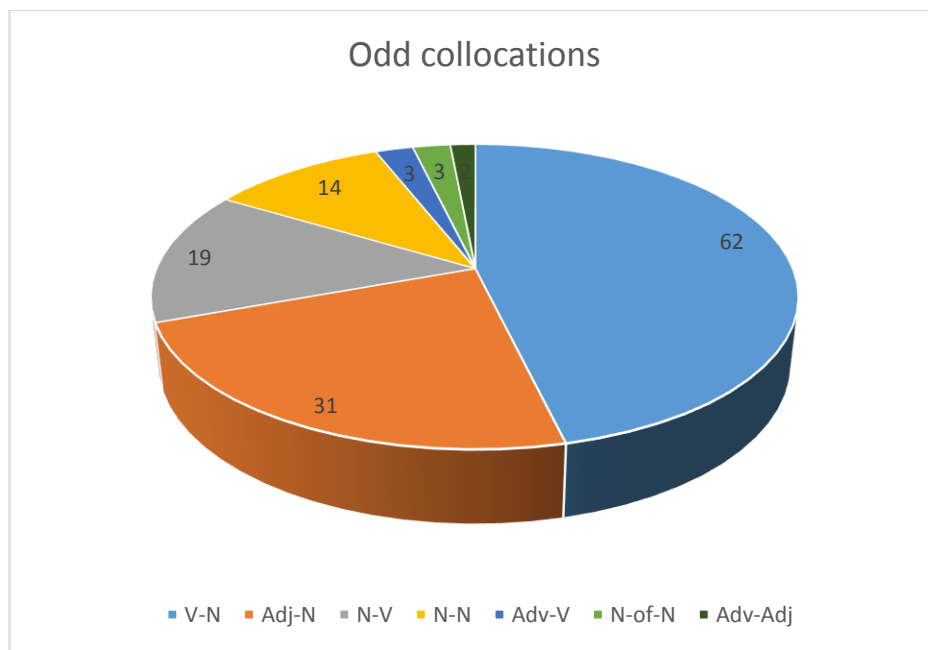
The highest number of odd collocations in a text was 20 (T13B). There were some texts without any errors or with only one (T6B, T7B, T14B, T17B, T43B) (see appendix A in the CD). Table 4.34 below presents odd collocations from the second set of essays. The second column is the odd collocations, the third column presents error types, and the last column is suggested corrections by the native informants. Odd collocations were also categorized into three main groups: (1) oddness in the collocating words, which were subdivided into oddness at the base, at the collocate, and at the combination as a whole; (2) oddness not at the collocating words; and (3) inappropriate meaning. Combinations highlighted were those looked up in the collocation dictionary

Table 4.34 *Odd collocations from the second set of essays*

N	Odd collocations	Error types	Suggested correction
I	**V-N**		
1	bring stress to people	coll	cause stress to people
2	take (any) time	meaning	have time
3	pay (a lot of) effort	coll	make (a lot of) effort
4	remain (social) relationship	coll	maintain (social) relationship
5	choose (negative) reaction	coll	have (negative) reaction
6	retain a (tremendous) responsibility	meaning	have a tremendous responsibility
7	provide people with disadvantages	coll	problem with "provide"
8	put people under disorders	coll-whole	put people under the threat of
9	suffer from … traffic jam	coll-base	suffer terrible traffic
10	take the stress	meaning	withstand the stress
11	interfere with (our) inside	coll-base	interfere with our mental health/wellbeing
12	lead to (an) unability	coll-base	lead to an inability
13	cope under pressure	~~coll~~	cope with pressure

N	Odd collocations	Error types	Suggested correction
14	be facing with (the) problem	~~coll~~	face a problem/be faced with
15	better public transportation	coll	improve (public) transportation
16	benefit (both you and the) traffic	coll	improve the traffic
17	requires (higher) requirements	coll	lead to (higher) requirements
18	to get achievements	coll	have a thirst for achievements
19	make stress (appear)	coll	create/cause so much stress
20	focus on the burden (of working)	coll	mainly stem from the burden
21	have stress	coll	suffer from stress
22	increase the change	coll-base	increase the chance (of getting cancer)
23	save (your) nerve cells	coll	strengthen/regenerate (your) nerve cells
24	get (their own new) cars	meaning	buy … cars
25	foster people (to carpool)	coll	encourage people (to carpool)
26	decline exhaust fume	coll	reduce exhaust fume
27	making a (close) conversation	meaning	have a (close) conversation
28	clear away the stress	coll	reduce/relieve stress
29	arrange (their) timeline	coll-base	manage their time
30	result in stress hormone	coll-base	result in increased production of stress hormones
31	working under (excessive) sounds	coll-base	working in/under excessively noisy conditions
32	enduring stress	coll	experiencing stress/enduring high level of stress
33	regulate (cars') demand	meaning	regulate demand for cars
34	reshape the minds (of people)	coll-base	reshape people's views/opinions
35	better … life	coll	improve … life
36	plant in (people) the concept of	coll	ingrain/instil in (them) the

N	Odd collocations	Error types	Suggested correction
			concept of
37	boost (positive) result	meaning	encourage (positive) result
38	explain for this phenomenon	coll	explain the phenomenon
39	propose materialism	coll	encourage materialism
40	shift (too much) emphasis	meaning	place (too much) emphasis
41	discuss on work	~~coll~~	discuss work
42	endure through a power cut	~~coll~~	endure a power cut
43	outshone the … ways (of resting)	coll	certainly exceeded the ways of resting
44	face with many pressure	~~coll~~	face sth/be faced with many pressure
45	shout the worker	~~coll~~	shout at worker
46	reach to the victory	~~coll~~	reach their goal/objective
47	face with many problems	~~coll~~	face/be faced with sth many problems
48	consider about (their) work	~~coll~~	consider (their) work
49	train ... skill	meaning	develop/improve … skill
50	deal deadlines	~~coll~~	deal with deadlines
51	please your goals	coll	achieve our goal
52	bring about stress	~~coll~~	bring stress
53	dry people out	coll	weary people out
54	draw … attraction	coll-base	capture the attention of
55	affect to the environment	~~coll~~	affect the environment
56	limit the overpopulation	meaning	combat/fight the overpopulation
57	face with … difficulties	~~coll~~	face sth/be faced with difficulties
58	suffer stuffs	coll-base	suffer problems
59	trouble with relationship	coll-whole	result in relationship troubles
60	release stress (out of their life)	coll	minimize stress

N	Odd collocations	Error types	Suggested correction
61	take a (comfortable) chat	coll	have a (comfortable) chat
62	stand stress	coll	face/experience stress
II	**N-V**		
1	stress … provides	coll	stress …causes
2	reasons create	coll-base	(another) factor causing stress
3	stress was formed	coll	stress arises
4	stress have	coll	no suggestion
5	pace of living getting (faster)	coll-whole	life is moving at faster pace
6	productivity takes (responsibility)	coll-whole	demands for increased productivity are responsible for
7	government & individual … join (hands)	coll	government and individual work together
8	(modern) life propose (materialism)	coll-whole	no suggestion
9	technological advancement has outshone	coll-whole	no suggestion
10	the pace of relationship … become	coll-whole	the speed at which we form relationship become
11	stress has	coll	no suggestion
12	the loop … carry on	coll-whole	the cycle … go on
13	women role as housewives	coll	women have the role/serve the role
14	failures stretch	coll	failures overwhelm us
15	stress covers	coll	stress dominates/overwhelms
16	money … turn (people)	coll	money changes (people)
17	stress comes out	coll	stress starts
18	phenomenon … drawn	coll-whole	no suggestion
19	smoke … affect	coll	smoke … impact
III	**Adj-N**		

N	Odd collocations	Error types	Suggested correction
1	hurry pace	coll	hurried pace
2	working burdens	coll	burden of work/employment burden
3	multitask person	coll	multi-tasker/multitasking person
4	wrong living style	coll	wrong style of living/poor living style
5	unhealthy status	coll-base	unhealthy state
6	bad chemical	coll	harmful chemicals
7	stressful world	coll	stress world that we live in
8	numerous habitant	coll-base	many inhabitants
9	closely conversation	coll	have a close conversation
10	social houses	coll-base	social housing/public housing
11	great examples	meaning	no suggestion
12	excessive sounds	coll-base	excessive noise
13	dangerous health troubles	coll-base	dangerous health conditions
14	concerned issue	coll	concerning issues
15	relative cause	coll-whole	no suggestion
16	stressful people	coll	people experiencing stress
17	high-speed life	coll-base	high-speed living/living a high-speed life
18	fast relations	coll-base	build relationship faster
19	breaking days	coll-whole	some days away from work/leisure time
20	good status	coll	high status/well-being
21	gigantic properties	coll	significant property
22	burning trouble	coll-base	burning issues
23	caffeine-contained substances	coll	caffeine containing substances
24	minimal fears	coll-base	minimal concern

N	Odd collocations	Error types	Suggested correction
25	caused sign	coll-whole	no suggestion
26	healing method	meaning	keeping a well-balanced diet
27	sufficient nutrition	meaning	nutritious/well-balanced diet
28	constant exercising	coll-whole	regular exercise
29	official shift	coll	regular shift
30	temporary world	coll-base	temporary situation/environment
31	inside factors	coll	internal factors
IV	**N-N**		
1	soil sources	coll-whole	no suggestion
2	human being's peace	coll-whole	peace of mind
3	impact effect	coll-whole	profound impact
4	living style	coll-whole	lifestyle/style of living
5	collar woman	coll-whole	white collar worker
6	chaos stuffs	coll-whole	chaotic things
7	stress activities	coll-base	stressful activities
8	life standard	coll-base	standard of living/living standard
9	working pressure	coll-base	work pressure
10	cars' demand	coll-whole	demand for cars
11	vehicles' demand	coll-whole	demand for vehicles
12	commit suicide rate	coll-base	suicide rate
13	today world	coll-whole	today's world/world today
14	life quality	coll-whole	quality of life
V	**N-of-N**		
1	(a great) deal of chemicals	coll-whole	large amount of chemicals
2	state of emotion	coll-whole	emotional state
3	stress of relationship	coll-whole	relationship stress
VI	**Adv-V**		

N	Odd collocations	Error types	Suggested correction
1	work more and more	coll	work increasingly longer hours
2	tax greatly	coll	tax heavily
3	strictly affect	coll	seriously affect
VII	**Adv-Adj**		
1	significantly polluted	coll	extremely polluted
2	interestingly comfortable (chat)	coll	interesting and comfortable (chat)

There are two cases where combinations were not of the collocation patterns being considered, i.e. *state of emotion* and *stress of relationship*. The first nouns of the combinations, *state* and *stress*, are not quantifiers to be allocated to the N-of-N pattern; however, as I have discussed in the methodology chapter, in order to avoid missing odd collocations, combinations like those were still examined. These cases were eventually identified as odd collocations since they were used as a result of learners not knowing the appropriate collocations, i.e. *emotional state* and *relationship stress*.

4.3   Analysis of variety in collocational use

To have a thorough understanding of my data I needed to investigate not just the appropriateness but the variety of collocational use. This involved an exploration and comparison of this dimension of the students' written texts. Based on the recording of the number of times a particular combination occurred in the texts, I counted and compared the number of combinations that had been repeatedly used more than twice in a text. Special notice was paid to these combinations because, as discussed in the Literature Review chapter, if not for the purpose of emphasis, the same combination should not be repeatedly used. In this way, I could understand whether or not the OOCD had any impact on the variety of collocational use. Again, this comparison was made between two written texts from the same participants. Table 4.35 below compares the number of combinations that were used more than twice from the two sets of essays. The second and the fourth column show the number of combinations from the two sets of written texts, the 'collocations' columns present combinations

repeatedly used in the writings, and the numbers in parentheses indicate the number of times they were used in the essays.

Table 4.35 *Repeatedly-used collocations*

N	First set of essays		Second set of essays	
	Number of combinations	Collocations	Number of combinations	Collocations
T1A	1	Studying abroad (4)	0	
T4A	1	Studying abroad (8)	0	
T6A	3	Students have (3) Foreign country (4) Studying abroad (3)	1	People have (5)
T7A	1	Studying abroad (3)	0	
T10A	1	Going abroad (3)	0	
T11A	1	Studying abroad (6)	0	
T12A	2	Foreign students (3) Culture shock (3)	0	
T13A	1	Studying abroad (9)	0	
T14A	3	Studying abroad (6) Students study (5) Foreign country (6)	0	
T15A	2	Senior positions (3) Senior posts (3)	1	Traffic problems (3)
T16A	2	Studying abroad (4) Students study (4)	0	
T17A	4	Students study (4) Students get (5) Foreign countries (9) Studying abroad (5)	1	Reduce stress (3)
T19A	0		0	
T20A	0		1	Noise pollution (3)
T21A	2	Students learn (3) Studying abroad (3)	0	
T22A	2	Foreign students (6) Students suffer (3)	1	Morden life (3)

T23A	3	Studying abroad (9) Foreign country (3) Cultural knowledge (5)	0	
T24A	1	Students have (3)	0	
T25A	2	Studying abroad (7) Foreign countries (6)	1	Positive thinking (4)
T27A	1	Going abroad (3)	0	
T28A	2	High cost (3) Study abroad (4)	1	Cause the stress (4)
T31A	2	Culture shock (3) Studying abroad (4)	0	
T35A	0		0	
T36A	2	Take … money (3) A lot of money (3)	1	Human beings (3)
T38A	2	Foreign country (3) Studying abroad (5)	2	Outside factors (3) Reduce stress (3)
T40A	3	Foreign country (3) Studying abroad (3) New environment (3)	0	
T42A	3	Foreign country (3) Studying abroad (5) Students have (4)	6	Public transportation (7) Use money (3) Use transportation (3) Use buses (3) People use (4) Traffic problems (3)
T43A	2	Foreign country (4) Studying abroad (3)	0	
T44A	3	High cost (3) Culture shock (3) Students have (4)	1	Polluted environment (4)

A decision on whether or not the dictionary helps learners avoid repeatedly using the same collocations could only be made if evidence of students using the OOCD to search for another way of expressing the same idea was detected. Besides looking at the recording sheets, which I will be presenting in the next section, I needed to look

at collocations extracted out on the spreadsheets from the individual essays. The purpose was to look for collocations with similar meanings to collocations looked up in the dictionary. This is an attempt to avoid missing collocations that the learners used after the dictionary look-up to avoid repetition, but did not state this in the recording sheets. On the spreadsheets all collocations were presented according to the order of their appearance in the essays; collocations looked up were marked in yellow. As discussed earlier, this way of presenting collocations on the spreadsheet was to enable the process of analysing collocations to run more smoothly and accurately. No collocations were identified as being used to avoid repetition in this process.

4.4   Questionnaire data analysis

To find the frequency and percentage of agreement-disagreement among the participants regarding the survey questions, descriptive statistical analysis on SPSS was used to process the data. Information from 33 questionnaire papers was imported into SPSS.

4.5   Recording sheets analysis

The recording sheets contain both quantitative and qualitative data. As well as recording how learners used the OOCD as a supporting tool and tracing back what collocations were looked up, they gathered statistical information on types of collocation, the percentage of successful look-ups and evaluation of satisfaction with individual look-ups, based on a five-point Likert scale from *very dissatisfied* to *very satisfied*. The data were processed partly using SPSS and partly using NVivo. NVivo was used to process one qualitative question (question 8: Further comments/Why did you use other dictionaries? If yes). The recording sheets were scanned to be imported into NVivo. All the responses from the qualitative questions in the recording sheets are referred to as RS.

Data imported into SPSS includes questions 2, 3, 4, 5, 6, and 7 in the recording sheets, which are reiterated here. These are also the variables in the SPSS file:

Q2: What types of collocations were you looking for? (the choice of 7 collocation types)

Q3. Did you find the word you were looking for? (Yes/No)

Q4. Why did you look up the word? (Checking/Finding)

Q5. Did you find the instruction on how to use it? (Yes/No)

Q6. Did you use the OOCD in combination with other dictionaries? (Yes/No)

Q7. How satisfied were you with what you found? (1 Very dissatisfied – 5 very satisfied)

There were 144 searches in total, including 18 searches for other purposes than collocations. In response to the question about the types of collocations they were looking for, participants chose 'other' rather than a specific type of collocation if they wanted to either check for meaning, spelling, or part of speech of a certain word. These were to be excluded since the search was not for collocation-searching purposes. The average number of collocation check-ups per text was approximately 4. The lowest number of check-ups to complete the writing task was 1 and the highest was 8. 119 out of 126 searches for collocations were done while students were doing the writing. Of the 7 searches done after completing the writing, 6 were to look for Adj-N collocations and 1 to look for V-N collocation. Table 4.36 below summarises the number and types of collocations looked up from each recording sheet.

Table 4.36 *Distribution of collocations looked up in the OOCD of each student*

	Students	V-N	N-V	Adj-N	N-N	N-of-N	Adv-V	Adv-Adj	Total
1	S1	1			1	1	4		7
2	S4		2	1					3
3	S6		1	2					3
4	S7		1	5				1	7
5	S10	3		1		1		1	6
6	S11	2	1		1	1		2	7
7	S12	3		3					6

	Students	V-N	N-V	Adj-N	N-N	N-of-N	Adv-V	Adv-Adj	Total
8	S13		2	1				2	5
9	S14	1	1	2					4
10	S15	2		2			1		5
11	S16	1							1
12	S17			1			1		2
13	S19			2		2			4
14	S20			2					2
15	S21		3		1				4
16	S22	3		2	1				6
17	S23			4					4
18	S24	2	1	3			1		7
19	S25			4			1		5
20	S27	7					1		8
21	S28			1					1
22	S31			1					1
23	S35		1					1	2
24	S36	1	1	5					7
25	S38	1					1		2
26	S40			1					1
27	S42		2	4					6
28	S43		1	3			1	1	6
29	S44			1			3		4
	Total	27 21.4%	17 13.5%	51 40.5%	4 3.2%	6 4.8%	14 11%	7 5.6%	126

As I have presented in the data collection section, all collocations looked up in the dictionary were highlighted in the essays by the participants after they had finished the writing. From the recording sheets I found that there were five cases in which the learners used the OOCD to search for another way of expressing ideas (see Table 4.37 below). For most of these cases the participants responded quite clearly to the purpose of the look-ups except for the second *stress* in recording sheet 27B.

Minimize stress was used after that search, and the participant only wrote on the recording sheet 'to look for synonyms'. It was probably used to avoid repeating *avoid stress* or *handle stress*, which the participant had used before in the writing. In the table below, the second column is the headwords searched from the OOCD, the third column is collocations used after the look-ups, the next column is learners' comments on the search, and the last column is collocations that the students had used earlier in the essays.

Table 4.37: *Collocations searched in the OOCD to avoid repetition*

RS	Word searched	Collocations checked up	Further comments on recording sheets	Collocations of similar meanings
11B	*stress*	*create stress*	To find another way to say *cause stress*	*cause stress*
24B	*problem*	*urgent problem*	To find a way to say *pressing problem*	*pressing problem*
27B	*stress*	*create stress*	I use the dictionary to find word to express *cause stress*	*cause stress, lead to stress*
27B	*stress*	*minimize stress*	To find synonyms	*avoid stress, handle stress*
31B	*problem*	*Common problem*	Students often use *big* or *major* with *problem*, so I want to use other adjectives to make the essay sound smooth and not repeated.	*serious problem, big problem*

13 out of 126 look-ups were done in combination with other dictionaries. The reasons for using the OOCD in combination with other dictionaries were mainly found to correlate with collocates, such as to know the exact meanings of collocates to use appropriately or to look for collocates which could not be found in the OOCD to express their intended meaning. As some participants stated, they sometimes used other dictionaries to search for synonyms of collocates found in the OOCD to avoid repeatedly using the same collocations, especially when the OOCD only gives a very

limited number of collocates. Other reasons were checking spelling, looking for a preposition following the headword entry, or headword entry not found in the OOCD. Table 4.38 below summarises all collocations that learners used other dictionaries along with the OOCD for help.

Table 4.38: *Words searched by using the OOCD in combination with other dictionaries*

N	Texts	Word entry searched in the OOCD	Collocation patterns	Reasons for using other dictionaries
1	T1B	interfere	V-Adv	To find collocates to express the intended idea
2		triggers	N-N	To find meaning of the noun 'trigger'
3		leave	V-Adv	To find meanings of collocates
4	T10B	along	No response	Not found in the OOCD
5	T11B	deniable	Adj-N	To check spelling
6	T16B	face	other	To find preposition following 'face'
7	T19B	goods	Adj-N	To find synonyms of collocates
8	T21B	hazard	Adj-N	To find meanings of collocates
9		law	V-N	To find synonyms of collocates
10		regulate	other	To find meanings of collocates
11	T40B	problem	Adj-N	To find meanings of collocates
12	T41B	value	Adj-N	To find collocates to express the intended idea
13	T44B	global	No response	

There were two cases in which learners did not respond to the question asking which collocational pattern they wanted to look for. They did not use collocations relevant to these words in their essays either. As for the case 'other' (in T21B) the learner looked for something else rather than collocates of a particular pattern.

4.6   Interview data analysis

In what follows, I will present the process of preparing interview data, the choice of the analysis method, and a description of the phases through which the transcriptions were processed.

4.6.1   Preparation of data for analysis

The transcription of interviews was done immediately after they had been conducted. The dataset comprised eight interviews from student informants. It was almost eight months between the transcription and the analysis, so it was important for me to re-listen to the audio records to check if the transcripts were accurate. They were then printed out so that I could mark them up easily. For ease of reference in quoting their response in the findings section with their names remaining anonymous, I named the interviewees in alphabetical order from A to H. As well as being used to process part of the recording sheets, NVivo was used to process this dataset in order for me to keep track of the available data easily and flexibly. Working on paper helped me make sense of the data easily; hence, before importing the data into NVivo, I processed them manually by using different coloured highlighters to highlight what I saw as variables or themes.

4.6.2   Choice of thematic analysis

Braun and Clarke (2006, p. 79) write that thematic analysis is 'a method for identifying, analysing and reporting patterns (themes) within data.' This definition does not differentiate this analysis method from others which are also essentially thematic, such as discourse analysis, interpretative phenomenological analysis (IPA) or grounded theory. The choice of this analysis method for the study was mainly based on the benefits that it brought compared to others. One of its most prominent benefits is its flexibility - the researcher is free to choose theoretical assumptions, research questions, the types of data and even how to implement the analysis (Clarke et al., 2015, p. 225). According to Braun and Clarke (2006), thematic analysis can be used to analyse the majority of qualitative data including interviews, focus groups, qualitative surveys, diaries, story completion tasks, or even secondary sources like newspaper articles. The choice of thematic analysis allows researchers to choose individual elements rather than the 'complete package' (of theoretical framework,

kinds of research questions, data types), in contrast to other qualitative approaches like interpretative phenomenological analysis, Discourse Analysis, Conversation Analysis or grounded theory. IPA, for instance, is 'theoretically bounded' and is an approach used 'to explore in detail how participants are making sense of their personal and social world' (Smith and Osborn, 2015, p. 25). For this analytic method, 'there is no attempt to test a predetermined hypothesis of the researcher' (Smith and Osborn, 2015, p. 28). Conversation Analysis, on the other hand, is often used in studies for studying naturally occurring data which, as Drew (2005, p. 112) describes, 'are not gathered through simulations, through experimental or quasi-experimental tasks, and are not fabricated', and is employed to explore 'how participants arrive at understanding of one another's actions during back-and-forth interaction'. Grounded theory shares with thematic analysis its flexibility in methodological strategies, but it develops analytic codes and categories from data rather than preconceived hypotheses, and relies on simultaneous data collection and analysis. In light of these, thematic analysis is the most suitable method for this study. This analytic method does not require the research to be 'directed towards theory development' (Braun and Clarke, 2006, p. 81) but does not restrict it from building theories from data if they 'emerge'.

I chose an inductive approach to process the dataset to 'stay as close as possible to the meanings in the data' (Clarke et al., 2015, p. 225). Put another way, coding in this study is more data-driven than theory-driven. In this way, this form of analysis 'bears some similarity to grounded theory' (Charmaz, 2015, p. 83). It is, nevertheless, true that pure induction is impossible since, as Clarke et al. (2015, p. 225) put it, 'analysis is shaped by a researcher's theoretical assumptions, disciplinary knowledge research training, prior research experiences, and personal and political standpoints.' All the themes and categories were identified based on the explicit or surface meaning of the data.

I followed Braun and Clarke (2006) who contend that themes do not emerge or are discovered but rather exist in the researcher's head; therefore, the process of identifying themes is influenced by the researcher's theoretical position. Themes in

this study were identified based on both their relevance to the research questions and their prevalence throughout the data. Since the data was 'driven by the researcher's theoretical or analytic interest' and themes were coded for specific questions, this analysis is specifically called 'theoretical thematic analysis' (Braun and Clarke, 2006, p. 84).

4.6.3   0Phases of data analysis

I followed the step-by-step guide for thematic analysis recommended by Braun and Clarke (2006, p. 87), shown in the table below.

Table 4.39 *Phases of thematic analysis*

Phase	Description of the process
Familiarizing yourself with the data	Transcribing the data (if necessary), reading and rereading the data, noting down initial ideas.
Generating initial codes	Coding interesting features of the data in a systematic fashion across the entire data set, collating data relevant to each code.
Searching for themes	Collating codes into potential themes, gathering all data relevant to each potential theme.
Reviewing themes	Checking if the themes work in relation to the coded extracts (Level1) and the entire data set (Level 2), generating a thematic 'map' of the analysis.
Defining and naming themes	Ongoing analysis to refine the specifics of each theme, and the overall story and analysis tells, generating clear definitions and names for each theme.
Producing the report	The final opportunity for analysis. Selection of vivid, compelling extract examples, final analysis of selected extracts, relating back of the analysis to the research question and literature, producing a scholarly report of the analysis

Phase 1: Familiarisation with the data

Since the data had already been transcribed and printed out, familiarisation involved rereading the transcripts several times to become 'thoroughly acquainted with the data' (Sapsford and Jupp, 2006, p. 169) since 'each reading has the potential to throw up new insights' (Smith and Osborn, 2015, p. 40). While reading, I highlighted words or strings of words that are relevant to the research questions and made notes at some points to easily make sense of them when I got back. Consulting some notes taken while interviewing assisted me in comprehending their perceptions on dictionary use and in-depth reasons for the evaluation. In five of the eight interviews, the informants used English; and in the other three, the informants mainly used Vietnamese and code switched at some points. They were transcribed exactly as in the recordings. I decided not to translate the full Vietnamese interviews except for some strings of words that I later used as extracts to illustrate analytic claims, since it is easier for me to fully make sense of the data when it is expressed in my mother tongue.

Phase 2: Generating initial codes

With the research questions in mind, I started generating initial codes. Admittedly, it is 'tempting to skip this phase and start theme identification immediately' (Clarke et al., 2015, p. 234). When I looked back at the codes for the first interview there were only seven, and most of them were much like themes rather than codes since they were quite general (e.g. *evaluation*, *likeable*, *dislikeable*, and *desirable features*, or *how OOCD is used*). I found that one-word codes like 'evaluation' did not make sense when I was not working on that data item, and because of that I could hardly identify 'analytically coherent connections between multiple codes' (Clarke et al., 2015, p. 235). Good codes, according to Clarke et al. (2015, p. 235) 'make analytic sense without needing to see the data; they provide a shorthand analytic "take"'. I kept working on the next interviews with a strong awareness in mind that this coding phase should be carried out with great care in order to generate deep engagement with the data.

Because of the efficiency of NVivo 11 as an assisting tool, I used it for coding all the interviews and then for collating them to prepare for the next phase. NVivo facilitates researchers by allowing them to trace back the codes in the transcripts by

just a click on the link above the coded reference. In this way, the context is not lost through the coding process, as, it has been suggested, has often been the case (Bryman, 2015). This software also allowed me to code many extracts under different patterns or to tweak the codes relatively easily.

Phase 3: Searching for themes

Once the coding phase was complete, I started to look for themes. This is not simply opting for themes among those that already exist, but, as Clarke et al. (2015, p. 236) describe, it requires the researcher to 'create a plausible and coherent thematic mapping' of the data. In particular, this involved a lot of considerations to collapse, regroup, and sort different codes into potential themes. At this point, NVivo proved itself to be a helpful tool. I could use its mind-map function to 'play around' with codes to envisage their relationships. The criteria for identifying themes suggested by Braun and Clarke (2006) are that they not only tell you something about the data but also determine a coherent aspect of it. With those criteria clear in mind, I came up with several themes relevant to the research questions. Figure 4.4 below shows how I used the mind-map to initially group themes and key ideas.

Figure 4.4 *Initial mind map of themes and sub-themes*
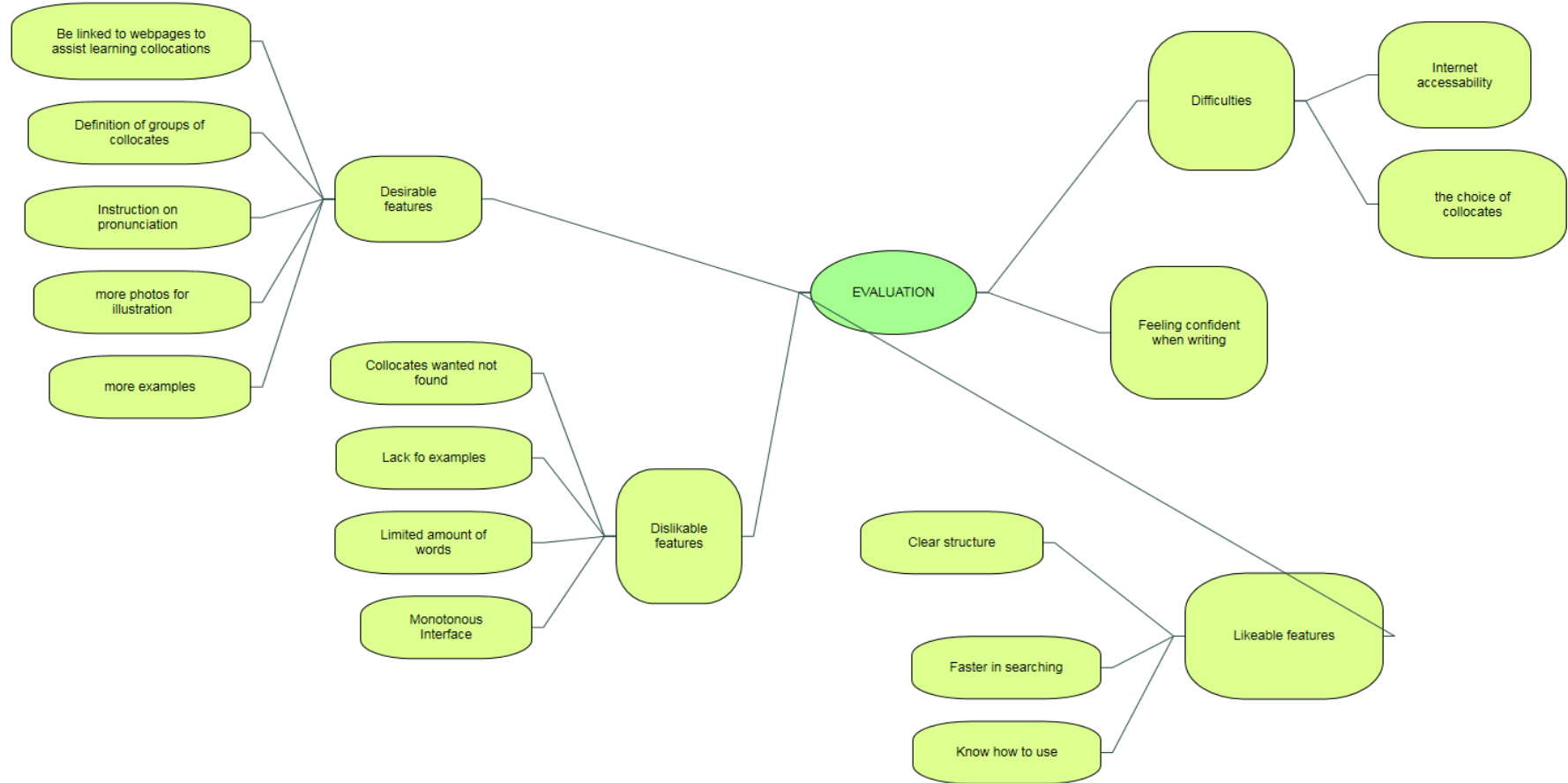
Phases 4, 5, 6: Reviewing, defining themes and writing up

Following the instructions of Braun and Clarke (2006), I reviewed the themes, which involved two levels: the level of the coded data extracts and the level of individual themes in relation to the whole data set. I then determined what the 'overall meaning' of each theme was and what aspect of the data it captured. However, it was not a straightforward process, but in fact involved producing several possible mind maps as well as renaming potential themes. After a lot of elaboration, I developed the mind map of four main themes presented below.

Figure 4.5 *Mind map of themes and key ideas*

- Frequency of use
- When using the dictionary
- Why using in combination with other dictionaries

1. How learners use the OOCD

- Usefulness
- Clear structure

2. What learners like about the OOCD

4. Suggested improvements

3. What learners dislike about the OOCD

- Definition for collocates
- Pronunciation
- Illustrative examples
- Images to illustrate
- Link to webpages

- Non-plentiful content
- Monotonous interface
- Difficuties in use

Chapter 5 : FINDINGS

5.1   Introduction

In the previous chapter I presented the analysis of students' written texts, questionnaire surveys, recording sheets, and interview data. In this chapter I will address the research questions based on combining the results found from different data sets. This chapter is divided into four sections corresponding to the four research questions, which I will be repeating at the beginning of each section. It is hoped that findings about how learners use the dictionary for collocation look-up, learners' collocation use with the dictionary support, and their perceptions of the use of the dictionary as a supportive tool, can provide deeper insights into the role and impact of the dictionary on learners' collocation use.

5.2   Research question 1

How do advanced learners use collocations?

The results of the analysis of the first set of written texts were data to answer this research question. Of the total 1,825 combinations extracted from this first set of essays, 1,290 combinations were identified as strong collocations, 415 as casual combinations, 2 as idioms, and 117 as odd collocations. The distribution of different types of combinations from each essay is summarized in Table 5.1 below. As can be seen in the table, the number of collocations used in an essay far outnumbers the number of casual combinations and idioms. On average, 45 collocations are used in one essay. The highest number of collocations in an essay is 64 (T23A) and the lowest is 20 (T19A). The highest number of odd collocations in a text was 12 (T11A), and some texts had no or just one odd collocation (T7A, T17A, T19A, T20A, T35A, T43A).

Table 5.1: *Distributions of different types of combinations of each essay*

N	TEXT	Strong collocations	Casual combinations	Odd collocations	Idioms
1	T1A	30	15	3	0
2	T4A	51	11	6	0
3	T6A	37	8	2	0

4	T7A	53	10	1	0
5	T10A	39	12	3	0
6	T11A	50	14	12	0
7	T12A	48	9	9	2
8	T13A	50	10	7	0
9	T14A	44	11	3	0
10	T15A	51	26	4	0
11	T16A	45	4	3	0
12	T17A	46	5	0	0
13	T19A	20	15	1	0
14	T20A	36	18	1	0
15	T21A	43	15	5	0
16	T22A	48	21	3	0
17	T23A	64	21	3	0
18	T24A	44	11	2	0
19	T25A	51	16	4	0
20	T27A	53	21	4	0
21	T28A	50	15	8	0
22	T31A	50	16	2	0
23	T35A	49	11	1	0
24	T36A	44	42	7	0
25	T38A	31	8	5	0
26	T40A	50	12	4	0
27	T42A	39	25	10	0
28	T43A	48	10	1	0
29	T44A	36	3	3	0

The collocations most frequently produced by at least three learners comprise:

- V-N: *help students* (11 occurrences), *have a chance* (10 occurrences), *go to a country* (6), *make students* (6), *make friends* (5), *bring advantages* (4), *face disadvantages* (4), *has advantages* (4). The rest of these collocations occurred three times: *adapt to the environment*, *affect the result*, *affect health*, *bring disadvantages*, *comeback hometown*, *earn money*, *eat (local) food*, *express*

ideas, *have difficulties*, *improve skill*, *learn a language*, *overcome difficulties*, *spend time*, *suffer shocks*.

- Adj-N: *foreign students* (12), *new environment* (9), *new country* (7), *developing country* (7), *new culture* (6), *new friends* (5), *part-time-jobs* (5), *different countries* (5), *different cultures* (5), *culture shocks* (4). Collocations occurring three times were: *daily life*, *financial problem*, *following reasons*, *good chance*, *high cost*, *high salary*, *long time*, *modern facilities*, *native people*,

- N-V: *students study* (15), *students face* (11), *students get* (10), *students learn* (9), *students live* (9), *students work* (6), *students go* (5). Collocations occurring three times were: *countries have*, *students overcome*, *students tend*, *countries have*.

- N-N: *culture shock* (10), *education system* (4), *language barrier* (4)

- Adv-V: *go abroad* (9), *live far away* (8), *think carefully* (3)

All of these collocations are undoubtedly frequent in native speaker English. They have fairly high Log Dice scores (>6), except for *students face* and *go abroad*; their Log Dice scores are 4.7 and 5.4, respectively. They were all used correctly except for one case in which a learner misused the preposition following the verb *adapt* e.g. **adapt with the environment* and another case in which a learner misused *culture shock* into *shock culture*. No collocations of N-of-N and Adv-Adj pattern occurred more than twice. 117 (8.4%) out of 1,407 collocations of all the seven patterns were identified as awkward.

5.2.1 Odd collocations

Collocations of V-N pattern are the most problematic to Vietnamese advanced learners (51% of all odd collocations). Odd collocations of Adj-N pattern are the second highest with 23 errors (19.7%), while collocations of N-of-N and Adv-Adj patterns have the least number of odd collocations equally, only 3 (2.6%) cases in each. Figure 5.1 below presents a distribution of odd collocations by their patterns.

Figure 5.1 *Distribution of odd collocations by patterns of the first set of written texts*

Odd collocations

Legend: V-N, Adj-N, N-N, N-V, Adv-V, N-of-N, Adv-Adj
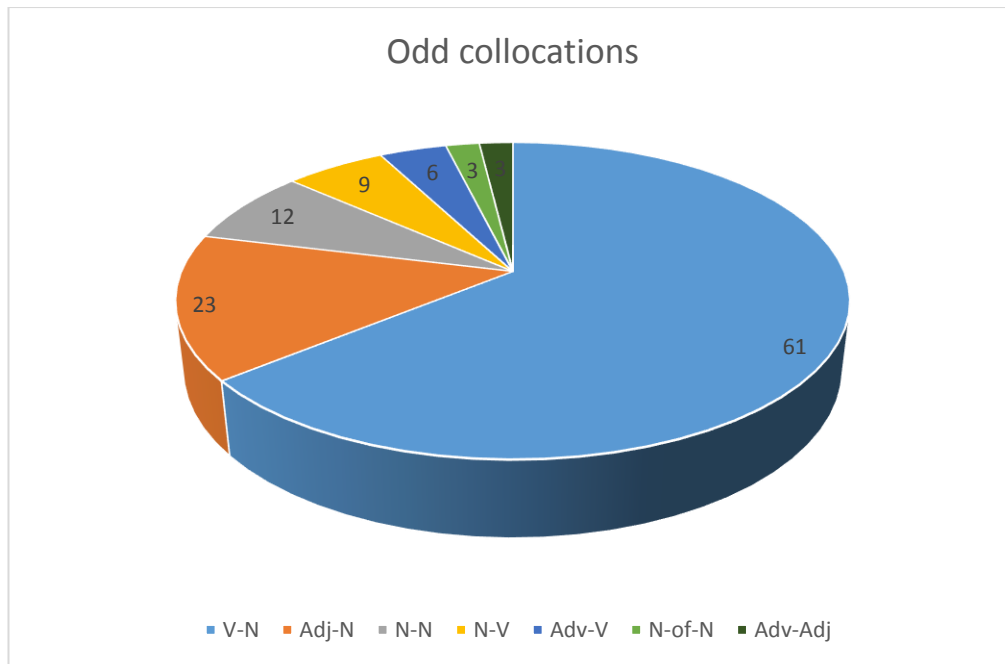
Table 5.2 below summaries the ratio of collocations used and odd collcations. V-N and Adj-N patterns have the highest number of odd collocations; the ratio of odd combinations over the number of collocations used is higher in collocations of N-N pattern (15.2%). Adv-Adj and N-of-N patterns have only 3 odd collocations equally, but the proportion of odd collocations over collocations used, 9.4% and 8.4% respectively, is higher than that of Adj-N collocations (6.3%). Collocations of N-V pattern have the lowest percentage of odd collocations over collocation used, at only 2.9%.

Table 5.2 *Ratio of collocations used and odd collocations*

	V-N	N-V	Adj-N	N-N	N-of-N	Adv-V	Adv-Adj
Number of collocations used	518	316	367	79	36	59	32
Number of odd collocations	61	9	23	12	3	6	3
Percentage	11.8%	2.9%	6.3%	15.2%	8.4%	10.2%	9.4%

Only two students produced the same odd collocations twice; they are *approach **to** a (new) culture/ language* (T1A) and *studying career* (T13A). No students made the same errors three times.

5.2.2 Types of errors

Findings from the learners' essays show that oddness can lie in any element of collocations. As can be seen from Table 5.3 below, oddness due to mismatching the base and the collocate are the most common in almost every collocation pattern.

Table 5.3 *Distribution of the types of errors in each collocation pattern*

Collocation patterns	Errors in collocating words	Errors not in collocating words	Inappropriate meaning
V-N	24	22	15
N-V	9	0	0
Adj-N	20	0	3
N-N	12	0	0
N-of-N	3	0	0
Adv-V	4	0	2
Adv-Adj	2	1	0

As for V-N collocations in this study, errors in collocating words lie mostly in the choice of verbs (17 out of 24 odd collocations) (see Table 4.31, page 137). Below are some examples of odd collocations due to the wrong choice of collocates in V-N collocations. Combinations after the arrows are expected collocations suggested by native informants.

dominate drawbacks → outweigh drawbacks

perceive knowledge → increase/acquire knowledge

raise students' independence → increase students' independence

shorten the gap → reduce/ minimize the gap

disturb their studying → impair their studying

It seems that although learners are familiar with *face/ deal with/ cope with a problem*, they do not recognize that *face* cannot be varied with *differences*. Simply combining *perceive*, meaning *come to realize or understand* (from OALD), with *knowledge* results in an awkward combination. The choice among near-synonyms (*raise ≈ increase, shorten ≈ reduce/minimize*) was problematic for learners as well.

With regard to collocations of Adj-N pattern, nearly half of the unacceptable collocations (11 out of 23 odd collocations) had errors in the Adjective collocates. In two of these cases, errors arose because learners either invented words (e.g. *malnutrious* from *malnutrition*) or did not remember the spelling of the word that they wanted to use. They are:

    malnutrious eating habit → *poor eating habit*

    psycholic disorder → *psychological disorder*

Odd collocations also resulted from learners over-generalizing the rule of forming Adjective modifying nouns by adding –ing to verbs. I found three cases in which errors are associated with this:

    learning career → *academic career*

    studying responsibility → *responsibilities of the studies*

    studying results → *study results/results of studying*

Errors in the collocates were also due to the derivational process. From the errors found, it seems that learners know the rule for forming compound adjectives but fail to add –ed to the ending element. Adding prefixes meaning 'not' to an adjective to create a contrastive word also confuses them. Some examples are:

    old-fashion minds → *old-fashioned minds*

    well-equipment classroom → *well-equipped classroom*

    unappropriate action → *inappropriate action*

The use of the collocate adjective *good* with very high frequency to express positivity does not always guarantee acceptability. *Good certificate* is an example. The use of this collocation shows that the learners either chose a safe solution or struggled to find an appropriate collocate to express their idea.

Errors in the collocates were also detected in other collocational patterns, such as N-V (7 out of 9 odd collocations), N-N (1 out of 12), N-of-N (2 out of 3), Adv-V (2 out 6). Adv-Adj is the only collocation pattern that did not have any unacceptable collocations of this type. Some examples of these odd collocations are:

    barrier causes (N-V) → *barrier results in/creates*

    the trend intrigues (N-V) → *the trend engenders*

    electricity fee (N-N) → *electricity costs*

large number of money (N-of-N) → large amount of money

affect strictly (Adv-V)→ seriously affect

learn seriously (Adv-V) → learn effectively

A wrong choice of collocates that have similar meanings again results in odd collocations, e.g. *cause ≈ results in/creates, fee ≈ costs*. In N-N collocations, learners construct them by combining equivalent nouns regardless of language convention, and this unfortunately results in awkward collocations. The confusion between quantifier nouns *number* and *amount,* one accompanying countable nouns and one uncountable nouns, is another cause of the error. There is only one error associated with this, though. The construction of collocations based solely on semantic meaning and syntactic knowledge of the target language causes odd collocations of Adv-V pattern as above.

Errors were also found in the base as well as in the collocates. Once the choice of a base is inappropriate in a particular context, the combination as a whole is often inappropriate. As in the below example, the use of the noun base *stuffs* was deemed inappropriate in the context. Hence the combination *suffer stuffs* as a whole was incorrectly used in meaning in the context and the native informants suggested it be changed to *problems*.

(T42A) *Students' goal is focusing on studying while in reality it often takes them time to adapt to a very new environment. They have to ***suffer stuffs** related to culture differences. (→ suffer problems)*

As can be seen from Table 4.31, errors in the base are detected in some collocational patterns: V-N (3 out of 24 errors in collocating words), N-V (1 out of 9), Adj-N (3 out of 20), N-N (3 out of 12), Adv-Adj (1 out of 3). These are some examples:

suffer stuffs (V-N) → suffer problems

chance (of work) is waiting (N-V) → a strong likelihood of work is waiting

advanced way (Adj-N) → advanced form

study career (N-N) → academic career

more opened (Adv-Adj) → more open

As for collocations of V-N, N-V and Adj-N as in the examples above, oddness is caused by the inappropriate choice of bases, whereas collocations of N-N and Adv-Adj in the last two examples are different. When constructing *study career,* learners

simply combined the two nouns based on their semantic meanings. *Study* and *academic* have a similar meaning in Vietnamese, but the choice of *study* instead of *academic* might be because *study* occurs much more frequently than *academic* (39,738 and 5,581 times on the BNC, respectively), and hence is the word that first come to the learners' minds. Errors such as **more opened* show that learners are aware of the suffix-adding rule to transform a verb into an adjective; however, they failed to recognize that *open* is itself an adjective.

Errors in the combination as a whole is the last subtype of errors in collocating words. For these collocations, the whole combination is incorrect rather than the base or the collocate individually. There were eight cases in which native judgements did not give specific suggestions on how to correct them. Most of these did not co-occur on the BNC except for *expense + cost* and *appoint + level*, but they are not in the rigid order that I was investigating. As Table 4.31 shows, odd collocations of this type were detected in V-N (4 collocations), N-V (1), Adj-N (6), N-N (8), N-of-N (1), Adv-V (2), Adv-Adj (1). Examples of the unacceptable combinations are:

*appoint to (higher) level (V-N) → No suggestion

*the loop will keep on (N-V) → *the cycle will go on*

*rewarding trend (Adj-N) → No suggestion

*expense cost (N-N) → No suggestion

*their need of studying (N-of-N) → *their study requirements*

*study safely (Adv-V) → No suggestion

*definitely valuable (Adv-Adj) → No suggestion

The confusion between N-N and N-of-N construction led to odd collocations in a few cases. **Their need of studying* and **life standard* were used while *their study requirement* and *standard of living* were expected. These examples also show that learners construct combinations rather freely, ignoring the conventions of the language.

Errors not caused by the wrong choice of collocating words are mostly due to the incorrect use of prepositions (22 out of 23 errors). They were mostly found in collocations of V-N pattern. There are potentially three different kinds of errors with prepositions. I found examples of all three:

(1) Learners added prepositions where they were not required:

*approach **to** a new language → approach a new language

*prepare **for** their knowledge → prepare their knowledge

(2) Learners omitted prepositions where they are required:

*bring students a new environment → bring students to a new environment

*suffer homesickness → suffer from homesickness

(3) Learners misused prepositions:

*adapt **with** the environment → adapt to the environment

*disagree **to** the idea → disagree with the idea

Learners also have problems with prepositions when using the verb *face* (5 out of 11 times that *face* was used). It seems that when there are two combinations relating to one word that could be used to express a similar idea (e.g. *face sth* and *be faced with sth*, *concern someone* and *someone be concerned about*), they tend to blend them as in the sentences below:

(T38A) *In conclusion, besides a lot of benefits that studying abroad brings about, we must ***face with many difficulties***. (→ face many difficulties)*

(T40A) *Besides many advantages, students can also ***face with many drawbacks***. (→ face many drawbacks/ are faced with many drawbacks)*

(T28A) *The very first problem that ***students*** normally ***concern*** is language barrier. (→ students are concerned about / concern students)*

Oddness is not just due to the wrong combination of the elements but to the inappropriate use of their meanings in context. Odd collocations of this type were found more in the V-N collocations (15 collocations) than in other collocational patterns such as Adj-N or Adv-V, with only 2 errors equally. The problem with this error type might be because learners do not know the precise meaning of collocations to use appropriately in context. Combinations such as *normal students*, *live normally, have drawbacks, come back ... hometown, or enlarge their knowledge* are correct collocations, but according to native informants they were not used semantically appropriately in the following sentences. *Normal students* in (T27A) has been used in contrast with students studying abroad, suggesting that the student writer used it to

refer to students studying in their home country. As such, the combination *domestic students* or *traditional students* should be the combination of choice.

> (T27A) *Secondly, students who study abroad have more experience than a* ***normal student**. (→ traditional/domestic student)*
>
> (T25A) *When a person cannot **live normally** with others, that one may be led into stress. (→ live peacefully)*
>
> (T38A) *Students studying abroad **have some drawbacks**. (→ face/experience some drawbacks)*
>
> (T12A) *Students do not only suffer culture shock when they are studying in a foreign country but also after they **come back to their hometown**. (→ return their hometown)*
>
> (T6A) *It is widely believed that studying abroad is a great opportunities for students to **enlarge their knowledge**. (→ increase/improve their knowledge)*

Another way to approach the question of which collocations are particularly difficult for the group of learners under investigation is to find combinations inappropriately produced by more than one learner. In this respect, collocations involving these nouns are problematic: *knowledge* (5 odd collocations used by 4 learners), *drawback* (4 by 4L), *environment* (4 by 4L), *independence* (2 by 2L), and *hometown* (4 by 4L), *problem* (2 by 2L), *time* (2 by 2L).

Failure to use collocations semantically appropriately in context could also be because learners wanted to opt for a safe solution. The use of the verb *have* in the sentences below is an example. *Have* is a very high frequency word and can be accompanied by a wide variety of nouns with the meaning *to possess/own* or *to undergo/experience*. The use of *have* in those combinations is unfortunately incorrect in these particular contexts. *To get a part-time job* and *to face/experience some drawbacks* are expected to be used in these contexts instead.

> (T4A) *The cost of studying abroad is really high, so students have to try their best to get scholarship and they also **have a part-time job** to support their living.*

(T36A) *In conclusion, students studying abroad* ***have some drawbacks*** *relevant to cultural shocks and health conditions, so it's necessary for them to have a good preparation.*

The confusion of *make* and *do*, *take* and *make*, while frequent in the early and intermediate stages of learning (Nesselhauf, 2005; Koya, 2003) was not found to be troublesome to language use of more advanced learners in this study. In particular, of the 16 times combinations containing *make* or *do* occurred, there were no cases where learners confused them. What appears as a particular problem in the present data instead is the confusion of verbs of direction - *come back/go back/go outside* for *return* or *leave* (in the four out of six times these verbs were used). These combinations are not wrong but are slightly awkward; they were categorized into inappropriate meaning errors. Some of these collocations read as follows in the students' essays:

(T12A) *Moreover, students do not only suffer cultural shock when they are studying in a foreign country but also after they* ****come back*** *their* ***hometown***.

(T21A) *In conclusion, I believe students should try to* ****go outside*** *of their* ***hometown*** *or motherland to learn if they have the chance.*

(T24A) *People who study in their own country can still live independently; however, they tend to* ****go back hometown*** *with their families in some occasions.*

(T44A) *Last but not least, studying abroad gives students more opportunities to get a good job when they* ****come back*** *their* ***hometown***.

Quantitatively, of the 15 times combinations with *come* and *go* were used, 4 were incorrectly used. It seems that at this level learners still get mixed up with when to use *come* and when to use *go*. This might be because these verbs have nearly the same basic meaning. What tends to direct the learners in their use of these verbs, however, is the preposition following the verbs, *back* or *to*, rather than the verbs themselves. *Back* means *trở lại*; *to* means *đến*, which show directions in Vietnamese. They are inclined to use either *come back* or *go back* to show movement towards the speaker or the person spoken to and *come to* or *go to* for away movement. When scrutinizing cases where these verbs are used correctly, I found that they are cases

when either *come* or *go* can be used depending on whether the speaker sees things moving towards or away from another person or place. As in the two examples below, *go to* is used when the writers see things from the '*students''* viewpoint while *come back* is used when they see things from the '*home*' position.

> (T40A) *When **going to** another country, students have a good chance to widen their knowledge*.

> (T24A) *Additionally, students just **come back** home on Tet holiday or when they finish their courses*.

5.3   Research question 2

How do students use the OOCD to support their writing?

I tackled this question by combining findings from analysis of the survey, the recording sheets and the interviews presented in chapter 4. The answer to this question is presented in two sections: (1) when learners approach the dictionary for help and (2) how they use the dictionary to look for collocations.

5.3.1   When learners approach the dictionary for help

The survey data showed that quite a lot of participants (69.7%) often used the OOCD as an assisting tool whenever they do a piece of academic writing; more than one-quarter of them (27.3%) only occasionally used it, and only one participant has never used it since it has been introduced. In response to this same question in the interviews, seven out of eight replied that they use the dictionary every time they write, but the extent of their search depends on the individual writing task. Most of them only use it to look for collocates when they are doing academic writing, except for one who added that she was in the habit of learning new vocabulary, so the OOCD was a good source of collocations to learn from. The one informant that did not use the dictionary frequently supposed that collocations can be found in any general dictionary. When asked how often she uses the OOCD, she said:

> **H**: *not very often, usually I only check meanings of a word and in an English-English dictionary it has some collocations already in there. Only when I can't find collocations that I need then I will go to check further in the OOCD.*

Participants sometimes approach the dictionary not because they do not know collocations to express an idea but to look for a different way of expressing it to

avoid repetition. Another wise reason for consulting the dictionary is to look for a hint for an idea rather than a collocation to express some intended idea. They shared:

> *E: It (the dictionary) helps me avoid repetition in my writing. For example, when I want to find the verb for success, I can use achieve, obtain, or have … there are a lot of collocates that I can use.*
>
> *G: It (the dictionary) helps me prevent repetition. Sometimes I search collocate of a word not because I don't know the collocations to express my idea but to look for another way of saying it.*
>
> *H: When I think the word (collocate) that I use is so common, I check the dictionary for another way of saying it. I sometimes search the dictionary for a hint rather than looking for a word to express an idea that I have already had in my mind.*

Findings from the recording sheets and the interviews suggest that most of the look-ups (94.4%) were done while students were doing their writing. In other words, learners tended to approach the dictionary for help immediately when need be rather than at the end when the writing was finished. In particular, as recorded in the recording sheets, 119 out of 126 look-ups were carried out while students were doing their writing. Three of the participants from the interviews shared that looking for collocations immediately when they get stuck is their habit. They said:

> *E: usually when I have difficulty with collocations, I will check it right away. That's my habit.*
>
> *F: I check collocations from the dictionary in conjunction with writing for both in-class and at-home writing. (translated from Vietnamese)*
>
> *D: I don't use the dictionary to support whenever I write. In fact, I only use it when writing something academic. As for free writing, I don't use it. I'm kind of lazy and don't reread my writing after finishing, so I often consult the dictionary immediately when I get stuck. Using it to look for collocations right away reassures me. (translated from Vietnamese)*

One student reported that it is hard for her to get rid of the habit no matter whether it is beneficial or not. This reason seems to be closely associated with a feeling of certainty that was shared by most of the participants. It helps them feel confident that what they have written is correct and complete. One participant said:

> G: *I often consult the dictionary immediately when I get stuck. Using it to look for collocations right away reassures me.*

The interview data show that the time constraint for in-class writing is another reason why they tend to consult the dictionary while writing. Having no spare time for drafts induces them to search for help to complete every sentence of the writing. Two of the interview participants also shared that they only search at the end of the writing for collocations in which one of the elements is optional, such as Adj-N, Adv-Adj, or Adv-V, and this was to add something or to check if combinations they had used were correct.

Findings from the analyses of the recording sheets also show that on average learners use the dictionary 4 times for looking up collocations when doing the writing. The highest number of check-ups in an essay was 8 and the lowest was 1. More check-ups (52.8%) were done to look for collocates to construct collocations than for checking if the collocations they intended to use were correct (43.8%). They approach the dictionary for help with collocations of Adj-N pattern the most (40.5%). 27 out of 126 check-ups (21.4%) were of V-N collocations. For this most problematic collocation pattern, this number of learner consultations is quite low. There are in total only 4 look-ups for the N-N pattern, whilst errors of this pattern occur quite a lot (14 and 12 in the first and the second set of essays, respectively).

5.3.2   How learners use the dictionary to look for collocations

The learners did not have problems with the alphabetic search, which is a skill needed to use paper dictionaries effectively (Koren, 1997). The interview data show that to look for a collocate to complete the intended phrasal meaning, after typing in a base word, learners scan through the list of collocates provided quickly. As two of the participants shared in the interviews, this is quite an easy step since collocation patterns are all set in red capitalised letters. Collocates of similar meaning are grouped together and are in bold. They can quickly locate the position of words of some particular part of speech. One said:

> *H*: *if I search for verb collocates of a noun … first of all I have Adjective-Noun, and then Verb-Noun. It is presented in a fixed order. Collocation patterns are in red, so I can easily know the order of collocates.*

When asked if they use the OOCD in combination with other dictionaries, all informants responded that they often used it together with other dictionaries - an English-English, English-Vietnamese dictionary or a thesaurus to look for meanings of collocates. This seems consistent with the survey data, with 66.7% participants responding thus. Data recorded from the recording sheets nevertheless shows a quite different picture. In only 12 out of 126 look-ups (9.5%) do learners use this dictionary together with other dictionaries. This is understandable since their responses from the interviews and surveys were just their general estimations without considering some factors such as topic of the writing, the kind of writing (academic or free writing), and writing conditions (with or without time constraint). As one participant shared, due to the limited time span of this writing, when searching for collocates of a word in the OOCD, she often opted for a collocate that she already knows rather than consider choosing other collocates that she did not know. When doing assignments at home, in contrast, to enhance the writing she often considered choosing 'strange words' (collocates) after searching for their meanings from other dictionaries. Another informant shared that, when time allowed, in order to avoid repeatedly using some combination she used a thesaurus to look for another way of expressing the same idea. She said:

> **G**: *actually I often use it (the OOCD) with a thesaurus. Like … when I have a word in mind, I look for its synonyms and then look for collocates of a certain synonym (from the OOCD) and make a comparison to see if those collocates can be used with both words and then come up with a combination for my writing.*

Having had this same idea, one participant shared on the recording sheet that:

> **RS 25**: *I don't want to use repeatedly the same collocation and because the dictionary only provides a limited number of collocates of a headword, I looked for synonyms of a collocate from thesaurus.*

This 'creative' strategy is somewhat risky since obviously combining words based on synonyms of collocates suggested by the OOCD might lead to an unacceptable combination. Synonyms of words do not necessarily convey exactly the same meaning, so this strategy might lead the learners astray by opting for a synonym that is not appropriate in a certain context (East, 2008, p. 20). Another participant expressed her worry that searching for meanings of collocates from other dictionaries

could distract her from her writing and decided not to use the OOCD in combination with other dictionaries unless it does not provide her with a combination that she wants:

> *A: when I write and look up a word in this dictionary (the OOCD), I try myself not to look at another one because I'm afraid that I might get into, you know … sometimes when I look up a word in a normal dictionary, I shift from word to word because I find it quite interesting and fascinating to me. So I'm afraid that if I use this dictionary with another dictionary I might get distracted from my writing. But sometimes when I could not find a suitable collocate, I have to use an English-English dictionary.*

Flow of thought is deemed another factor affecting how learners use the dictionary. When writing with new ideas continuously coming to mind, they will leave a blank and then go back to find words to complete the idea. Conversely, if they have not come up with what to write next, they will consult the dictionary to find collocates to accomplish the phrasal meaning. Some even shared their 'strategies' for dictionary use:

> *G: If I have no time for a draft and I don't know which word can collocate the word that I want to use, I will leave a blank there and check it later. If I have no time for a draft and have to write straight down, I will check the word immediately when getting stuck.*
>
> *C: Formerly I looked for collocations while writing. Now I try to avoid that habit since it sometimes distracts my writing. I train myself to use the dictionary after finishing the writing. I reread it and use the dictionary to add collocates where possible. (translated from Vietnamese)*
>
> *H: I use the OOCD while writing. Well actually it depends on the flow of thought. If at that time no new ideas come to mind, I will stop to look up collocations from the dictionary to complete that sentence. If, however, the flow of thought continues, I will leave a space and come back to search for a word to fill out later. It is often with collocations of which one element is not required like Adj in Adj-N, Adv in Adv-Adj and Adv in Adv-V collocations. In this way, I can avoid being distracted.*

5.4   Research question 3

Does the use of the OOCD aid learners to use collocations correctly in writing?

The answer to this question was based on the comparison of the results of collocation use in the first and the second set of essays, with reference to recording sheets to

trace back collocations looked up in the dictionary. It is divided into two sections. In the first section, I present findings in terms of the appropriate use of collocations based on the comparison of:

- ○ the number of odd collocations from the two sets of essays;
- ○ the number of odd collocations from the first and second essay of each student;
- ○ the ratio of odd collocations over the number of collocations used in the two sets of essays;
- ○ types of combinations used (strong collocations, casual combinations, and idioms) to see if there are any changes in the number of collocations used when learners were allowed to seek help from the dictionary;
- ○ the number of strong collocations with very high Log Dice score (>7) in the first and second essay of each learner;
- ○ the distribution of types of odd collocations (oddness in collocating words, oddness not in collocation words, and inappropriate meaning) in the first set and second set of essays;

In the second section, findings in terms of the variety of collocation used in the two sets of essays are then discussed.

5.4.1   Appropriate use of collocations

Findings from the analysis of the second set of essays, regardless of whether or not collocations were consulted in the dictionary, show that learners' collocation use in general did not improve. Counter to my expectation, the total amount of odd collocations in the second set of essays did not decrease but increased, with 17 collocations more than before the intervention. As can be seen in Figure 5.2 below, all types of odd collocations found in the first set of essays were found in the second set. Collocations of V-N patterns are still collocations with the highest amount of odd collocations (62 collocations). Adj-N collocation patterns have the second highest amount of odd collocations (31 collocations), 8 collocations more than the first set of essays. Odd collocations of Adv-V and Adv-Adj types were slightly reduced.

Figure 5.2: *Number of odd collocations in the two sets of essays*
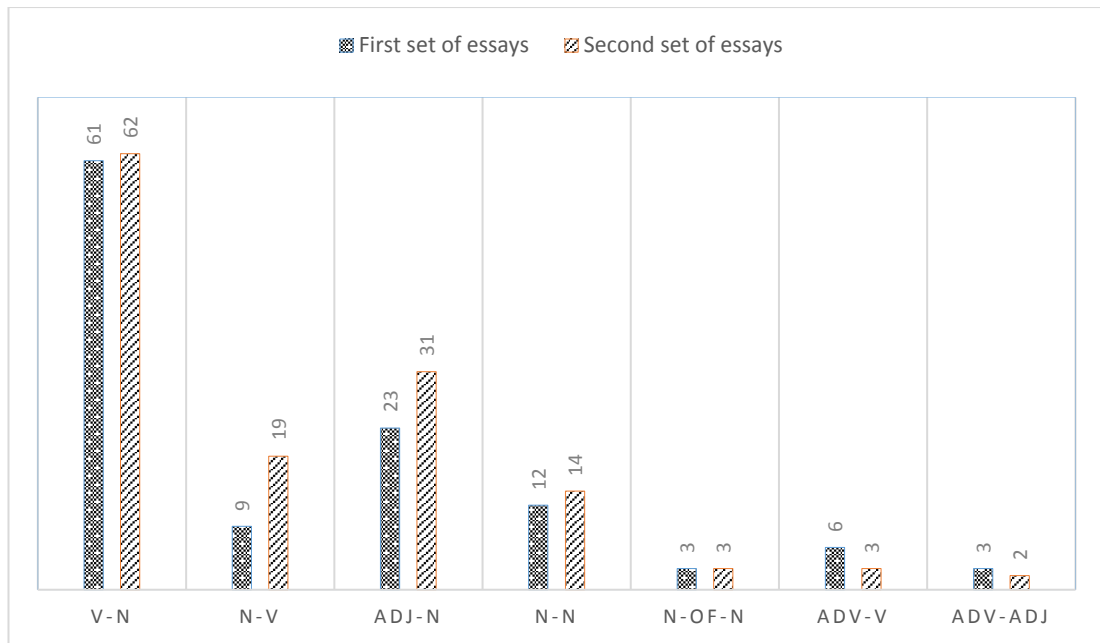
The increase in the number of odd collocations when writing with the OOCD support compared to without initially suggests that the OOCD does not help learners reduce odd collocations. As can be seen in Table 4.36 on page 152, Adj-N is the collocation pattern that learners looked up the most (51 out of 126 look-ups); it is, unexpectedly, the collocation pattern with odd collocations increasing substantially (8 collocations). However, in order to better understand whether the OOCD helps learners in improving collocational use, I need to compare written texts (without and with the OOCD support) of individual participants. In Table 5.4 below, I summarise the number of odd collocations of all the written texts in pairs. If we categorize essays with three or more odd collocations fewer than in the first set of essays as improvement of collocation use, those with two odd collocations fewer as slight improvement, the number of odd collocations remaining the same or just reducing by one as no improvement, and the number of odd collocations increasing as worse, we have the results presented in Table 5.4 below. Columns 2 and 3 present the number of odd collocations in each pair of writing tasks, while the last columns give a general assessment of improvement.

Table 5.4 *Comparison of odd collocations without and with the OOCD support*

Written texts	Number of odd collocations (- dictionary)	Number of odd collocations (+ dictionary)	Number of look-ups	Assessment of improvement

T1A - T1B	3	3	7	No improvement
T4A - T4B	6	5	3	No improvement
T6A - T6B	2	0	3	Slight improvement
T7A – T7B	1	0	7	No improvement
T10A – T10B	3	2	6	No improvement
T11A – T11B	12	3	7	Improvement
T12A – T12B	9	6	6	Improvement
T13A – T13B	7	20	5	Worse
T14A – T14B	3	1	4	Slight improvement
T15A – T15B	4	5	5	Worse
T16A – T16B	3	3	2	No improvement
T17A – T17B	0	1	1	Worse
T19A – T19B	2	5	4	Worse
T20A – T20B	1	6	2	Worse
T21A – T21B	5	7	4	Worse
T22A – T22B	3	6	6	Worse
T23A – T23B	3	10	4	Worse
T24A – T24B	2	4	7	Worse
T25A – T25B	4	8	5	Worse
T27A – T27B	4	2	8	Slight improvement
T28A – T28B	8	6	1	Slight improvement
T31A – T31B	2	2	1	No improvement
T35A – T35B	1	4	2	Worse
T36A – T36B	7	11	7	Worse
T38A – T38B	5	3	2	Slight improvement
T40A – T40B	4	3	1	No improvement
T42A – T42B	10	6	6	Improvement
T43A – T43B	1	1	6	No improvement
T44A – T44B	3	1	4	Slight improvement

From the above table, it can be seen that the highest number of odd collocations in a text was 20 (T13B), much higher than the figure in the first group, 12. Only 3 out of 29 learners (10%) showed clear signs of collocational improvement, with the number of odd collocations reducing by three or more. These students had carried out a number of dictionary look-ups, except for one case in which the student did only one look-up. Six learners (20%) improved slightly, with two odd collocations fewer than in the first piece of writing. Almost a quarter of the students did not improve their collocation use. 12 students (41%) performed worse, with the amount of odd collocations increasing. There was a noticeable case with the number of odd collocations in the second writing more than double that in the first piece of writing - 20 and 7 respectively.

Four out of 29 learners looked up collocations in the dictionary just once, and the comparison of collocation use in these students' essays brought up different results. In particular, one essay improved slightly; two did not improve; one got worse. Scrutinizing the essays for which learners sought help from the dictionary a number of times (7 or 8), I found that the results of collocation use in those essays are quite similar to those in which learners only searched once, in that results spread evenly from improvement to getting worse. The essay with 8 searches shows a slight improvement, whereas for those essays with 7 search times one improves, two do not improve, and three get worse. A Pearson correlation coefficient was also computed to assess the relationship between the number of look-ups and the number of odd collocations used in the essays. The result showed that there was statistically significant relationship between the two variables ($r = .107$, $p = .579$). This means that the number of dictionary consultations seems not to be related to the effectiveness of the production.

Comparing the ratio of odd collocations over the number of collocations used in the two sets of essays in Table 5.5 below, I found that N-N collocation is still the collocation with the highest ratio of odd collocations over collocations used (16.9%). Adj-N collocation is the third highest (9.4%), higher than in the first set of written productions despite being the most searched-for collocation pattern − 54 out of 126

look-ups. The proportion of odd collocations over collocations of V-N patterns is still the second highest, with the number of odd collocations approximately the same as that of the first set of writing - 62 and 61 respectively.

Table 5.5: *Ratio of odd collocations over collocations used in the first and second set of essays*

Collocation types	First set of essays			Second set of essays		
	Total	Oddness	Percentage	Total	Oddness	Percentage
V-N	518	61	11.8%	452	62	13.7%
N-V	316	9	2.9%	313	19	6.1%
Adj-N	367	23	6.3%	330	31	9.4%
N-N	79	12	15.2%	83	14	16.9%
N-of-N	36	3	8.4%	45	3	6.7%
Adv-V	59	6	10.2%	54	3	5.6%
Adv-Adj	32	3	9.4%	41	2	4.9%

Comparing the numbers of collocations, casual combinations, and idioms used in the two sets of essays (see Figure 5.3 below), I found that, contrary to what I had hypothesized, the number of collocations in the second set of essays decreased while casual combinations increased. This suggests that even with the availability of the dictionary learners did not use more collocations than without.

Figure 5.3: *Distribution of combinations from the two sets of essays*

Distribution of combinations from the two sets of essays
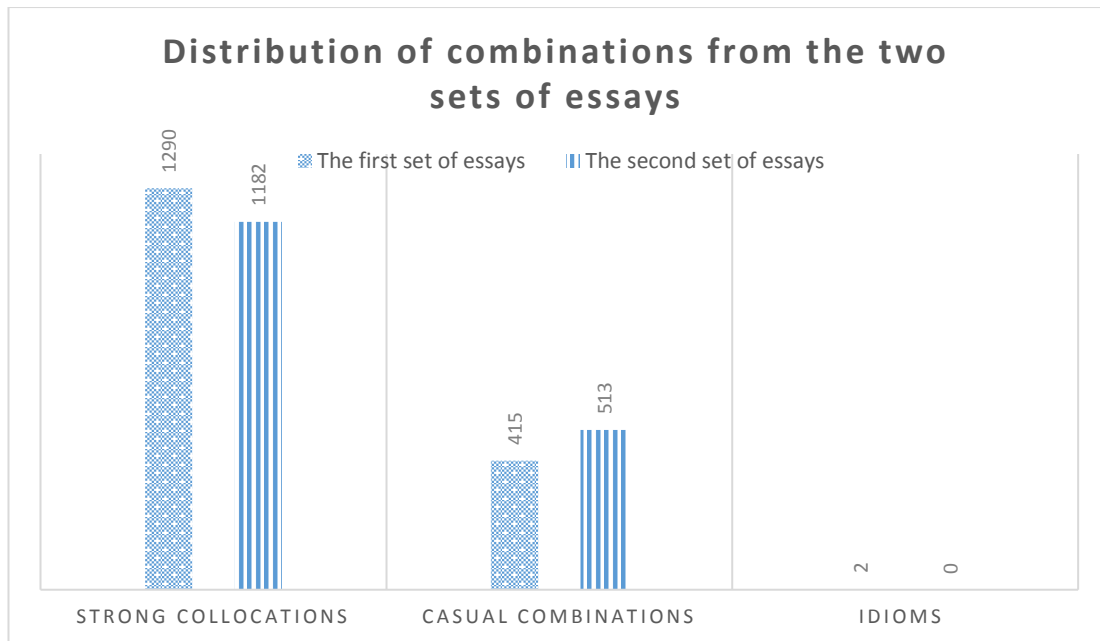
In an attempt to thoroughly understand learners' collocation use without and with the intervention, I counted and compared the number of collocations with a very high Log Dice score (>7) in all the learners' first and the second essays. Results presented in Table 5.6 below show that the number of strong collocations with a very high Log Dice score in the second set of essays increases considerably, with 60 collocations more than in the first set. However, the number of collocations used as a result of the OCCD check-up is only eight. 14 out of 29 students used more collocations (Log Dice > 7) in the second essay than in the first, but only two of these collocations were looked up in the dictionary. This suggests that the dictionary did not impact the use of collocations either.

Text	Strong collocations in the first essays	Strong collocations in the second essays	Collocations looked up from the OOCD	Text	Strong collocations in the first essays	Strong collocations in the second essays	Collocations looked up from the OOCD
T1	19	17	0	T22	19	17	0
T4	13	7	0	T23	20	20	2
T6	11	18	0	T24	15	17	0
T7	19	11	2	T25	12	20	0
T10	10	8	0	T27	15	12	0
T11	18	6	1	T28	18	8	0
T12	19	17	0	T31	15	18	0
T13	11	25	0	T35	17	24	0
T14	11	17	0	T36	12	16	0
T15	21	16	0	T38	13	18	0
T16	13	13	0	T40	8	8	0
T17	11	17	1	T42	6	19	0
T19	4	16	0	T43	13	19	0
T20	10	19	0	T44	16	10	1
T21	16	23	1				
				Total	**406**	**466**	**8**

Table 5.6: *Number of strong collocations (Log Dice >7) in the first and the second essay of each student*

Table 5.7 below presents the distribution of types of odd collocations in the second set of essays. It can be seen that similarly to the first set of essays, the wrong choice of collocating word is the most common error in collocations of V-N and Adj-N patterns, and is the only type of odd collocation in the other patterns.

Table 5.7: *Distribution of the types of odd collocation in each collocation pattern*

Collocation patterns	Errors in collocating words	Errors not in collocating words	Inappropriate meaning
V-N	39	13	10
N-V	19	0	0
Adj-N	28	0	3
N-N	14	0	0
N-of-N	3	0	0
Adv-V	3	0	0
Adv-Adj	2	0	0

Collocates, verbs in V-N, N-V and adjectives in Adj-N combinations, are still the most problematic element to the learners although they could consult the dictionary for help. In particular, 27 odd collocations of V-N pattern, 11 of N-V, and 14 of Adj-N involved the wrong choice of collocates (see Table 4.32). Some examples of these odd collocations are:

> *bring stress* (V-N) → *cause stress*
>
> *remain relationship* (V-N) → *maintain relationship*
>
> *stress formed* (N-V) → *stress arose*
>
> *stress covers* (N-V) → *stress overwhelms*
>
> *bad chemicals* (Adj-N) → *harmful chemicals*
>
> *inside factors* (Adj-N) → *internal factors*

Odd collocations due to over-generalizing the rule of forming Adjective modifying nouns by adding –ing to verbs or derivational processes were still found. However, it is important to note that none of these were looked up in the dictionary:

> *working burdens* → *employment burden*

> **breaking days* → *days away from work*
>
> **hurry pace* → *hurried pace*
>
> **multitask person* → *multitasking person*
>
> **concerned issues* → *concerning issues*
>
> **caffeine-contained substances* → *caffeine-containing substances*

The use of the collocate adjectives *good* and *bad* with very high frequency to express positivity or negativity again resulted in unacceptable combinations. **Good status* and **bad chemicals* are examples. This proves that the choice of a 'safe' solution is not always safe. The OOCD search of the noun bases *status* and *chemical* brings about quite a lot of adjective collocates of the two nouns: *great, high, superior/inferior, low, lowly* and *pure/organic/inorganic, synthetic/dangerous, harmful, hazardous, poisonous, toxic/agricultural, household, industrial*. If they had searched the dictionary, they could have chosen appropriate collocates (e.g. *great, high, superior* in combination with *status,* and *dangerous, harmful, hazardous, poisonous, toxic* in combination with *chemicals*) to express the intended meanings.

Similarly to the first set of essays, in the second set errors in the collocates were also detected in other collocational patterns: Adv-V (all 3 odd collocations), and Adv-Adj (all 2 odd collocations) (see Table 4.32). The construction of collocations based solely on sematic meaning and syntactic knowledge of the target language causes these odd collocations. Some examples of these are:

> **tax greatly (Adv-V)* → *tax heavily*
>
> **significantly polluted (Adv-Adj)* → *extremely polluted*

Odd collocations at the base were also found in collocations of V-N (10 out of 39 odd collocations), N-V (2 out of 19), Adj-N (10 out of 28), and N-N pattern (4 out of 14) in the second set of essays. The confusion of words that have similar forms (e.g. *status* and *state*) or that are associated with each other in some way (e.g. *timeline* and *time*, *collar* and *white-collar*) led to the construction of odd collocations:

> **arrange their timeline* (V-N) → *arrange their time*
>
> **reasons create* (N-V) → *factors causes*
>
> **unhealthy status* (Adj-N) → *unhealthy state*
>
> **collar workers* (N-N) → *white-collar workers*

Odd collocations as a whole were also found in the second set of essays. There were six cases in which native judgements did not give specific suggestions on how to correct them. Odd collocations of this type were detected in V-N (2 collocations), N-V (6 collocations), Adj-N (4 collocations), N-N (10 collocations), and N-of-N (3 collocations). Examples of the unacceptable combinations are:

trouble with relationship (V-N) → result in relationship troubles

(technological) advancement has outshone → no suggestion

breaking days (Adj-N) → leisure time/days away from work

soil sources (N-N) → no suggestion

state of emotion (N-of-N) → emotional state

The confusion between N-N and N-of-N construction was also observed in the second set of essays. *Living style, *life quality, *stress of relationship* were used where *style of living, quality of life, and relationship stress* were the expected collocations.

Similarly to the first set of written texts, in the second set odd collocations not in collocating words are associated with prepositions (all 13 collocations). Learners also added, omitted, or misused prepositions. These oddnesses were only found in collocations of V-N pattern. Some examples of these odd collocations are:

explain for the phenomenon → explain the phenomenon

dealing deadlines → deal with deadlines

cope under pressure → cope with pressure

Learners also have problems with prepositions when using the verb *face.* Of the seven times *face* occurs, four times they were used incorrectly. Of two students using *face* in both the first and the second essays, one student used it correctly and one incorrectly in both essays.

Oddness due to inappropriate meaning of collocations in a particular context was also found in the second set of essays: namely 10 odd collocations of V-N and 3 Adj-N. Some examples are:

train skills (V-N) → develop skills

limit overpopulation (V-N) → combat overpopulation

sufficient nutrition (Adj-N) → well-balanced diet

If in the first set of essays errors pertaining to the inappropriate use of the verb *take* did not occur, they were detected in the second set. There is one case in which a learner confused *take* and *have*; **take a (comfortable) chat* was used while *have a (comfortable) chat* was required instead. In another case *take* was mistakenly used when another verb was appropriate, e.g. *withstand the stress*. As Table 4.32 shows, the most prominent errors are those associated with the noun *stress*. Of a total of 53 collocations containing *stress*, eight odd collocations are of V-N pattern and six N-V pattern. It is important to notice that none of these combinations were looked up in the dictionary.

Comparing odd collocations in the first and the second essays of each learner, I found that most of the learners made errors with different base words. One of the learners constructed combinations containing the noun base *disadvantages* awkwardly in both essays: *give ... disadvantages* (T22A) and *provide ... disadvantages* (T22B). Another learner produced odd collocations of Adj-N pattern by forming Adjectives modifying nouns from verbs: **studying career* (T13A) and *working burdens* (T13B). Another learner used combinations containing the verb *face* incorrectly in both essays: **face with drawbacks* (T40A) and **face with difficulties* (T40B); he/ she is also the one who used the same odd collocation **temporary world* in the two essays. Again, none of these odd collocations were looked up in the dictionary.

Results from recording sheets showed that 126 out of 144 look-ups were for collocation purposes. Almost all of the collocations looked up in the dictionary were used correctly. There were only two cases where collocations were not used appropriately in terms of meaning. They were:

T21B *In this way, we can avoid the negative consequences and ***boost** the positive **results**. (→ **encourage** the positive results)*

T12B *Adults ***retain** a tremendous **responsibility** in caring for their children and teaching them. (→ **maintain** a tremendous responsibility)*

This suggests that the dictionary would be able to aid learners to use collocations correctly in their writing as long as they were aware of the collocations that pose problems in order to approach the dictionary for help.

5.4.2   Variety of collocation use

Table 4.35 (p. 149) displays repeatedly-used collocations from the two sets of essays. Findings from the study show that the total number of repeatedly-used collocations in the second set of essays reduced considerably (35 collocations less than the first set of essays). Looking at this dimension of individual students' essays, I found that 26 out of 29 students used fewer number of collocations repeatedly in the second essays compared to the first. In one case (T42), the number of repeatedly-used collocations in the second essay was double that of the first essay. However, it is noticeable that many of the repeatedly-used combinations from the first set of essays (31 out of 52) are combinations used in the question title such as *students have*, *studying abroad*, and *foreign country*. It seems that, except for *foreign country* (which can be replaced by *overseas/strange country*), there are no equivalent collocations to express *students have* or *studying abroad*. An OOCD search of the noun *students* does not give any suggestions for verbs following that noun. As for the verb *study,* the OOCD suggests some adverbs, but these are not equivalent to the adverb *abroad* (i.e. *carefully, closely, in depth, in detail, intently, intensively, extensively, widely/fully, thoroughly/systematically/thoughtfully*). This means that even with the support of the dictionary it is hard for learners to avoid repeating them if they want to express those ideas. Therefore, I presume that the essay question does play a certain role in the amount of repeatedly-used collocations.

As discussed earlier (section 4.2.2), it should not be taken for granted that the reduction in the number of repeatedly-used collocations was due to the positive impact of the OOCD. This can only be concluded if evidence of students using the OOCD to search for another way of expressing the same or nearly the same idea to avoid repetition could be detected. From the analysis in the previous chapter, I detected a few cases of the students using the OOCD to avoid repetition.

T11B: *cause stress, create stress*

T24B: *pressing problem, urgent problem*;

T27B: *cause stress, create stress, lead to stress*;

avoid stress, handle stress, minimise stress;

T31B: *serious problem, big problem, common problem.*

It is evident from this study that the dictionary can help students use a wide variety of collocations to avoid repetition as long as they wish to do so.

5.5 Research question 4

How do learners evaluate the use of the OOCD in support of their writing?

The answer to this research question was based on combining results from the survey, the recording sheets and the interviews. What learners like/dislike about the dictionary and their suggested improvement as narrated in the interviews brought deeper insights into how learners evaluate the use of the online collocation dictionary as a supportive tool and why they evaluate it as they did. The answer to this question relates to three aspects: assessment of satisfaction, perceived effectiveness, and accessibility and ease of use.

5.5.1 Assessment of satisfaction

Findings from the recording sheets show that students are satisfied with more than three quarters (76.4%) of the results found in the dictionary. The satisfaction assessment was also based on their responses to the questionnaire survey regarding different aspects pertaining to using the dictionary to search for collocations. In particular, as can be seen in Table 5.8 below, 91% of the participants felt confident when expressing ideas in writing, and 93.9% of the participants believed that the dictionary helped them expand their collocation knowledge. In response to the question asking if they would use the dictionary to support their writing in the future, almost all of the participants shared that they would use it. The high proportion of participants (94%) reporting that they would introduce the dictionary to their peers also proves that they highly appreciate the dictionary as an assisting tool.

Table 5.8 *Students' satisfaction towards OOCD use*

Items	Variables	N	Strongly agree %	Agree %	Disagree %	Strongly disagree %

5	Confident in expressing ideas	32	15.2%	75.8%	6.1 %	0%
6	Helpful for expanding collocation knowledge	33	63.6%	30.3%	6.1%	0%
7	Using more collocations	33	33.3%	57.6%	9.1%	0%
8	Help improve my writing	33	48.5%	42.4%	9.1%	0%
9	Preferring the OCD to other dictionaries	33	24.2%	57.2%	18.2%	0%
10	I will use the OOCD to assist my writing in the future	33	51.5%	45.5%	3%	0%
11	I would recommend using the OOCD to my friends	33	45.5%	48.5%	6.1%	0%

Learners' high level of satisfaction about the use of the dictionary as a supportive tool for collocation search was confirmed by the interview data. It involves features that the participants like about the dictionary. All the interview participants contended that the OOCD is a useful tool in supporting learners' collocation search. Their responses were full of positive expressions such as 'very helpful', 'convenient' (participant A, C, G, H), 'really good' 'very useful' (participant B, E), 'time-saving' (participant C), 'reliable' 'easy to use' (participant D, F), 'beneficial' (participant G). Most of them expressed their feeling of confidence when using collocations found from the OOCD. The following quotes are representative of their evaluation of the usefulness of the dictionary and why they feel confident when they use the dictionary as an assisting tool:

> *A: It's very helpful because it helps me to prevent translating Vietnamese to English word for word. It's also quite convenient for a student like me to find a natural expression of a word.*
>
> *B: It's really good, very useful, it saves me time in writing. Most of the time I find it demanding and challenging, and I can't come up with a word before or after a word that I need to use in my writing, I just look*

up and find out one. Finding and discovering one collocation make me more confident because I know that I'm in the right direction, the right way.

RS: *It is convenient. When I search the word stress, I can find many useful collocates that I can use later. I came across level of stress and used it later in my writing.*

Two of the participants expressed their trust in the dictionary since it is provided by a famous publisher, Oxford University Press, especially when comparing it to other sources for collocation search like "hello chao" or "google translate". One said:

D: *I think it is a reliable source for collocation check-up since it is compiled by a famous publisher. Compare to other sources that I had used to look up collocations such as 'hello chao' or google translate, I think the possibility of incorrect collocation use will be lower.*

Results from the survey questionnaires also showed that most participants (90.9%) thought that their writing would improve due to enhanced collocation use and that they would use more collocations if they could use the dictionary to assist with their writing. From the interview data, one participant shared that the dictionary offers a wide choice of native-like lexical collocations, which could help her avoid repetition in her writing. Another participant also shared that she is in the habit of learning new words and the dictionary is a good source of collocations to learn from. They said:

E: *I think the use of the OOCD is very beneficial and useful. First, it makes our writing more natural like the way native speakers write. Second, it helps me avoid repetition in my writing. For example, when I want to find a verb for success, I can use achieve, obtain, or have … a lot of collocates I can use. It is very useful when you are at intermediate or advanced level in IELTS. One of the criteria for getting high score is you have to use a wide range of vocabulary with natural control of lexicons. So I think the OOCD really helps me a lot in writing.*

G: *I often learn new vocabulary every day and when learning I usually learn them in combination with their collocates and sometimes learning how to make sentences. In this way it makes me easily remember the new vocabulary.*

It is, however, important to emphasize that that is what learners believe they gain from the dictionary. Results from the analysis of their writings show that the use of the dictionary in reality does not always bring as good results as they thought. In particular, for many of them (12 out of 29), their collocation use did not improve but

conversely got worse. In contrast to what they anticipated, the use of the dictionary did not help them use more collocations; there are 1,290 strong collocations in the first set of essays while in the second set only 1,184 strong collocations were used. The number of errors in the second set of essays increased instead of decreasing, with 17 more errors than in the first set. The number of repeatedly-used collocations in the second set of essays reduced substantially (35 collocations fewer than in the first set); however, as discussed in section 5.4.2, this was not really the positive impact of the OOCD. 31 out of 52 combinations repeatedly used were from the question titles (e.g. *students have*, *study abroad*, *foreign country*), and with the exception of *foreign country* (which could be replaced by *overseas country* or *strange country*), even when using the dictionary learners could not find equivalents to *students have* or *study abroad*.

The results from the survey data also show that preference for using this dictionary over others for collocation look-up was high among the students (81.4%). However, they also reported in the interviews some negative attitudes towards the dictionary. Reasons for not preferring this dictionary to others seem to be that, as two of the participants shared in the interviews, it would be sometimes a waste of time if they do not find collocations that they want. They then have to either look for them in another dictionary or paraphrase their ideas. Even if the OOCD provided them with collocates of a head word, it sometimes took them time to search for their meanings from other dictionaries before making a choice. One said:

> *F: It would waste me more time if I did not find the collocates that I wanted, or if I found some collocates but did not know their meanings.*

Another no less important factor contributing to the students' dissatisfaction is the dictionary not providing descriptions of pronunciation of headwords. Though this neither directly affects nor is necessary for their collocation use in writing, it literally is a drawback to dictionary users for their other receptive skill, speaking. It is true that, as can be seen in Figure 5.4 below, pronunciation description is not given. They compared it with other dictionaries:

> *A: It does not provide pronunciation description of all headwords. It could be ok if I use it for writing purpose but for speaking skill it is a disadvantage. Other dictionaries provide not just pronunciation but*

indicate differences in pronunciation between British and American English. Learners at this level, like me, can read the phonemic transcription, so human voice articulating words is not very necessary.

E: The dictionary should provide us with pronunciation. Maybe some words are easy and sometimes we neglect or ignore the correct pronunciation, but I think it's much better if it has pronunciation.

Figure 5.4: *OOCD search of stress, suffer, difficult*

5.5.2 Perceived effectiveness

Looking at the effectiveness of the dictionary use rating across all participants in this experiment gives the results shown in Table 5.9 below. As can be seen in this table, most of the participants (90.9%) contended that they can easily look up collocates of a word from this dictionary. Though lower, recording sheets recorded 81.9% of times participants found the collocations that they wanted to express their ideas in writing. Interview data also validated this. Two participants shared that most of the time they found collocations that they wanted to use. One said:

> **B**: *When I want to find a verb for **success**, I can use **achieve**, **obtain**, or **have** ... a lot of collocates I can use. It is very useful when you are at intermediate or advanced level in IELTS.*

Another participant shared that the dictionary is effective in that in one search she could find different collocates that can go with the headword, so she could use them later in her writing to avoid repetition. For example, when searching for the word *stress*, she came across the collocation *level of stress*, which, as she reported, was made use of soon after that.

Table 5.9 *Effectiveness of dictionary use*

Items	Variables	n	Strongly agree	Agree	Disagree	Strongly disagree
1	Easily look up collocates of a word	33	27.3%	66.7%	6.1 %	0%
			Always	**Often**	**Occasionally**	**Never**
12	Get instructions on collocation use	33	21.2%	51.5%	24.2%	3%
13	Get information needed	33	0%	69.7%	27.3%	3%
14	Use the dictionary in combination with other dictionaries	33	9.1%	57.6%	30.3%	3%

However, more than a quarter of the participants (27.2%) responded that they only occasionally or even could not find instructions on how to use collocations through examples. This coincides with the result of the observation data (see Table 5.10 below), which showed that in nearly a quarter of the searches (22.2%) participants did not find instructions on collocation use. The percentage of responses to this question (86.1%) is not really high though. This is because, as I explored in the interviews, although they could not find examples illustrating how the collocations they found were to be used, they knew how to use them grammatically correctly by looking at other examples.

Table 5.10 *Summary of findings from recording sheets*

Items	Questions	Yes	No	Missing
3	Did you find the word you were looking for?	118 81.9%	24 16.7%	2 1.4%
5	Did you find instruction on how to use it?	92 63.9%	32 22.2%	20 13.9%
6	Did you use the OCD in combination with other dictionaries?	12 9.5%	114 90.5%	0

The survey data showed that 30.3% of the participants only occasionally or even never found the information they wanted. 66.7% of the participants responded that they had often used the collocation dictionary in combination with other dictionaries, mainly to look for meanings of collocates. This could relate to non-plentiful content, a dislikeable feature about the dictionary that participants shared in the interviews. Three of the participants said that they were not very satisfied with content of the dictionary because it is not rich. It only has a limited amount of words; it does not contain academic words like *abnormal*, *anomalous, acquire,* or *face* (a verb) and many more words that they learn in SAT and GRE. Also, for each word that it presents, it provides fewer meanings than other general dictionaries. An example that one participant used to illustrate the point is the word *policy*. The OOCD gives two meanings, while in the Oxford Advanced Learner's Dictionary *policy* has three meanings:

OOCD: 1. *Plan of action*

2. *Insurance contract*

OALD: 1. *A plan of action agreed or chosen by a political party, a business*;

2. *A principle that you believe in that influences how you behave*;

3. *A written statement of a contract of insurance*

Three of the participants expressed their concern, stating that the dictionary does not provide many collocates that can accompany the headword being considered. They did not give specific examples to illustrate their points, though. Failure to present combinations that they believe to be used by native speakers confuses them. They also suggested that it can hinder their creativity in combining words:

> *A: I believe that English has many words and in the OOCD it just lists down some of the most common combinations. I sometimes feel … like I know that native speakers they use some combinations that are not in the dictionary. I feel that the OOCD hinder my creativity. I don't know if I can combine some new adjectives that I've learned with a noun or not, so I sometimes get confused and not comfortable when I use a collocation from the dictionary.*

> *B: On the online one [dictionary], it [a sought word] has only one meaning while in fact it has many meanings. The dictionary has limited number of words. It is not various as the Oxford Advanced Learners' Dictionary. Sometimes I type a word and cannot find it and then I have to use another source or google … I'm wondering whether I can use my own way or style of combining words.*

H: I'm really concerned about whether I can use my own combinations of words according to their meanings … It doesn't have some very common words like acquire or get. When I searched the words require, it presents only two adverbs that can accompany it while as I know it can go with other adverbs like legally, normally or merely.

Their concern regarding content of the dictionary is undeniably true. There is a substantial difference in the number of word combinations and examples between the online dictionary and its installed electronic version. The online dictionary provides around 150,000 combinations and 50,000 examples while the electronic collocation dictionary presents over 250,000 combinations and over 75,000 examples (McIntosh, 2009). It seems important for the dictionary users to know that compared to the installed electronic version the online dictionary provides fewer word combinations and examples. Learners can rely on it as a facilitator to look for suggestions for collocations, but should not restrict themselves to the collocations provided.

Lack of examples to illustrate how a collocation should be used is also one of the factors that makes the dictionary less effective. As they reported, having no illustrative examples sometimes contributed to their perplexity. One participant shared on the recording sheet that:

RS: There is no example in some cases, so I don't know if I use the combination found in the dictionary appropriately in meaning in a particular context.

However, the reality shows that only 12 out of 126 (10.3%) of their look-ups were in combination with other dictionaries. The big difference between the two sets of data with regard to this feature is understandable since, as they shared in the interviews, they often opt for collocates that they know when writing under time constraints rather than look for their meanings in another dictionary. Two of the participants shared in the interviews their expectation of the dictionary to have the meanings of each collocate group presented. They hinted that in this way the dictionary would be more effective when saying *It would waste me more time … if I found some collocates but did not know their meanings.*

5.5.3 Assessment on accessibility and ease of use

Table 5.11 below presents the survey results pertaining to learners' evaluation of accessibility and ease of use. Findings show that language learners can access the dictionary easily when the internet is available (90.9%). This is also what six of the interview participants indicated they liked about the dictionary; they stated that it can be accessed from any technological device and can be opened very quickly, while for other electronic dictionaries it often takes a little while. Yet learners can only access it if the internet is available. To some others, this turns out to be a drawback since internet is not available everywhere. Even if it is available it does not always guarantee fast and efficient access. Internet speed decides how fast a look-up is. In reality participants in this research at times experienced slow and interrupted searches, which might have affected their flow of thought. Commenting on this, one participant expressed her dissatisfaction:

> *E: Internet is available in the University campus but it does not mean that you can get access to the online dictionary from anywhere in the campus. You might be dropped out during your search if internet is not strong enough.*

The online collocation dictionary could be a solution for the problem of time involved in flicking through the dictionary pages and subsequent disruption of the flow of writing which concerned students in Dziemianko's study (2010). Although this present study did not aim to compare between online and paper dictionaries and consultation time was not measured, based on their own experience the participants seemed to make some comparisons, stating that fast search time is an advantage of the dictionary. All participants in the survey agreed with the statement that the dictionary saved them time for each check-up. This was confirmed in the interviews. Two of the participants shared:

> *A: I have a paper book of collocation dictionary, and I think that the OOCD is a faster way to look up a word.*
>
> *B: You know, in comparison with Oxford Advanced Learner's Dictionary, I think the OOCD is much better because it is faster for me to turn it on, and when I type a word in, the result immediately appears.*

One participant reasoned that that could be because other installed dictionaries contain more words and information, so they are much heavier. For each search, she had to wait for a few seconds:

E: *It (the OOCD) is running faster than its installed version, which could be because the installed version contains more collocations, more examples. Usually, I have to wait for a while for the result to show up.*

This could be their general assessment based on their intuition of the time for a word being looked up to appear on the screen compared to the time to find it in a paper dictionary. If we consider the check-up time as the whole process starting from a word being typed in until a collocation is found, learners' responses to this feature could have been different. One participant shared in the interview:

F: *It would waste me more time if I did not find the collocates that I wanted, or if I found some collocates but did not know their meanings.*

Table 5.11 *Accessibility and ease of use*

Items	Variables	N	Strongly agree	Agree	Disagree	Strongly disagree
2	With internet availability I can access this dictionary easily	33	51.5%	39.4%	9.1 %	0%
3	Save time for each check-up	33	66.7%	33.3%	0%	0%
4	The layout of meanings, grammatical use and frequently used expressions is user-friendly	33	21.2%	63.6%	15.2%	0%
			Always	**Often**	**Occasionally**	**Never**
16	I have difficulty in making a choice of collocates found	33	12.1%	21.2%	54.5%	12.1%
17	It takes me time to look for collocates from other dictionaries	33	15.2%	27.3%	48.5%	9.1%

84.8% of survey participants responded that the layout, grammatical use and frequently used expressions are user-friendly. Responding to this question, three interview participants shared that the dictionary is clear and well-organized. This could be an element contributing to the high level of satisfaction with the dictionary, as discussed earlier. The use of different colours, uppercase/lowercase letters, or words in bold or italic help them scan for the position of the part of speech of words that they want to look for fast and easily (see Figure 5.5 below). Also, the

arrangement of words with similar meanings together assists them in using collocations regardless of the limited number of examples. They shared:

> *E: I like the way it organizes the words in the dictionary … like if I search for verb collocates of a noun … first of all I have Adjective + noun, and then verb + noun. It is presented in a fixed order. Collocation patterns are in red, so I can easily know the order of collocates (part of speech).*

> *G: It puts collocates of similar meaning together. Even if you don't know meanings of some collocates, you can find that in that group they have similar meaning and figure out meanings of words in that group. It also has examples, so it helps me to figure out how to use that word.*

Figure 5.5: *The OOCD entry for the word **stress***

However, two of the participants shared that its interface is also an aspect that needs to be improved. Compared with other webpages that also assist English learners, one participant assessed that it is quite plain and tedious, not professional. She supposed that if there were more pictures to illustrate, it would be more attractive, and hence could help learners learn more easily. In contrast, another participant argued that its target audience is upper intermediate to advanced learners, and they are supposed to know the meanings of those collocates, so it need not add pictures for definition

illustration. She explained that *most of the time I have no difficulty in understanding collocates.*

Two of the participants have also shared that advertising appearing in the centre of all of their search pages annoyed them. They even expressed concern that that could distract learners from their writing. They said:

> *B: You know ... because it is a free online dictionary, it is often inserted with advertisings. Those are moving pictures and are very outstanding, so they sometimes irritate me. For me it is ok. I'm not distracted by those advertisings but it could distract others.*

> *E: I don't like the advertising appearing on the page. I think it's normal. You are using the dictionary for free, so you have to accept that. When writing in class, it doesn't distract me because I have to concentrate on my writing, but when I write at home, I sometimes get distracted by advertisings about clothes or programs to help check grammar in writing.*

87.9% of the participants reported in the questionnaires that they have difficulty in making a choice of collocates for a headword, so sometimes it takes time to look for collocates from other dictionaries (according to 90.9% of participants). This has also been confirmed from the interviews. As one of the participants stated, she has almost no difficulty in using grammatically correct collocations found in the OOCD in her writing. What matters is that she does not know the difference in meaning between collocates instead

> *C: when I search an adjective collocate of a noun, it gives me a list of adjectives that can go with that noun, but then I can hardly know which to use in a particular context. I sometimes don't know the difference in meanings among those collocates to choose appropriately. (translated from Vietnamese)*

A participant with a similar view commented on the recording sheet that:

> *RS: Although it's quite easy to know how to use the combination found grammatically correctly, the dictionary does not offer learners links to look for meanings of collocates. I think that's the difference between online dictionary and electronic dictionary. As I know for many electronic dictionaries, you can cross search a word and it clearly saves learners time.*

However, another participant perceived things differently. She explained:

G: it puts collocates with similar meanings together. Even if you don't know meaning of some collocate, you can still figure it out from collocates of that same group. Also, it has examples, and they help me to know how to use that word in a context.

Another difficulty that participants sometimes experienced in using the dictionary is that when typing in a word in the search box, it does not provide a list of words suggested based on the first initial letters of the word being searched. One participant expressed:

B: The dictionary does not give any suggestions of words based on the initial letters of the word that I'm searching. This means that I have to remember exactly the word that I want to look for. If there is any mistake in the word that I type in like a letter missing or letters not in correct order, results will not be found. Compared to other electronic dictionaries, this is a hindrance. If the dictionary had that feature, it could help users fasten the searching speed by not having to type in the whole word.

Another issue, though not really a difficulty, that one participant wondered about is the presentation of the N-N collocation in the dictionary. If, as confirmed by the dictionary compilers, the presentation of collocations in the dictionary is at the base entry, which learners will think first, then the presentation of N-N collocation seems not to comply with the rule. Vietnamese learners, in order to express an idea like *chính sách giáo dục* (*education policy*), tend to start thinking of the second noun *policy* (*chính sách*), which is not the base. This means that in order to search for a N-N collocation, learners have to remember which of the two nouns is the base. One participant said:

G: Although I have been trained on how to look for collocations from the dictionary and I know that to search for a collocation of N-N structure like work experience I need to type in the first noun work, I sometimes forgot. Actually, if I search for experience, I can still find the combination but work now is an adjective. It's a bit confusing. I think it's different from others (collocation patterns). Like for the case of a bunch of flower, it's clear that you start searching with flower, but for this case we don't start with experience, which I think is the main noun.

5.6   Summary of the chapter

This chapter set out to answer the four research questions. From the analysis of the data sets, I found that collocations are also problematic to Vietnamese advanced learners. Learners produced odd collocations due to the wrong choice of collocates more often than any other error types. They tend to approach the dictionary for help as soon as they need it rather than waiting until they finish their writing, and this was found to have brought them confidence in expressing ideas and a feeling of completeness. Writing with the dictionary support did not in fact help learners improve their collocation use. The dictionary, however, is highly evaluated by learners. Some suggestions on how to improve the dictionary were made to contribute to increasing the dictionary's user-friendliness.

Chapter 6 : DISCUSSION OF THE FINDINGS

6.1 Introduction

In the previous chapter I addressed the research questions on the basis of the results found. In this chapter I will discuss five key points that emerged from the findings, including learners' collocation use, benefits of the dictionary, suggested strategies for dictionary use, suggested improvements for the dictionary, and effects of the OOCD on learners' collocation use.

6.2 Learners' collocation use

Results from the first set of texts (without the OOCD support) in this study show that 117 (8.4%) out of 1,407 collocations of all the seven patterns were identified as odd collocations. This result appears to run counter to other previous research which confirmed that collocational oddities account for a large number of collocations used - 36% or up to 56% (Nesselhauf, 2005; Granger, 1998; Dang, 2014). This could be partly because collocation in these studies was defined based solely on the phraseological approach, and therefore collocations in these studies only include combinations that are very restricted in their combination. The possibility of learners making errors could be higher.

Another possible explanation for the big difference in the amount of errors could lie in the tasks that learners were required to do. In particular, Bahns and Eldaw (1993) asked learners to do cloze and translation tasks in order to test learners' productive knowledge of specific groups of collocations. In this study, however, to explore learners' use of collocations (namely what kinds of collocations they can use correctly and what kinds they have trouble with in their free language production), I asked them to write an essay on a given topic. Learners in this study could possibly avoid using collocations that they were not confident in, and therefore the number of errors could be fewer.

With regard to types of odd collocations, the findings are consistent with many of those in previous studies. In particular, the study found that V-N collocations are the most problematic for Vietnamese advanced learners (51% of all errors), in line with the results of other studies on Vietnamese, Polish and German-speaking learners (Dang, 2014; Biskup, 1992; Howarth, 1996). The order of collocational patterns with the most to the fewest errors is almost exactly the same as that in Dang (2013); the collocational pattern with the most errors spotted after the V-N pattern is adj-N; N-of-N and Adv-Adj patterns have the least errors. The findings are also consistent with those of Nesselhauf (2005) and Bahns (1991), who state that errors can lie in any element of collocations. The wrong choice of collocating words accounts for the majority of errors, at 73 out of 117 (62.2%). This same result has been found in Nesselhauf's (2003) study on collocation in a learner corpus, Nesselhauf being among the few researchers who have investigated all elements of V-N collocations. 181 out of 352 V-N collocation errors were identified as such in her study (Nesselhauf, 2003, p. 171). This study also confirms a claim from a piece of research not on learners' collocation production in free writing but on their use of collocations in translation by Bahns and Eldaw (1993), that as for V-N collocations verbs pose more problems than any other lexical elements.

There are several possible explanations for why V-N has the highest number of errors and collocates of such a pattern are the most troublesome. The first and quite obvious probability is that V-N is the most frequent combination (518 out of 1,407 combinations of all seven patterns). It occurs in every sentence (Howarth, 1996; Aisenstadt, 1981), and thus problems with it are likely to occur more frequently.

Another possible explanation for why V-N has the most errors is that unlike Adv-Adj, or Adj-N collocation pattern in which one element is optional, it is often impossible for the learner to either not use or to choose a safer collocating word (e.g. *very*) (Bahns and Eldaw, 1993). Take, for example, the collocation *get a job* in the sentence *The cost of studying abroad is really high, so students often have to get a part-time job to support their living* (T4A). The collocation was appropriately used in

the sentence, and there seem to be no safer collocating verbs to replace *get*. *Have* has a similar basic meaning to *get* and a higher frequency than *get*. The replacement of the verb will bring about a slight change in meaning though, and therefore it is inappropriate in the given context. Obviously there is no 'all round' solution for V-N collocation as with the case of Adj-N or Adv-Adj.

It could also be because paraphrasing a V-N collocation is often difficult if not impossible (Bahns and Eldaw, 1993). Failure to use the collocation '*create stress*' to express the idea can result in the sentence not making sense or sounding awful. It was paraphrased as this in a piece of student writing: *This makes stress appear so common that it has become a major problem in many countries*, which would have been more appropriately phrased as *This creates/causes so much stress that it has become a major problem in many countries*.

Also, there seems to be a confusion of semantic near-synonyms of appropriate verbs. *Perceive* (+ knowledge), *raise* (+ students' independence), *shorten* (+ the gap), and *face* (+ differences) were used, while their near synonyms *acquire*, *increase*, *reduce,* and *confront* respectively were required instead. In other words, words are potentially confusing to language learners if they have a similar translation in learners' L1. *See, look, watch* and *view* all mean *xem* in Vietnamese, and it might be problematic for Vietnamese learners to express the idea *xem tivi* (*watch television*) unless they encounter it frequently enough. This explanation seems applicable to errors of other collocational patterns as well. It would not always be an easy choice among words of similar translation such as *small, little, tiny, toy, mini, minor*, which all mean *bé* in Vietnamese. To express the idea *cô bé* (*little girl*) learners might say *small girl*, which means a girl who is small rather than referring to a girl who is young in age. Combining *girl* with *minor, toy, tiny, or mini* might result in combinations that sound awkward or inappropriate in meaning in a certain context. If that is true, Cop's (1990) suggestion that collocations should be learned according to semantic sets of a base as in the below figure is entirely reasonable.

Figure 6.1: *Semantic sets of the base* ***decision***

VERB + NOUN

Arrive at/ come to/ make/ reach/ take

Overrule/ overturn/ quash/ reverse

Announce/ give

→ **DECISION**
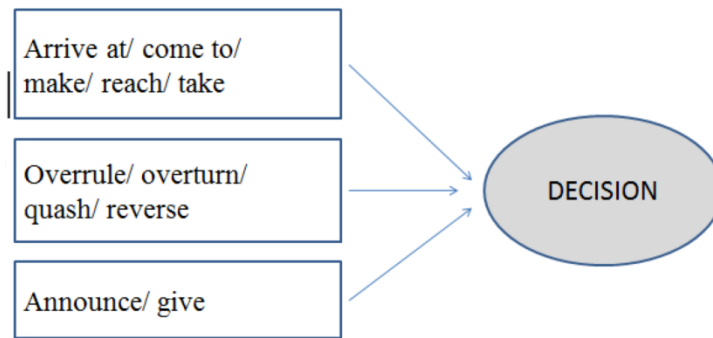
One other possible explanation for this is learners' ignorance of semantic prosody (discussed in section 2.2.1). An example of errors found from the study that are associated with this is *provide people with disadvantages. Provide* has a positive semantic prosody (Stubbs, 1995); that is, the word is observed to co-occur frequently with a wide range of evaluatively positive things. Its combination with *disadvantages*, which has a negative meaning, is therefore inappropriate. As such, a contrastive analysis of collocations and an introduction to semantic prosody would possibly be useful to L2 learners.

More errors at the collocate than at the base could also be because when learners want to express something they think first of the base and then look for the collocate to complete their phrasal meaning (Coffey, 2010). In other words, the choice of the collocates depend on or are restricted by the base (Laufer 2010), which is a 'lexical-semantic constant' (Cop, 1990, p. 41). Take, for example, the base *stress*. To express the idea *to make stress happen*, one can say *cause stress*, *create stress* or *lead to stress* but not *bring stress*. This cause seems to be closely associated with semantic near-synonyms, as I have just discussed.

The study found that N-N collocations account for 10.2% of all errors, but the ratio of errors over the number of collocations for that pattern is highest (15.2%). Learners tend to combine words based on direct translation from their L1. The way of forming N-N collocations by combining two nouns is not irrational since there are a lot of N-N collocations equivalent to their L1 (e.g. *mountain bike: xe đạp leo núi, bus stop: trạm xe búyt, education policy: chính sách giáo dục*). Generalizing the rule of

forming N-N collocations by combining nouns together regardless of the conventional word combination of the language unfortunately resulted in errors. This also suggests that learners appear to rely more on the open choice rather than the idiom principle. Learners' generalization of this rule somehow reflects the difference in nature of this collocational pattern compared to the other six patterns. That is, the choice of collocates to construct collocations of the other patterns depends on the choice of the bases, while it seems that, as for collocations of N-N pattern, the choice of the collocate does not depend on the choice of the base. In other words, the base-collocate relation of N-N collocations does not function in the same way as that of the other six patterns. Take, for example, the choice of the collocate meaning *to do* that can accompany nouns such as *a cake* and *homework*. Collocates *make, bake* will accompany *a cake* and *do* will accompany *homework* (suggested by the OOCD), while it seems that the choice of the collocate *program* or *policy* in *education program* or *education policy* does not depend on the choice of the base *education*. The construction of collocation of this pattern is rather the combination of two elements of equal status. If this is true, Wray's (2002) suggestion that collocations be learned as big words is apparently reasonable.

Though L1 influence on collocation use was not explored in this study and so was not investigated emphatically, L1 seems to have some influence on learners' L2 collocations. This was claimed in many previous studies with learners of different language backgrounds, Polish, Germany, and Thai (Granger, 1998; Biskup, 1992; Nesselhauf, 2005; Bahns and Eldaw, 1993; Phoocharoensil, 2012). In this study there are cases where learners seem to make errors with congruent collocations (e.g. *face* instead of *confront differences, shorten* instead of *reduce the gap, raise* instead of *increase students' independence*), which are supposed to be less likely to be troublesome to them. The possible cause for these errors is the puzzlement with synonyms that I have just mentioned above. *Face, deal with, cope with, encounter,* and *confront* all mean *đối mặt với* in Vietnamese. Therefore, it is sometimes challenging for learners to use combinations which do not sound awkward in certain contexts.

It seems that though troublesome (Wolter and Gyllstad, 2011), some collocations with no direct translation were still used correctly in the learners' writing (e.g. *make mistakes, meet the needs, take a chance* etc.). Part of the reason for this might be that they are collocations with fairly high frequency of occurrence (the Log Dice score of *make mistakes* is *8.2; meet the needs* 9.8; *take a chance* 6.8) and so learners come across them quite often in their language learning process. They, therefore, become the language intake. Where learners do not know L2 collocations, they will produce collocations based on their L1 collocations. These are just initial possible indications of learners' L1 influence. The influence can be both positive and negative, so this is an important issue for further research.

Some of the odd collocations that the study found seem to be associated with learners' linguistic and cultural background. *Meet robberies ≈ gặp kẻ cướp*, *live normally ≈ sống bình thường*, *government and individual…joint hands ≈ nhà nước và người dân cùng chung tay* reflect the way Vietnamese expresses ideas. The appropriate ways to express these ideas, e.g. *be victim of a crime, live peacefully, or the government and individual work together* respectively, are not necessarily unknown to the learners but did not come to their minds. This would be an interesting point to investigate further.

The findings show that learners did not have much problem with collocations containing delexicalized verbs (e.g. *have a look, make a decision*) except for two cases in which they were not used appropriately in meaning. There are two possible explanations for this. Delexicalized verbs *take/have* and *make/do* occur very frequently, so could have been introduced to learners at earlier stages. The other possible reason could be L1 influence. Some combinations in Vietnamese seem to exhibit the delexicalization phenomenon as well. To express ideas such as *make a decision, make a suggest, make a claim, or make a comment*, people say *đưa ra quyết định, đưa ra đề nghị, đưa ra nhận xét,* and *đưa ra khẳng định* respectively. Similarly to English, the nouns in these combinations (*quyết định, đề nghị, nhận xét, khẳng định*) are the elements that carry meaning. The combinations as a whole mean exactly the same as the corresponding verbs: for instance, *đưa ra quyết định ≈ quyết định.*

The findings reinforce a claim in some previous studies that besides the wrong choice of collocate and base, odd collocations are also associated with incorrect use of prepositions (Phoocharoensil, 2012; Nesselhauf, 2005). The addition, omission, or misuse of prepositions appears to be linked to their L1. Learners tend to add a preposition as in *approach **to** a new environment* when there appears an 'element functioning as prepositions between verbs and nouns to express the scope of or direction to the noun object' of the combination in Vietnamese (Châu et al., 2011). *To* is mistakenly added since it is equivalent to *với* in *tiếp cận **với** môi trường mới.* Instead of writing *suffer from homesickness*, students wrote **suffer homesickness* which is equivalent to *chịu đựng nỗi nhớ nhà* (with *chịu đựng* equivalent to *suffer, nỗi nhớ nhà* to *homesickness*). Similarly, instead of *adapt to the new environment*, **adapt **with** the new environment* was used. In this case the preposition *with* is the direct translation of *với* in Vietnamese. Regarding these errors, further research should be done to investigate the potential influence of Vietnamese on learners' L2 collocations pertaining to prepositions.

Of the three most common reasons for collocation errors that I have mentioned in the Literature Review chapter (lack of awareness, cross-linguistic influence, lack of knowledge of collocation properties), lack of awareness of collocation phenomena does not seem to be an issue because in the present study learners were all introduced to it and to the collocation dictionary which they could use to seek help with collocations. No claim could be made with regard to cross-linguistic influence because, as discussed earlier, it was not the aim of this study to investigate that; there appears, however, to be some indication of errors pertaining to this. The other cause, lack of knowledge of collocation properties, is also present in many errors in this study. To deal with this, learners tend to either translate directly from their L1, which results in errors in some cases as illustrated above, or try to paraphrase the intended meaning without using a collocation. For instance, instead of using the collocation *study requirements*, they paraphrase it into *the need of studying* as in *Students choose to study in foreign countries mostly for seeking a modern and appropriate environment which can satisfy their need of studying* (T27A). This 'creative strategy' was also found in the study of Biskup (1992) on German learners and was also

claimed to result in collocational errors. Some other odd collocations of this kind found in the study are:

(T25A) *State of emotion* → *emotional state*

(T25A) *The stress of relationship* → *the relationship stress*

These odd collocations seem to indicate that N-of-N pattern is perceived by the learners as a construction that can be used freely in English.

6.3   Benefits of the OOCD

Findings from the study show that learners were generally positive towards the use of the OOCD for collocation searching. They regarded the dictionary as useful and convenient. For the purpose of collocation check-ups, the encouraging assessment is understandable since the focus of this specialized dictionary is on collocations, whereas, as I have mentioned in the Literature Review chapter, there is a basic lack of collocations in other general British monolingual dictionaries (Hottsrnonn, 1991, p. 230), or they are hidden in examples (Laufer 2010). Interview data shows that it was assessed as convenient that in one search learners could find different collocates of a headword, which they could use soon after. That could be seen as an advantage of the OOCD over other general dictionaries. When writing about a particular topic, *causes of stress* for example, learners would probably need collocations pertaining to *stress*, e.g. *cause, create/reduce, relieve/avoid, remove/cope with, manage, handle stress*. In one search, learners could find some useful collocations to employ.

The dictionary was also assessed as convenient partly because it is a free online dictionary. With the widespread availability of the internet, learners can access the dictionary from any technological device at anytime and anywhere. This is a great benefit of the dictionary. This finding is in agreement with Chon (2009), who showed that the availability of the online dictionary also helps learners to get rid of the burden of carrying with them bulky paper dictionaries or installing an electronic dictionary on their technological devices ready for use. However, convenience does not necessarily lead to progress in learning. In this study no progress has been found as result of learners being allowed to access the OOCD to support their writing.

Findings from the study also show that speed of search was one of the advantages of the OOCD. With regard to this aspect, learners' assessment corresponds to findings by Dziemianko (2010), Nesi (2000), and Roby (1999), which confirm that the process of electronic dictionary consultation is less time-consuming. This could encourage learners to do more exploratory browsing (Nesi, 2000) and so learners can learn more from the language input (Laufer, 2010). However, concerns that information that can be retrieved so quickly and painlessly from electronic or online dictionaries will be forgotten easily (Nesi, 1999) are not irrational. In order to find the answer to this, further research should be done to compare the retention rate of collocations searched in an online collocation dictionary and a paper collocation dictionary.

Results show that there are some, though not many, indications that the dictionary could help learners to avoid repeatedly using some collocations. The presentation of collocates in semantic sets split up by a dash (see Figure 6.2) is reasonable and is supposed to be known by the learners since they were introduced to the structure of the dictionary at the beginning of the course. Therefore, it is highly dependent on how learners intend to use the dictionary.

With regard to the layout of meanings, grammatical use and frequently used expressions, the dictionary was assessed as well-organized and user friendly. The OOCD makes use of the strengths that most of the electronic dictionaries do: using different colours, uppercase/lowercase letters, or words in bold or italic to help the user quickly position words that they want to look up (Chon, 2009; Nesi, 2000). In-depth interview data, however, shows quite different perceptions of this aspect. Some participants believe that the interface of the online dictionary should be improved to make it more eye-catching, for example by adding pictures; in this way it is hoped that it helps learners learn and remember new words easily. However, I believe that there is no need to expand the dictionary in that way because the audience the dictionary is aimed at are upper intermediate to advanced level, and more importantly its primary purpose is to provide collocates for productive use rather than providing meanings of headwords. The former reason, the target audience, was in reality backed up by a participant's response in the interview.

6.4   Suggested strategies for dictionary search

Findings from the interviews show that different strategies of dictionary use have been applied. Some decisive factors that have impacted on how learners used the dictionary are time constraints, flow of thought, types of writing (academic or free writing), and topic of writing. Based on these findings, I discuss and suggest some strategies for dictionary searches below.

When writing under time constraints, learners tended to use the dictionary immediately, as soon as they needed it. This brought them confidence that their use of collocation in the writing was correct and complete. This learner behaviour is understandable and appropriate since, as the findings show, most of the time they could quickly find collocations that they wanted. As they stated, they often opted for collocates of which they knew the meaning, since time did not allow them to do further searches. This decision seems quite reasonable since besides the matter of time, further searches for meanings of collocates might lead them astray or distract them from the writing. The use of the OOCD at the end to polish their writing after they had finished is deemed a wise move. They used the OOCD to look for collocations of which one element is optional, such as Adj-N, Adv-V, or Adv-Adj. However, in reality not many of them made use of the dictionary in this way.

Conversely, when writing without time constraints they were inclined to choose new collocates, which required them to check for meanings in other dictionaries. This, according to them, is to enhance collocation use in their writing. If it is true that the improvement of vocabulary depends on how a word is searched for rather than the number of times it is searched for (Bruton, 2007; Laufer and Hill, 2000), this strategy is then reasonable because it could not only help them avoid using collocations inappropriately, but also enhance the possibility of turning language input into language intake after several careful searches of a collocation. However, in order for this to be confirmed further research needs to be carried out.

6.5   Suggested improvements

Findings from the study show that content of the dictionary is one of the issues that needs to be improved. In particular, it was suggested that the dictionary should provide more examples to help learners with usage. Regarding this, I argue that the dictionary has given a fair number of examples as illustration for usage. There are almost always examples for each semantic set; in cases that there are not, the collocations are quite straightforward to use. Learners are supposed to be able to construct them based on their syntactic knowledge without difficulty. Take, for example, the headword *challenge* in Figure 6.2 below. Adjectives that can collocate the noun with its first meaning *sth new and difficult* are many, but there is only one example provided. Similarly, in the case of *challenge* as a verb, one group of adverb collocates (e.g. *successfully*, *unsuccessfully*) is given without examples to illustrate. However, for learners at upper intermediate level upwards the use of these collocations is deemed quite simple. They only need to combine them based on their syntactic knowledge. This argument is backed up by a response from the interviews, stating that there is no difficulty in using collocations provided by the dictionary in terms of syntactic structure. Evidence from the study also shows that the learners used all collocations looked up in the dictionary syntactically correctly. However, the suggestion should not be ignored since an electronic dictionary does not have space restriction as a paper dictionary does. With good planning it can provide learners with more examples, or at least one for each collocate group, without detracting from clarity and accessibility.

Figure 6.2: *OOCD search of the word **challenge***

challenge *noun*
1 sth new and difficult
ADJ. big, considerable, enormous, great, huge, radical, real, serious, significant, strong | difficult, tough | major, main | fresh, new | exciting, interesting | economic, environmental, intellectual, political, technical, technological *Liszt's piano music presents an enormous technical challenge.*

VERB + CHALLENGE *be*, *pose*, *present* | *face*, *meet*, *respond to*, *rise to*, *take on/up* *The gallery has risen to the challenge of exhibiting the works of young artists. He has taken on some exciting new challenges with this job.*
CHALLENGE + VERB **face sb** *the challenges facing nurses in casualty*
2 that shows that sb refuses to accept sth
ADJ. **serious** | **direct** | **effective** | **legal**
VERB + CHALLENGE **be**, **present**, **represent** *The demonstration represents a direct challenge to the new law.*
PREP. ~ **to** *a legal challenge to the president's power*
3 invitation to compete
ADJ. **leadership**, **title** *his title challenge to the heavyweight champion*
VERB + CHALLENGE **issue**, **mount** *plans to mount a leadership challenge within the party* | **accept**, **take up** *I accepted his challenge to a game of chess.* | **beat off**, **fight off** *Our team will have to fight off the challenge from better trained teams.*
PREP. ~ **from** *a challenge from the other political party*

challenge *verb*
1 question whether sth is right/true
ADV. **directly** *The newspaper was directly challenging the government's legitimacy.* | **seriously** | **effectively** *She was effectively challenging the whole basis on which society was run.* | **successfully** *The story was completely untrue and was successfully challenged in court.*
PREP. **on** *She challenged him on his old-fashioned views.*
2 invite sb to compete, argue, etc.
ADV. **seriously** *No one has seriously challenged the champion.* | **successfully**, **unsuccessfully**
PREP. **for** *She was poised to challenge for the party leadership.* | **to** *The count challenged him to a duel.*

Moving on to learners' suggestion that the dictionary provides meanings for each collocate group, it is not irrational to argue that meanings of collocates are not provided so as to help learners focus on reference work and because they are inferable from semantic sets or demonstrative instances (Coffey, 2010). Evidence from the study, however, shows that learners did not always use collocations found from the dictionary contextually inappropriately (see section 5.4.1), especially when they did their writing under time constraints. The search for the meaning of collocates from other dictionaries under such circumstances proved less feasible. This suggestion is, therefore, worth considering. Providing the meaning of each collocate group is probably supporting learners to use collocations semantically appropriately. This is also expected to help reduce search time for their meanings in other dictionaries. In this way, the dictionary could be helpful to learners at lower levels, whose vocabulary store is more limited.

A cross-searching function is an interesting suggestion from the participants to solve the problem of the dictionary not providing the meaning of each collocate group. In

fact, this is not a new idea since most electronic dictionaries make good use of this feature. Electronic dictionaries with this search function would encourage dictionary users to do more exploratory browsing (Nesi, 2000) with less time consumed than with a paper dictionary (Dziemianko, 2010).

The findings also show that in order to use the dictionary learners have to remember the exact spelling of the word that they want to look for, which sometimes poses difficulty. That is because the dictionary does not give any suggestions for words that users want to look for based on initial letters being typed into the search box. Incorrect spelling of a word will result in the word not being found. Clearly, this is something of a hindrance to the dictionary user. Compared to other online dictionaries available such as Cambridge Dictionary, Macmillan Dictionary, and Oxford Learner's Dictionary, this is obviously a drawback. It would take learners time to look for the spelling of a word in another dictionary in case they do not remember the exact spelling. So when building a useful dictionary for the production of collocations, dictionary compilers need to take this issue into consideration.

With regard to layout of collocates, syntactic structure, and frequently used expressions, learners in the study evaluated these quite positively. However, there are still some suggestions for the presentation of collocates. For the sake of clarity, collocates of the same semantic sets should be presented separately by putting them on a new line or using a bullet point to signal them. This suggestion is reasonable and is deemed easy to carry out because like other electronic dictionaries, this online dictionary is not space-bound. In reality, collocate groups are presented separately on different lines in its electronic version.

In this study learners expressed their concern about the dictionary not providing description of pronunciation, a feature that all electronic or online dictionaries make use of. It is true that for writing purposes learners did not need them, but for the other productive skill, speaking, this is a fundamental weakness. When building a useful productive collocation dictionary for learners, the dictionary compilers need to take this into consideration.

From the findings about learners' odd collocations, it seems that Nakamoto's (1992) argument for typicality, not predictability, to be the central factor in determining what collocations should be included in a collocation dictionary, is rational. In this study some collocations that are predictable to Vietnamese learners since they are direct translations from learners' L1, such as *confront difficulties* (with *confront* meaning *đối mặt với* and *difficulties* meaning *những khó khăn*) and *increase knowledge* (with *increase* meaning *nâng cao* and *knowledge* meaning *kiến thức*), are still problematic to Vietnamese learners.

6.6   Effects of the OOCD on learners' collocation use

The empirical results show that writing with the supportive tool did not help learners improve their collocation use. The number of odd collocations in the essays written with the dictionary did not decrease but increased instead (to 17 more odd collocations in the second set of essays). Nearly half the students (12 out of 29) performed worse than when writing without using the dictionary. The number of collocations used in the second set of essays is fewer than in the first set (1,182 collocations in the first set and 1,290 in the second set). There is no clear indication of learners using a wider variety of collocations. However, most collocations looked up in the dictionary were used correctly, except for two cases where they were not used appropriately in meaning in the contexts. If we put aside the factor of different question titles in the first and the second essay, which might have led to different results, it seems that not knowing collocations that present problems was an important factor conspiring towards a higher number of odd collocations in the second essays written with the dictionary support. These could be collocations that are easily comprehensible and do not look problematic to learners in the language input, e.g. *strong coffee, follow instructions, offer help, regular service* etc. (Laufer, 2010); learners therefore did not notice when encountering them in language input or checking them in the dictionary when writing. As such, teachers need to bring these matters to learners' attention. It is evident that learners did not make use of the dictionary effectively. If they know collocations that might pose problems and approach the dictionary for help, there is a high possibility that they can use collocations looked up in the dictionary successfully.

Another possible explanation for why the dictionary did not help learners improve collocations as hypothesized is that it did not provide learners with the collocations that they needed. It is obvious that the online dictionary does not provide learners with as many collocations as its electronic version does (see section 5.5.1). Nor does it provide learners with collocations that are predictable (Benson, 1989b). However, as I have argued earlier (see section 2.2.1), it is not easy for lexicographers who are from different linguistic and cultural backgrounds from learners to decide which collocations are predictable and which are not. In reality, learners are still struggling with collocations that lexicographers consider 'predictable', such as *see a doctor* (Nakamoto, 1992), or *improve (public) transportation, improve the traffic, improve life, reduce exhaust fume,* and *internal factors*, as found in this study. The dictionary not helping learners improve collocation use could also be because learners could not find collocations they needed while in fact the dictionary provides them. Mistyping a word in the search box would result in the word not being found, and this could be the reason why learners did not find them. It could also be because learners coming from different linguistic and cultural backgrounds do not express ideas in the same way as native speakers do; hence they do not know that they could choose the collocates provided to express their intended meanings. *Strong tea* is an example. To express the idea *trà đắng* (*strong tea*), Vietnamese learners tend to look for a collocate which means 'being condensed or concentrated', and therefore might ignore the collocate *strong* suggested in the dictionary. Another example is *hail* (*mưa đá*). To express the idea *mưa đá* with *mưa* meaning *rain*, learners could possibly start searching for *rain* and then look for a collocate meaning 'being frozen or icy'. The case of *hail* in fact is nothing to do with collocation, but where appropriate should be brought to learners' attention.

The study found that students used the OOCD in combination with a thesaurus to find synonyms of collocates suggested by the OOCD. As they stated, for some headwords the OOCD does not provide many collocates. This seems to be a risky strategy and might result in odd collocations. One example relevant to this is *suffer from stress* (T20B). To avoid repeated use of this combination a student used **endure the stress* after searching for synonyms of *suffer*. Hence using the OOCD in

combination with a thesaurus in such a way is not recommended, or learners need to be advised to be vigilant when using synonyms of collocates from a thesaurus.

Findings from the survey and interview data show that the majority of learners felt more confident when expressing ideas with dictionary support. They believed that the dictionary helped them expand their knowledge of collocations and therefore would help them improve collocation use in writing. They also reported that they could easily find collocates of the searched-for words, and hence could save them time for each check-up. Comparing these with the empirical results of the study, I can conclude that the dictionary has psychologically positive impacts on the learners rather than a practical impact on their collocation use. What the dictionary brought about was a feeling of confidence and security that they had a supportive tool to rely on and that their collocations were native-like, rather than any actual improvement in their collocation use. In actual fact, empirical research concluding that such interventions did not bring about a better result is not rare. Nesselhauf's study (2005) on German-speaking learners' collocation use is one example. In that study Nesselhauf found that there is mostly no difference in the number of errors nor in the number of attempts to use collocations for compositions written with and without consulting dictionaries. Ard's study (1982) on the use of bilingual dictionaries in the writing of ESL students with Spanish, Arabic, and Japanese backgrounds also found that access to these dictionaries did not assist learners in writing accurately; conversely, learners made more of some types of errors, which are believed to be dictionary-based.

The results show a contrast between the lack of improvement in learners' collocation use and their positive evaluation of the dictionary. The possible explanation for this is that when evaluating the dictionary the learners tended to compare it with general dictionaries whose collocations are obviously not the focus, while the highly likely reason for the lack of improvement in collocation use lies in the reality that the learners did not know collocations that pose difficulties in order to consult the dictionary for help. So, as discussed in section 6.2, the dictionary can only assist learners to better use collocations when they are made more aware of the potential collocational errors.

Chapter 7 : CONCLUSION AND IMPLICATIONS

7.1   Introduction

This final chapter begins by providing an overview of the study. Section 7.3 comments on some contributions of the study. In the central part of this chapter, section 7.4, implications of the results, comprising implications for language educators, language learners, and dictionary compilers, are discussed. The limitations of the study are addressed in the next section. Section 7.6 briefly points out ways forward in the study of collocations and dictionary support for language learners. The final section presents some concluding remarks.

7.2   Overview of the study

The study aimed to investigate the impacts of the Oxford Online Collocation Dictionary on learners' collocation use, and whether or not the use of the dictionary supported learners to use collocation correctly. There were two phases of data collection. The first phase involved the collection of learners' written productions without the support of the dictionary, which was also the baseline for the research. Findings from the data showed that learners made errors with all collocation patterns, and collocation patterns in which learners made the most mistakes was V-N collocation. Then came collocations of Adj-N, N-N, N-V, Adv-V; N-of-N and Adv-Adj have the least errors. The study also found that errors can lie in any elements of the combination, and collocating words accounted for the majority. There appears to be some influence of L1 on L2 collocation use for Vietnamese learners, though the methodology did not allow this to be investigated empirically. Another problem with Vietnamese learners' collocation use found in this study is that errors are sometimes not due to the wrong combination of elements but to the inappropriate use of the combination in a given context. One finding that has not been found or claimed in other previous studies is that, in Adj-N collocations, errors are sometimes associated with the derivational process (e.g. *old-fashion<u>ed</u> minds*).

Data collected at the second stage comprised the second set of essays written with the support of the dictionary, questionnaires, recording sheets and interviews. Results

from the observation sheets and interviews indicated that learners approached the dictionary for help most frequently with collocations of Adj-N pattern, followed by the V-N pattern. The learners reported that they often sought help from the dictionary to complete every sentence while doing the writing rather than waiting until the end, emphasizing that that was their habit, and that gave them confidence that what they had written was correct and complete.

Questionnaires and interview data showed that learners sometimes used the dictionary in combination with other dictionaries. The reasons for consulting other dictionaries were finding meanings of collocates, finding collocates to express the intended meaning which could not be found in the OOCD, or finding synonyms of a collocate found in the OOCD to avoid repetition. Observation data, however, showed that to complete this writing task, the learners made relatively little use of other dictionaries. The most likely reason for the less frequent use of other dictionaries than was reported in the questionnaires and interviews is the time constraint of in-class writing. Put another way, time plays a significant role in learners' behaviour regarding the dictionary.

The data also showed that learners could locate the position of collocation types that they wanted to look for without much difficulty. Collocations looked up in the dictionary were used correctly except for a few cases where they were not used appropriately in meaning. Looking at the findings from a holistic view, however, I could say that the results of collocation use are not encouraging. Errors still occurred, even more than when writing without using the dictionary. All types of errors found from the first set of writings were found in the second set. The issue seems to lie in whether or not the learners know their possible errors in order to seek help, rather than on what the dictionary can offer or what modification or changes need to be made. To put it in another way, it is highly likely that these errors can be avoided as long as the learners are aware of the problems that some collocations can pose.

The questionnaire data showed that the dictionary is highly evaluated by the learners as a supporting tool for collocation look-ups. The learners' satisfaction was represented through their responses regarding aspects such as their commitment to

future use of the dictionary, their introduction of the dictionary to their peers, and their beliefs about collocation improvement in writing. The students' evaluation of the effectiveness, accessibility, and ease of use is also very encouraging. The interview data, however, provide a more nuanced insight into the learners' perceptions of dictionary use. Specifically, electronic dictionaries in general and online dictionaries in particular are used because they facilitate quick searches; however, internet connections decide access speed. In addition, the dictionary does not always provide the learners with collocates to express their intended meaning, and when this is the case, looking for it in another dictionary would take them more time. Another drawback of the dictionary is that it does not provide the learner with the meanings of collocates, and to make up their mind regarding which collocate to choose, they need to carry out a further search. Lack of plentiful content is also one of its limitations, which affects its quality to some extent. Complaints about the dictionary's monotonous interface as well as its lack of suggestions for looked-up words based on their initial letters were among the disadvantages reported in this study.

In general, the dictionary has more of a psychological impacts on the learners than practical impact on their collocation use. The benefit that it brings is a feeling of confidence and reassurance to the learners that their collocation use is native-like and that they have a tool to rely on whenever they have problems with collocation.

7.3   Contributions of the study

Although there are quite a lot of studies investigating the impact or role of dictionaries on collocation enhancement of language learners, they are either concerned with general dictionaries or with language learners of different language backgrounds and language levels. There is a gap in the existing knowledge with regard to the support of this specialized dictionary for learners' collocation use and their perceptions of this dictionary. Therefore, the study contributes to existing knowledge by providing an in-depth understanding of these issues.

The study is believed to be rigorous and significant in that it implemented a triangulation technique in the use of data-gathering instruments to arrive at the

answer to its research questions. By applying rigorous practices, the study made significant contributions to theory and methodology alike. Theoretically, to start with, the study provides a thorough understanding of the typical odd collocations that Vietnamese learners often make. V-N is the most problematic collocational pattern for Vietnamese learners; however, N-N collocation has the highest ratio of odd collocations over the number of collocations used. This reflects the difference in nature of the N-N collocation from the others, and indicates that the learners overgeneralize the construction of collocations of this pattern by combining two nouns regardless of conventional word combinations. Collocations of such a pattern would be better learned as big words rather than by combining two noun elements. The results also suggest that the use of collocates with very high frequency of occurrence to construct collocations does not always bring about an appropriate collocation, e.g. *a good certificate*. In addition, the wider the semantic field of a given lexical item, the more difficult it is for learners to construct collocations, e.g. *cause/create/lead to/result in/*bring/*produce + stress*.

The study also presents findings that can shed light on how this specialized dictionary was used by language learners. These findings are beneficial to language educators as well as language learners in practice. Being alert to the types of errors that they often make, the learners would be more cautious and know when they need to approach the dictionary for help. Both educators and learners might equip themselves with strategies to use the dictionary effectively to assist them on their life-long learning journey. These will be discussed in detail in the implications section. It is hoped that the study will provide additional evidence with respect to learners' needs from which dictionary compilers could make appropriate adjustments to the dictionary.

Methodologically, the study provides a practical approach to observing how learners access the dictionary for help. The recording sheets from the study by Atkins and Varantola (1997) was adapted in this study to record what was going on when the learners consulted the dictionary for help. As discussed in chapter 3, this recording activity was not carried out by the person using the dictionary because of certain drawbacks. In particular, the recorder was obviously unable to know what the

dictionary user wanted to look for or how they felt about each search to make an evaluation. The method of recording information in Harvey and Yuill (1997), using self-observation with introspective and retrospective information, resolves most of the problems with recording in the study of Atkins and Varantola (1997). However, it has its own problem. That is, learners might get distracted when they were required to complete some information on the recording sheets while they were writing. In this study, this pitfall is believed to have been reduced to a minimum since the recording sheets were assigned to be completed by both the observer and the dictionary user depending on what information was required.

Another methodological contribution of this study springs from its method for determining collocations. It involved clear and precise steps using statistical information (Log Dice score) from the BNC and native speaker informants in making judgements. The study shows that the exclusion of some combinations, not of the collocation patterns being considered, in the process of extracting collocations might result in missing errors. Take, for example, *state of emotion and *stress of relationship. They are not N-of-N collocations since the first noun is not a quantifier, but were still extracted for examination in this study. It was suggested that they are corrected to *emotional state* (Adj+N), as in (T25B) *Stress is an emotional state*, and into *relationship stress* (N-N), as in (T25B) *In daily life we are forced to work with computers and smart phones, which causes relationship stress* (or *stress to our relationship*). These suggested corrections are all strong collocations, and this proves that lack of knowledge of collocation properties leads to errors. Obviously, excluding at the early stage some combinations different to the patterns being considered might result in missing errors. This is an important contribution that research in the field needs to take into account.

One other important methodological contribution is the analysis of three or more element combinations. No other previous research has provided a thorough and clear way of dealing with these cases. In this study combinations such as *give students some disadvantages* or *keep stress under control* were extracted out as a whole for consideration. However, when checking for their conventionality and Log Dice score from the BNC, I separated them into sub-combinations: *give + students* and *give +*

disadvantages; *keep + stress* and *keep + under control*. This is because the search for the combinations as whole from the BNC would not bring up results. If either of the two sub-combinations do not co-occur more than the five-times threshold (e.g. *give + disadvantages*, *keep + stress*), the whole combination would be passed to native informants for judgement. I believe that this way of processing those combinations not only facilitated the native speakers' judgement but also helped me identify exactly where the oddness was derived from. *Keep (sth) under control* is a strong collocation with a fairly high Log Dice score of 6.4; *something* here is a quite flexible element and the use of *stress* in the combination is acceptable.

7.4 Research implications

7.4.1 Pedagogical implications

Although the overall result of collocation use with dictionary support is not encouraging, most of the collocations looked up in the dictionary were used grammatically and semantically correctly. This suggests that it is worth introducing the dictionary to language learners. The dictionary is not beneficial or harmful on its own; the key lies in how learners use it. The findings show that it is to some extent a helpful resource for collocation consultations, especially for advanced learners, who are expected to actively expand their store of vocabulary on their own rather than wait to be taught. The findings bring about a number of important implications for future practice.

When introducing the dictionary to the learner, it is believed to be important for the teacher to give them sufficient training on how to make full use of it. They need to emphasize to the learners that the dictionary provides support with collocations but the expectation of finding all of the possible collocates that can accompany the base word entry that they are searching for is not realistic. In order words, they need to make clear to the learners that the dictionary presents possible collocations; however, the list is not exhaustive. Awareness-raising about what the dictionary can offer is significant since in this study my failure to do that caused unnecessary confusion for the learners. As they reported, they were quite certain that some collocations are often used by native speakers but could not be found in the dictionary. That 'puzzlement' is not irrational as obviously, compared to its electronic version, this

online version contains far fewer collocations. Thus it is the educators' role to help learners address such unnecessary confusion.

Findings showed that there are some, though not many, indications that the dictionary helps learners use a wide variety of collocations to avoid repetition. Collocates of the same or nearly the same meaning are presented together in groups in the dictionary. Therefore, it is necessary for language teachers to introduce this feature of the dictionary to learners so that they can exploit the dictionary efficiently. Learners should also be encouraged to consult other dictionaries for meanings of collocates that they do not know before using it. This is because the research detected several cases where one collocation was chosen in place of another, which was more appropriate in meaning.

Language educators might also consider encouraging learners to use the dictionary not just to support them while doing the academic writing but to expand their store of vocabulary when doing free writing or other language tasks relevant to the production of collocations. This is because the study found that the results of collocation use depend more on quality of the search than on the quantity. The learners in this study tended to explore the dictionary in more depth and with reference to other dictionaries when doing tasks without time constraints. With an in-depth search for collocations, language input can become intake, and eventually learners can use collocations without relying too much on the dictionary.

I found from this study that more errors occurred in essays written with dictionary support than in those written without. However, almost all of those odd collocations were produced without the dictionary being consulted. A lack of awareness of the concept of collocation is less likely to be a cause for this awkwardness since the concept was carefully introduced at the beginning of the course. What seems to be important here is that learners need to be made more aware of their possible collocational mistakes. It might be helpful if language teachers focus the learners' attention more on collocations that have no direct translation from learners' L1. Errors associated with blending combinations containing the same word to express similar meanings, such as *face sth* and *to be faced with sth*, need to be brought to

greater attention since there is a high possibility that similar combinations, such as *comprise sth/be comprised of sth*, could pose a problem. Also, greater attention should be paid to collocations of N-N pattern since the ratio of the number of odd collocations over the collocations of this pattern used is 12/79 (15.2%), higher than the ratio of the V-N collocations (11.8%). The number of look-ups for help with this collocation pattern is rather low, at only 4. Learners should also be made aware that overgeneralizing the rule of forming N-N or N-of-N collocations based on combining two nouns could result in errors.

With the advent of electronic dictionaries, especially online dictionaries, learners can easily access them without worrying about carrying heavy bulky traditional paper dictionaries. Search speed is confirmed as one of the most outstanding advantages of electronic dictionaries (Chon, 2009; Chun, 2004), and this was reasserted in this study. The use of this online dictionary for collocation consultation while doing the writing was found to possibly distract learners' flow of thought. Therefore, it is a good idea to suggest learners use the dictionary sensibly depending on whether or not they are doing the writing under time constraints.

Another important implication for dictionary users is that they could use the dictionary to polish their writing by looking up collocations of which one element is not obligatory, such as Adj-N, Adv-V, or Adv-Adj. The dictionary could also be used as a facilitator to look for suggestions for collocates, especially when they have no ideas for how to complete the phrasal meaning.

7.4.2   Implications for dictionary compilers

In the discussion chapter (section 6.5), I presented some improvements suggested by the learners. In this section, I will briefly present some improvements that dictionary compilers need to take into account.

Interview data showed that the dictionary did not receive positive responses in terms of the number of examples for illustration. Yet all the collocations which were looked up in the dictionary are grammatically correctly used in the written texts. This suggests that illustrative examples, though not as plentiful as expected, are probably

enough for the target learners at advanced level. What dictionary compilers might need to consider is to add pronunciation for each word entry since, as students reported, it fails to assist learners in using collocations for the other productive purpose, speaking. This, to some extent, degrades learners' evaluation of the collocation dictionary for productive use.

Evidence from the study also showed that the dictionary sometimes failed to provide learners with information that they needed to use collocations correctly in meaning. Some students complained that it sometimes took time to search for the meaning of collocates that they do not know before making choices as to which collocate to use. Therefore, it would be a good idea for the dictionary compilers to provide the learner with the meanings of collocate groups. Another possible solution for this is that they might consider creating links between this dictionary and the Oxford Advanced Learner's Dictionary in providing the meaning of each collocate.

The presentation of collocations in the dictionary at the base entry, based on the argument that learners will think first of a base and then look for a collocate to complete the phrasal meaning (Hottsrnonn, 1991), is fundamentally appropriate. To Vietnamese learners, nevertheless, the presentation of N-N collocations in the dictionary seems to be the odd one out. The first noun of the combination is the noun base rather than the second, and to search for a collocation of this pattern learners have to start with the first noun, whereas Vietnamese learners tend to think first of the second noun. When searching for N-N collocations such as *bus stop*, *mountain bike*, or *education policy* starting with the second nouns, dictionary users will not find them at the N-N collocation section but at the Adj-N pattern (e.g. *bus top*, *mountain bike*) or not find them at all (e.g. *education policy*). It is therefore suggested that dictionary compilers might need to re-examine their presentation of collocations of this pattern in the dictionary.

7.5   Limitations of the study

The study was designed to provide some useful insights into the effects of the Oxford Online Collocation Dictionary on the collocation use of advanced language learners; it is not free from limitations. These limitations lead to a number of suggestions and

recommendations for future research on how the use of this dictionary can be further investigated. The limitations are as follows.

The participants in the study are English major students coming from the University of Social Sciences and Humanities, which is not representative of Vietnamese learners at advanced level. Including participants from different universities might bring about more variety of errors and would add more perspectives on the value of the dictionary.

Another limitation is that although I was fully aware of the possible impacts of my role as a teacher and a researcher in the study and many attempts were made to minimize them, it is hard to say for certain how far those resolutions worked. Playing the two roles at the same time, to some extent, affected the participants' responses to the questionnaires, interviews, and behaviour towards the dictionary. Making use of the advantages of the dual role means I had to accept the influence that it brought.

Third, the study investigated the impact of the online collocation dictionary on the collocation use of language learners and thus involved the comparison of collocation use without and with dictionary support. However, there was a period when learners were assigned some writing tasks to familiarise themselves with the dictionary. Any improvements in collocation use might, therefore, be attributable to that practise rather than to the support of the dictionary alone. An additional uncontrolled factor is the possibility that the learners' store of collocations was affected by the process of learning because besides academic writing, the learners were taking several other modules at the same time. There is no guarantee that these did not affect the learners' collocations.

Besides, although I did my best to make adjustments as well as to give clear instructions on how to complete the observation sheets, I recognized that some of the students still got confused with the question asking if they could find instructions on how to use the collocation looked up in the dictionary. This might result in the participants having skipped the question or not answering it accurately, and could therefore skew the research result.

Another limitation that might impact on the result of the study is that learners were encouraged to use the dictionary to support them with their writing tasks at home in order to become familiar with the dictionary. However, there was a small group of students living in rented houses who were unable to get access to the online dictionary due to the unavailability of the internet. To deal with this, I helped to install the electronic version of this dictionary onto their laptops. The two versions, electronic and online, in theory are 'twins', but in reality they are quite different in terms of the number and examples of collocations, as well as the interface. Although I encouraged them to get access to the online dictionary by using computers from the English Language Centre, stressing that this was the dictionary they would be using for the second writing, the likely possibility is that not all of them tried to get used to it. This also means that these students were not truly used to it.

Finally, in the process of determining whether or not a combination is acceptable, the study relied on native speakers' intuition, which is inherently quite subjective. My effort to achieve more objective results was to invite native speakers of both American and British English and process the combinations through two rounds. No matter what effort I put into it, this limitation is obviously totally unavoidable.

7.6   Suggestions for future research

The representative sample of the study, though not small, is not big enough, and the study is qualitative in nature. It is thus limited in its generalizability. To arrive at findings that can enable generalization, future research could consider carrying out the research on a larger sample with a quantitative research design. Furthermore, this research only targeted learners at upper intermediate to advanced level. It would be helpful to carry out the same research with learners at lower levels, pre-intermediate, intermediate or even at beginner level, to see if the dictionary could help.

This study only looks at the use of the dictionary for help in writing, but the dictionary can be used for help with the other productive skill, speaking. Future studies could thus examine if and how the dictionary helps learners to improve their collocation use in speaking. In reality, learners not only consult the dictionary when

they are doing the writing but also expand their store of vocabulary in their learning process. Accordingly, an investigation of learners' collocation competence after a period of use of the dictionary would be interesting to carry out. Research in that direction is worth conducting since in the longer term learners are expected to be able to use collocations confidently without relying on any supporting tools.

The findings indicate that there may be some influence, both positive and negative, of learners' L1 on their L2 collocation use. Understanding whether there is any influence and how far the influence extends would be very helpful to language learners and educators alike. This would be an interesting point to investigate further. From the study there emerged some initial indications of the influence of learners' L1 on their use of prepositions in three or more element collocations. Learners' cultural background appears to have some impact as well. Therefore, future research could focus exclusively on one of these aspects.

The study found that the proportion of odd collocations of N-N pattern over collocations used is the highest of all the collocation patterns that I was investigating. Learners' generalization of the rule of forming collocations of this pattern reflects the possible differences in nature of N-N collocations and other collocation patterns. The differences in the constraints for N-N collocations, as opposed to V-N or collocations of other patterns, might also be an interesting point for future research to explore.

Lastly, most of the general dictionaries contain collocations and the presentation of collocations in these dictionaries is constantly improving. It would be interesting to carry out research comparing the effects of the use of general dictionaries and this specialized dictionary on learners' collocation competence, as well as their perceptions of the dictionaries.

7.7   Concluding remarks

What I have learned over the last four years and the results of the study has greatly contributed to my professional career. The skills that I have gained during the research process are invaluable in education settings where I am playing multiple roles, from teacher to language learner to researcher. In particular, observing the

students using the dictionary to support them in collocation use has triggered some interesting ideas for making use of the dictionary in activities to widen learners' store of collocations in my EFL classes, learners at lower levels included. Looking back at the starting point when I faced a lot of challenges, I find myself considerably more matured as a researcher. These challenges have taught me a lot about myself and my own capacity. More importantly, the thesis has aroused my passion and enthusiasm for further educational research.

List of References

Aarts, B. 2013. *English syntax and argumentation.* Basingstoke, UK: Palgrave Macmillan.

Adam, K. and Iztok, K. 2012. *Corpus tools for lexicographers.*

Aisenstadt, E. 1981. Restricted collocations in English lexicology and lexicography*: Review of Applied Linguistics.* **53**, pp. 53-61.

Ard, J. 1982. The use of bilingual dictionaries by ESL students while writing. *ITL Review of Applied Linguistics.* **58**, pp.1-27.

Atkins, B. S. and Rundell, M. 2008. *The Oxford Guide to Practical Lexicography.* Oxford: Oxford University Press.

Atkins, B. T. and Varantola, K. 1997. Monitoring dictionary use. *International Journal of Lexicography.* **10** (1), pp. 1-45.

Atkins, B. T. S. and Varantola, K. 1998. Language learners using dictionaries: The final report on the EURALEX/AILA research project on dictionary use. In: Atkins, B.T.S. ed. *Using Dictionaries: Studies of Dictionary Use by Language Learners and Translators.* Tubingen: Niemeyer, pp.21-82.

Bahns, J. 1991. Lexical Collocations: A contrastive view. *ELT Journal.* **47** (1), pp. 51-63.

Bahns, J. and Eldaw, M. 1993. Should we teach EFL students collocations? *Science Direct.* **21** (1), pp. 101-114.

Baker, P. 2006. *Using Corpora in Discourse Analysis.* London & NewYork: Continuum.

Banh, L. T. 2004. *VỀ VIỆC BAN HÀNH BỘ CHƯƠNG TRÌNH KHUNG GIÁO DỤC ĐẠI HỌC KHỐI NGÀNH NGOẠI NGỮ TRÌNH ĐỘ ĐẠI HỌC..* Ha Noi Viet Nam.

Batia, L. and Hill, M., H. 2000. *What Lexical Information Do L2 Learners Select in a CALL Dictionary and How Does It Affect Word Retention?* Hong Kong/ Israel: Educational Resources Information Centre.

Béjoint, H. 1989. The teaching of dictionary use: Present state and future tasks. *Franz Josef Hausmann, Oskar Reichmann, et al. (Hgg.): Wörterbücher. Ein internationales Handbuch zur Lexikographie.* **1**, pp. 208-215.

Bell, J. 2014. *Doing Your Research Project: A guide for first-time researchers.* England: McGraw-Hill Education (UK).

Benson, M. 1985. Collocations and idioms. *Dictionaries, lexicography and language learning.* pp. 61-68.

Benson, M. 1989a. The collocational dictionary and the advanced learner. *DOCUMENT RESUME.* p. 106.

Benson, M. 1989b. The structure of the collocational dictionary. *International Journal of Lexicography.* **2** (1), pp. 1-14.

Benson, M., Benson, E. and Ilson, R. 1986. *Lexicalgraphic Description of English.* Amsterdam: John Benjamins.

Benson, M., Benson, E. and Ilson, R. 1997. *The BBI Dictionary of English Word Combinations.* Amsterdam: John Benjamins.

Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E. and Quirk, R. 1999a. *Longman Grammar of Spoken and Written English.* Harlow: Longman.

Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E. and Quirk, R. 1999b. *Longman grammar of spoken and written English.* MIT Press.

Biskup, D. 1992. L1 influence on learners' renderings of English collocations: A Polish/German empirical study. In:. *Vocabulary and Applied Linguistics.* Basingstoke: Palgrave Macmillan.

Blass, L. and Pike-Baky, M. 2007. *Mosaic.* England: McGraw-Hill Higher Education.

Bogaards, P. 1996. Dictionaries for Learners of English. *International Journal of Lexicography.* **9** (4), pp. 277-320.

Bogaards, P. 2003. Uses and users for dictionaries. In: Van Sterkenburg, P. ed. *A Practical Guide to Lexicography.* Amsterdam/Philadenphia: John Benjamins..

Braun, V. and Clarke, V. 2006. Using thematic analysis in psychology. *Qualitative research in psychology.* **3** (2), pp. 77-101.

Brett, P. 1996. Using multimedia: An investigation of learners' attitudes. *Computer Assisted Language Learning.* **9** (2-3), pp. 191-212.

Bruton, A. 2007. Vocabulary learning from dictionary referencing and language feedback in EFL translational writing. *Language Teaching Research.* **11** (4), pp. 413-431.

Bryman, A. 1988. *Quantitative and Qualitative in Social Research.* London: Unwin Hyman Ltd.

Bryman, A. 2015. *Social Research Methods.* Oxford: Oxford University Press.

Bulut, D. and AbuSeileek, A. F. M. 2009. Learner attitude toward CALL and level of achievement in basic language skills.*Journal of Instiute and Social Sciences of Erciyes University.* 23 (2), pp. 103-126.

Burke, J.R. 1997. Examining the validity structure of qualitative research. *Education.* **118** (2), pp. 282-293.

Burton-Roberts, N. 2016. *Analysing Sentences: An Introduction to English Syntax.* London and New York: Routledge.

Carter, R. and McCarthy, M. 1988. *Vocabulary and Language Teaching.* London: Longman.

Carter, R. and McCarthy, M. 2006. *Cambridge Grammar of English: a Comprehensive Guide; Spoken and Written English Grammar and Usage.* Cambridge: Cambridge University Press.

Charmaz, K. 2015. Grounded theory. In: Smith, J. A. ed. *Qualitative Psychology: A Practical Guide to Research Methods.* Los Angeles/London/New Delhi/Singapore: Sage.

Châu, M. N., Yamatomo, R., Shimizu, M. and Tomika, K. 2011. Một số lỗi sinh viên Nhật thường hay mắc phải khi viết Tiếng Việt. In: *Hội Thảo Nghiên Cứu và Giảng Dạy Tiếng Việt Lần Thứ Nhất, Osaka University, Japan*.

Chomsky, N. 2014. *Aspects of the Theory of Syntax.* MIT press.

Chon, Y. V. 2009. The electronic dictionary for writing: A solution or a problem? *International Journal of Lexicography.* **22** (1), pp. 23-54.

Christianson, K. 1997. Dictionary use by EFL writers: what really happens? *Journal of second language writing.* **6** (1), pp.23-43.

Chun, Y. V. 2004. EFL learners' use of print and online dictionaries in L1 and L2 writing processes. *Multimedia Assisted Language Learning.* **7** (1), pp. 9-35.

Clarke, V., Braun, V. and Hayfield, N. 2015. Thematic Analysis. In: Smith, J.A. ed. *Qualitative Psychology: A Practical Guide to Research Methods.* Los Angeles/London/New Delhi/Singapore: Sage.

Clear, J. 1993. From Firth principles: Computational tools for the study of collocation. In: Baker, M. et al. *Text and Technology: In Honour of John Sinclair.* Amsterdam/Philadenphia: John Benjamins.

Coffey, S. 2010. A new pedagogical dictionary of English collocations. *International Journal of Lexicography.* **24** (3), .p. 323.

Coffin, C., Donohue, J. and North, S. 2013. *Exploring English Grammar: From Formal to Functional.* London: Routledge.

Cop, M. 1990. The function of collocations in dictionaries. In: *Euralex Proceedsings, Budapest.* Akadémiai Kédó, pp. 35-46.

Cowie, A. P. 1981. The treatment of collocations and indioms in Learners' dictionaries. *Applied Linguistics.* **2** (3), pp. 233-235.

Cowie, A. P. 1988. Stable and creative aspects of vocabulary use. In: Carter, R. and McCarthy, M. eds. *Vocabulary and Language Teaching.* London: Longman.

Cowie, A. P. 1992. Multiword lexical units and communicative language teaching. In: Arnaud, P. J. L. and Bejoint, H. eds. *Vocabulary and Applied Linguistics.* London: Palgrave Macmillan, pp.1-12.

Cowie, A. P. 1994. Phraseology. In: Ronald, E. A. ed. *The Encyclopedia of Language and Linguistics.* Oxford: Pergamon, pp. 3168-3171.

Cowie, A. P. and Howarth, P. 1996. Phraseological competence and written proficiency. In: Blue, G. M. and Mitchell, R. eds. *British Studies In Applied Linguistics.* Clevedon/Philadelphia/Adelaade: Multilingual Matters, pp. 80-93.

Cowie, A. P. ed 1998. *Phraseology: Theory, Analysis, and Applications.* Oxford: Clarendon Press.

Dang, D. 2014. Errors in Lexical collocations by Academic Writing students at Faculty of English Linguistics and Literature. In: *English Language Teaching in Asean Coutries, Ho Chi Minh.* University of Social Sciences and Humanities, pp. 41-46.

Denzin, N.K. and Lincoln, Y. S. 2003. *Collecting and Interpreting Qualitative Materials.* Los Angeles/London/New Delhi/Singapore: Sage.

Downing, A. and Locke, P. 2006. *English Grammar: A University Course.* London and New York: Routledge.

Drew, P. 2005. Conversation analysis. In: Fitch, C. L. and Sanders, R. E. eds. *Handbook of Language and Social Interaction.* New Jersey: Lawrence Erlbaum Associates, pp. 71-102.

Dziemianko, A. 2010. Paper or Electronic? The Role of Dictionary Form in Language Reception, Production, and the Retention of Meanings and Collocations. *International Journal of Lexicography.* **23** (3), pp. 257-273.

Dziemianko, A. 2014. On the presentation and placement of collocations in monolingual English learners' dictionaries: Insights into encoding and retention. *International Journal of Lexicography.* **27** (3), pp. 259-279.

Durrant, P. and Schmitt, N. 2009. To what extent do native and non-native writers make use of collocations? *IRAL-International Review of Applied Linguistics in Language Teaching.* **47**(2), pp.157-177.

East, M. 2008. *Dictionary Use in Foreign Language Writing Exams: Impact and Implications.* Amsterdam/Philadelphia: John Benjamins Publishing.

Firth, J. R. 1935. The technique of semantics. *Transactions of the Philological Society.* **34** (1), pp. 36-73.

Firth, J. R. 1957. Applications of general linguistics. *Transactions of the Philological Society.* **56** (1), pp.1-14.

Gablasova, D., Brezina, V. and McEnery, A. M. 2016. Collocations in corpus-based language learning research: identifying, comparing and interpreting the evidence. *Language Learning.* **67** (51), pp.155-179.

Golafshani, N. 2003. Understanding reliability and validity in qualitative research. *The qualitative report.* **8** (4), pp. 597-606.

González Fernández, B. and Schmitt, N. 2015. How much collocation knowledge do L2 learners have? *ITL-International Journal of Applied Linguistics.* **166** (1), pp. 94-126.

Granger, S. 1998. Prefabricated Patterns in Advanced EFL Writing: Collocations and Formulae. In: Cowie, A. P. ed. *Phraseology: Theory, Analysis, and Applications.* Oxford: Clarendon Press.

Greenbaum, S. 1996. *The Oxford English Grammar.* Oxford: Oxford University Press

Halliday, M. A. K. and Hasan, R. 1976. *Cohesion in English.* London: Longman.

Hartmann, R. R. 2001. *Teaching and researching lexicography.* England: Pearson Education.

Harvey, K. and Yuill, D. 1997. A study of the use of a monolingual pedagogical dictionary by learners of English engaged in writing. *Applied Linguistics.* **18** (3), pp. 253-278.

Hatherall, G. 1984. Studying dictionary use: some findings and proposals. In: *Euralex proceedings, London.* pp.183-189.

Hausmann, F. J. 1989. Le dictionnaire de collocations. *FJ Hausmann et al*.pp. 1010-1019.

Herbst, T. 2010. Valency constructions and clause constructions or how, if at all, valency grammarians might sneeze the foam off the cappuccino. In: Schmid, H.-J. and Handl, S. eds. *Cognitive Foundations of Linguistic Usage Patterns. Berlin and Boston: De Gruyter Mouton.* Germany: Hubert & Co. GmbH & Co. KG, pp. 225-256.

Hill, J., Morgan, L. and Lewis, M. 2000. Classroom strategies, activities, and exercises. In: Lewis, M. ed. *Teaching Collocation: Further Developments in the Lexical Approach.* USA: Thomson Heinle.

Hoang, V. V. 2010. The current situation and issues of the teaching of English in Vietnam.

Hornby, A. S., Cowie, A. P. and Lewis, J. W. 1974. *Oxford Advanced Learner's Dictionary of Current English.* London: Oxford University Press.

Hottsrnonn, F.J. 1991. Collocations in monolingual and bilingual English dictionaries. In: Ivir, V. and Kalogjera, D. eds. *Languages in Contact and Contrast: Essays in Contact Linguistics.* Berlin and New York: Mouton de Gruyter, pp. 225-236.

Howarth, P. 1996. *Phraseology in English Academic Writing: Some Implications for Language Learning and Dictionary Making.* Germany: Tubingen Niemeyer.

Howarth, P. 1998. The Phraseology of Learners' Academic writing. In: Cowie, A.P. ed. *Phraseology: Theory, Analysis, and Applications.* Oxford: Clarendon Press.

Huang, L.-S. 2001. Knowledge of English Collocations: An Analysis of Taiwanese EFL Learners. In: *Texas Foreign Languguage Education Conference, Texas.* Educational Resources Information Center.

Hunston, S. 2002. *Corpora in applied linguistics.* Cambridge: Cambridge University Press.

Jacobs, G. M. 1989. *Dictionaries Can Help Writing--If Students Know How To Use Them.* Thailand: Educational Resources Information Center.

Jones, M. and Haywood, S. 2004. Facilitating the acquisition of formulaic sequences. In: Schmitt, N. ed. *Formulaic sequences.* Amsterdam: John Benjamins, pp. 269-300.

Jones, S. and Sinclair, J. 1974. English lexical collocations. *Cahiers de lexicologie.* **24** (1), pp. 15-61.

Joppe, M. 2010. The research process. 2000. *Retrieved February.* **25**.

Kent, D. 2001. Korean university freshmen's dictionary use and perceptions regarding dictionaries. *The Korea TESOL Journal.* **4** (1), pp. 73-92.

Kilgarriff, A.a.R., Smrž, P. and Tugwell, D. 2004. Itri-04-08 the sketch engine. *Information Technology. 105, 116*

Knight, S. 1994. Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities. *The Modern Language Journal.* **78** (3), pp. 285-299.

Koren, S. 1997. Quality versus convenience: Comparison of modern dictionaries from the researcher's, teacher's and learner's points of view. *TESL-EJ.* **2** (3), pp. 1-16.

Koya, T. 2003. A study of collocation in English and Japanese noun-verb combinations. *Intercultural Communication Studies.* **XII** (1), pp.125-145.

Koyama, T. and Takeuchi, O. 2004. Comparing electronic and printed dictionaries: How the difference affected EFL learning. Kansai University: Jacet Bulletin: 38. pp. 33-46

Kozlowski, A. and Seymour, L. 2003. The importance of collocation in English language teaching. In: *Contact, University of Ottawa*. TESL Ontario, pp. 41-49.

Krashen, S. 1982. *Principles and Practice in Second Language Acquisition.* Oxford: Pergamon.

Kroll, J. F., Michael, E., Tokowicz, N. and Dufour, R. 2002. The development of lexical fluency in a second language. *Second Language Research.* **18** (2), pp. 137-171.

Laforest, M. 1980. Towards a typology of lexical errors. *Anglistika.* **23**, pp. 1-30.

Laufer, B. 2010. The contribution of dictionary use to the production and retention of collocations in a second language. *International Journal of Lexicography.* **24** (1), pp. 29-49.

Laufer, B. and Waldman, T. 2011. Verb-noun collocations in second language writing: a corpus analysis of learners' English. *Language Learning.* **61** (2), pp. 647–672.

Le, V. C. 2011. *Form-focused instruction: A case study of Vietnamese teachers' beliefs and practices*. PhD thesis, University of Waikato.

Lea, D. 2007. Making a Collocations Dictionary. *Zeitschrift für Anglistik und Amerikanistik.* **55** (3), pp. 261-271.

LEA, D. & RUNCIE, M. Blunt Instruments and Fine Distinctions: a Collocations dictionary for students of English. Proceedings of the Tenth EURALEX International Congress. Copenhagen: Center for Sprogteknologi, 2002. 829.

Lew, R. 2004. *Which dictionary for whom?: receptive use of bilingual, monolingual and semi-bilingual dictionaries by Polish learners of English.* Poland: Motivex.

Lew, R. and Doroszewska, J. 2009. Electronic Dictionary Entries with Animated Pictures: Lookup Preferences and Word Retention. *International Journal of Lexicography.* **22** (3), pp. 239–257.

Lewis, M. 2000. Teaching Collocation. In: Lewis, M. ed. *Teaching Collocation: Further Developments in the Lexical Approach.* USA: Thomson Heinle.

Lewis, M. and Conzett, J. 2000. *Teaching Collocation: Further Developments in the Lexical Approach.* Hove: Language Teaching Publications.

Lincoln, Y. S. and Guba, E. G. 1985. *Naturalistic Inquiry.* London and Newdelhi: Sage.

Liu, C. 1999. An analysis of collocational errors in EFL writings. In: *The Proceedings of the Eighth International Symposium on English Teaching, Taipei.* Crane, pp. 483-494.

McAlpine, J. and Myles, J. 2003. Capturing phraseology in an online dictionary for advanced users of English as a second language: a response to user needs. *System.* **31** (1), pp. 71-84.

McIntosh, A. and Halliday, M. A. K. 1969. *Patterns of Language. Papers in General, Descriptive and Applied Linguistics.* London: Longman.

McIntosh, A. 2009. *Oxford collocations dictionary for student of English.*

Mel'cuk, I. 1998. Collocations and Lexical Functions. In: Cowie, A. P. ed. *Phraseology: Theory, Analysis, and Applications.* Oxford and New York: Clarendon Press.

Moon, R. 1998. *Fixed Expressions and Idioms in English. A Corpus- based Approach.* Oxford: Clarendon Press.

Nakamoto, K. 1992. *Lexical Collocation in Dictionaries of English as a Foreign Language.* MA thesis, Tokyo University of Foreign Studies

Nation, I. S. 2001. *Learning Vocabulary in Another Language.* Cambridge: Cambridge University Press.

Nattinger, J. R. and DeCarrico, J. S. 1992. *Lexical Phrases and Language Teaching.* Oxford: Oxford University Press.

Nesi, H. 1999. A user's guide to electronic dictionaries for language learners. *International journal of lexicography.* **12** (1), pp. 55-66.

Nesi, H. 2000. Electronic dictionaries in second language vocabulary comprehension and acquisition: The state of the art. In: *Proceedings of the Ninth Euralex international congress, EURALEX*, pp. 839-847.

Nesi, H. and Haill, R. 2002. A study of dictionary use by international students at a British university. *International Journal of Lexicography.* **15** (4), pp. 277-305.

Nesselhauf, N. 2003. The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics.* **24** (2), pp. 223-242.

Nesselhauf, N. 2005. *Collocations in a Learner Corpus.* Amsterdam: John Benjamins

Nesselhauf, N. and Tschichold, C. 2002. Collocations in CALL: An investigation of vocabulary-building software for EFL. *Computer Assisted Language Learning.* **15** (3), pp. 251-279.

Nuccorini, S. 2003. Towards an 'ideal' Dictionary of English Collocations. In: Van Sterkenburg, P. ed. *A Practical Guide to Lexicography.* Amsterdam and Philadelphia: John Benjamins.

Oshima, A. and Hogue, A. 2000. *Writing Academic English.* Cambridge: Pearson/Longman.

Partington, A. 2004. "Utterly content in each other's company": Semantic prosody and semantic preference. *International Journal of Corpus Linguistics.* **9** (1), pp. 131-156.

Pawley, A. and Syder, F. H. 1983. Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In: Richards, J. C. and Schmidt, R. W. eds. *Language and Communication.* New York: Longman.

Philip, G. 2011. *Colouring Meaning: Collocation and Connotation in Figurative Language.* Amsterdam/Philadelphia: John Benjamin.

Phoocharoensil, S. 2012. Cross-linguistic influence: Its impact on L2 English collocation production. *English Language Teaching.* **6** (1), pp. 1-10.

Roby, W. B. 1999. What's in a gloss. *Language Learning and Technology.* **2** (2), pp. 94-101.

Rubin, H. J. and Rubin, I. S. 2011. *Qualitative Interviewing: The Art of Hearing Data.* Los Angeles/London/New Delhi/Singapore/Washington DC: Sage.

Rychlý, P. 2008. A lexicographer-friendly association score. In: *Proceedings of Recent Advances in Slavonic Natural Language Processing, RASLAN, Masaryk University.* pp. 6-9.

Sapsford, R. and Jupp, V. 2006. *Data Collection and Analysis.* London/Thousand Oaks/New Delhi: Sage.

Schmitt, N. and Carter, R. 2004. Formulaic sequences in action: An introduction. In: Schmitt, N. ed. *Formulaic Sequences.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Sinclair, J. 1987. *Collins COBUILD English Language Dictionary.* London: Harper Collins Publishers.

Sinclair, J. 1991. *Corpus, Concordance, Collocation.* Oxford: Oxford University Press.

Sinclair, J. 2004. *Trust the Text: Language, Corpus and Discourse.* London: Routledge.

Smith, J. A. and Osborn, M. 2015. Interpretive phenomenological analysis In: Smith, J. A. ed. *Qualitative Psychology: A Practical Guide to Research Methods.* Los Angeles/London/New Delhi/Singapore: Sage.

Stenbacka, C. 2001. Qualitative research requires quality concepts of its own. *Management Decision.* **39** (7), pp. 551-556.

Stubbs, M. 1995. Collocations and semantic profiles: on the cause of the trouble with quantitative studies. *Functions of language.* **2** (1), pp. 23-55.

Stubbs, M. 2002a. Two quantitative methods of studying phraseology in English. *International Journal of Corpus Linguistics.* **7** (2), pp. 215-244.

Stubbs, M. 2002b. *Words and Phrases: Corpus Studies of Lexical Semantics.* Oxford: Blackwell.

Summers, D. 1988. The role of dictionaries in language learning. In: Carter, R. and McCarthy, M. eds. *Vocabulary and Language Teaching.* London: Routledge, pp. 111-125.

Tono, Y. 2001. *Research on Dictionary Use in the Context of Foreign Language Learning: Focus on Reading Comprehension.* Tubingen: Niemeyer.

Trochim, W. and Donnelly, J. 2008. *The Research Methods Knowledge Base.* Boston: Cengage Learning.

Tsai, K.-J. 2015. Profiling the collocation use in ELT textbooks and learner writing. *Language Teaching Research.* **19**(6), pp.723-740.

Webb, S., Newton, J. and Chang, A. 2013. Incidental Learning of Collocation. *A Journal of Research in Language Studies.* **63** (1), pp. 91-120.

Weinreich, U. 1969. *Problems in the Analysis of Idioms: Substance and Structure of Language.* California: University of California Press.

Wolter, B. and Gyllstad, H. 2011. Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge. *Applied Linguistics.* **32** (4), pp. 430-449.

Wray, A. 2000. Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics.* **21** (4), pp. 463-489.

Wray, A. 2002. *Formulaic Sequences and the Lexicon.* Cambridge: Cambridge University Press.

Wright, J. 1998. *Dictionaries.* Oxford: Oxford University Press.

Yoon, H. and Hirvela, A. 2004. ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing.* **13** (4), pp. 257-283.

Appendix 1: Link to access the OOCD

http://oxforddictionary.so8848.com/

Online Collocation Dictionary

A completely new type of dictionary with word collocation that will help students and advanced learners **effectively study, write and speak natural-sounding English**. This *online dictionary* is also very helpful for the education of the IELTS, TOEFL test.

Level: Upper Intermediate to Advanced

Key features of Oxford Dictionary Online

1. Collocations/collocation - common word combinations such as *'bright idea'* or *'talk freely'* - are the essential building blocks of natural-sounding English. The dictionary contains over 150,000 collocations for nearly 9,000 headwords.

2. The dictionary shows all the words that are commonly used in combination with each headword: nouns, verbs, adjectives, adverbs, and prepositions as well as common phrases.

3. The collocation dictionary is based on the 100 million-word British National Corpus. Internet searches were made to ensure most up-to-date usage for fast-changing areas of language like computing.

4. Over 50,000 examples show how the collocation/collocations are used in context, with grammar and register information where helpful.

5. The clear page layout groups collocations according to part of speech and meaning, and helps users pinpoint speedily the headword, sense and collocation they need.

6. Usage notes show collocation/collocations shared by sets of words such as languages and seasons.

7. It is an ideal companion volume to the Oxford Advanced Learner's Dictionary.

8. **Free Download -- OXFORD Collocations Dictionary**

☐ IELTS Speaking Topics (part 1,2,3)

- [IELTS Essay Writing Topics](#)

- [IELTS Writing Ideas](#)

- [JIC NEW](#)

- [Except or except for?](#)

- [food name -- Chinese-English](#)

- [A picture for House/home vocabulary](#)

- [what time is Noon, Afternoon, evening, Night](#)

Appendix 2: Recording sheets

Dictionary user's number:

Ordinal number	Words check-up	Types of collocation								Words found: Y/N	Purpose of looking up		Instruction on how to use: Y/N	Use in combination with other dictionaries Y/N	Detailed description	Evaluation
		V-N	N-V	Adj-N	Adv-V, V-Adv	Adv-Adj	N-N	N-of-N	Others		Checking (C)	Finding (F)				
1	Discuss (v)									Y	√		Y	N		Not satisfied Ok

Appendix 3: Questionnaires

Dictionary user's number: [ ]

The following statements are regarding your evaluation of the use of OOCD as an assisting tool for writing. Please use the scale below to tick (√) the response that most resembles your perspectives.

1. Strongly agree	or	1. Always
2. Agree		2. Often
3. Disagree		3. Occasionally
4. Strongly disagree		4. Never

Number	Statements	1. Strongly agree	2. Agree	3. Disagree	4. Strongly disagree
1	I can easily look up collocates of a word.				
2	With internet availability I can access this dictionary easily.				
3	Compared to paper dictionaries, this dictionary saves me time for each check-up.				
4	The layout of meanings, grammatical use and frequently used expressions is user-friendly.				
5	I am confident in expressing ideas when using this dictionary.				
6	The OOCD is helpful for expanding my collocation knowledge.				
7	With this assisting tool, I use more collocations if applicable.				
8	The OOCD will help improve my writing skills.				
9	I prefer the OOCD to general dictionaries for looking up collocations.				
10	I will use OOCD to assist my English writing in the future.				
11	I would recommend using OOCD to my friends.				

Number	Statements	1. Always	2. Often	3. Occasionally	4. Never
12	I can get instructions on how to use collocations through examples.				
13	When I search for information in OOCD, I get the information that I need.				
14	I use this dictionary in combination with other dictionaries to find out how to use a collocation.				
15	I use this dictionary to support my writing.				
16	I have difficulty in making a choice of collocates found in the OOCD.				
17	It takes me time to look for meanings of collocates from other dictionaries.				

Appendix 4: Interview Questions

With participant number:

Thank you very much for your participation in the research project titled:

Impacts of the OCD on advanced learners' collocation competence in L2 writing

1. How often do you use the OOCD?
2. Do you often use the OOCD in combination with other dictionaries?
3. How do you evaluate the use of the OOCD as a supporting aid to L2 writing?
4. What are likeable and dislikeable features of the OOCD?
5. What are desirable features of the OOCD?
6. Are there any difficulties in using the OOCD to support your writing?
7. Will you continue using the OOCD?

Appendix 5: Amended recording sheets

Dictionary user's number:

#	Question		Answer
1	What entry were you looking up?		Discuss (v)
2	What types of collocations were you looking for?	V-N	
		N-V	
		Adj-N	
		Adv-V, V-Adv	√
		Adv-Adj	
		N-N	
		N-of-N	
		Others	
3	Did you find the word you were looking for?		Y
4	Why did you look up the word?	Finding (F)	
		Checking (C)	√
5	Did you find the instruction on how to use it?		Y
6	Did you use OOCD in combination with other dictionaries?		N
7	How satisfied were you with what you found? *Choose the scale from 1-5 with 1: very dissatisfied to 5: very satisfied*		4
8	Other comments/why did you use other dictionaries? If yes		

Appendix 6: First writing

Duration: 45 mins

Participant's number

Choose <u>one</u> of the two topics below and write an essay of around 350 words.

1. *Many high-level positions in companies are filled by men even though the workforce in many developed countries is more than 50 per cent female. Companies should be required to allocate a certain percentage of these positions to women. To what extent do you agree?*

2. *Nowadays many students have the opportunity to study for part or all of their courses in foreign countries. While studying abroad brings many benefits to individual students, it also has a number of disadvantages. Do you agree or disagree?*

Topic number:

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

Appendix 7: Exercises to become familiar with the OOCD

1 Ideas into words

1.1 Look at the entry for IDEA (ADJ section). Find an Adj you might use to express the following ideas. Sometimes more than one adj is possible.

a. An idea that is helpful, rather than being negative or impractical
b. An idea that is slightly crazy, in a good way
c. An idea that is completely crazy, in a bad way
d. An idea that has not been carefully thought out
e. An idea that seems very impressive but is not really very practical

1.2 Look at the section marked Verb + idea. Find verb that you might use to express:

a. To find an idea
b. To suggest an idea
c. To suggest an idea in a very forceful way because you really want people to accept it
d. To think about an idea for a while before you decide whether or not it is a good idea
e. To talk about a number of different ideas before you decide which ideas are the best

1.3 Look at the section marked Idea + verb. Find verbs that you might you to express:

a. When you think of an idea
b. When an idea develops into something important
c. When an idea does not develop into anything

2. Using a noun entry

2.1 Match each of the adjs on the left with a suitable noun from the facing column

a bewildering	**ambition**	a blazing	**defeat**
a biting	**array of goods**	a crushing	**inflation**
a burning	**chance**	a haunting	**pain**

a convincing	**chasm**	a nagging	**row over money**
driving	**rain**	a piercing	**scream**
a fighting	**sum of money**	raging	**statement**
a gaping	**win**	a sprawling	**suburb**
a staggering	**wind**	a sweeping	**melody**

2.2 Quantifiers

Quantifiers are words used to talk about the amount of sth, such as a drop of water. Complete each sentence with a suitable quantifier.

a. There were just a few w_____ of cloud in the sky.
b. The recent s_____ of attacks has made residents afraid to leave their homes.
c. He is on medication to ease his frequent b_____ of depression.
d. I just caught a brief s_____ of their conversation as I walked by their table.
e. The constant s_____ of traffic pass our house makes it difficult to cross the road.
f. A p_____ of stray dogs was wandering around the abandoned plant.
g. He's been off school all week with a bad d_____ of flu.
h. A couple of c_____ of garlic will improve the flavour of the soup.
i. The manager terrified the younger staff with his o_____ of temper.
j. The burglars stole several p_____ of jewellery.

2.3 Cross out any verbs that do not collocate with the bold nouns. (VERB +)

a. He *got into/had/made* an **argument** with the barman and was thrown out of the hotel.

b. He had to do two jobs to clear/pay off/pay up his **debts**.

c. Someone *came up with/presented/put forward* the **suggestion** that we should have an auction.

d. The scientists failed to *arrive at/decide/draw* any firm **conclusions** from the study.

e. The company *agreed/came to/struck* a **deal** with the union after lengthy negotiations.

f. A **meeting** has been *arranged/programmed/scheduled* for next week.

g. The supervisor refused to *accept/receive/shoulder* the **blame** for the accident.

h. He *drummed/rattled/tapped* his **fingers** nervously on the desk as he spoke.

i. We *did/took/went* on a **trip** to a nearby island on a fishing boat.

j. I put up my hand to *shade/shelter/shield* my **eyes** from the sun.

2.4 Complete the story with a suitable verb in each gap. (+ VERB)

I lay in bed, unable to sleep. The wind had f-_____ freely at the party, and now my head was t_____ and my stomach was c_____.
Outside the wind h and the rain l_____ against the window. My nerves were o_____ e_____ as I remembered all the horror films I'd ever seen. Suddenly I heard the key t_____ in the front door. My heart began to h_____ in my chest as heavy footsteps e_____ on the stairs. My mind was r_____, trying to think how I could save myself. The bedroom door c_____ open slowly, and as my eyes a_____ to the darkness I could make out a figure at the end of the bed. The man's mouth f_____ open when he saw me. It was my neighbour. I was in the wrong house.

3. Using a verb entry

3.1 In each of the following sentences one of the adverbs in italics is not a common collocate of the verb in bold. Decide which it is & cross it out.

a. She argued *fiercely/heatedly/hotly* about her right to compensation.
b. They will *fiercely/heatedly/hotly* defend their rights.
c. He grinned *owlishly/sheepishly/wolfishly* at her.

d. I *ruefully/sheepishly/woefully* confessed to having forgotten the map.

e. His frugal lifestyle contrasted *brutally/markedly/starkly* with his wife's extravagance.

f. Her tragic story *brutally/markedly/starkly* illustrates how vulnerable children can be.

3.2 Complete each of these sentences with a verb phrase from the box.

Be determined to	**be happy to**	**can afford to**	**fail to**
hasten to	**offer to**	**serve to**	**take**
steps to			

a. I _____ accept the invitation to become patron of the charity.

b. The company was fined when it _____ comply with the regulations.

c. These unanswered questions _____ highlight the potential problems.

d. I _____ add that my knowledge of computers is pretty basic.

e. We must _____ ensure that such a disaster can never happen again.

f. The minister _____ resign when the affair became public.

g. She _____ fight for her rights.

h. Few patients _____ pay the full cost of treatment.

4. Using an ADJ entry

4.1 Match each of the bold adjs with a verb that can go before it, then match the combination with a suitable subject.

His mistake	emerged	asleep
_____		
His mistake	fall	costly
_____		
I nearly	grew	damp
_____		
The crowd	passed	empty
_____		
The driver	proved	impatient
_____		
The house	runs	parallel
_____		

The house	smells	unnoticed

_____

The road	stood	unscathed

_____

4.2 Complete each sentence with a suitable verb

a. He d _____ me crazy with his constant talking.

b. She was h _____ captive by rebels for six months.

c. Several cars were s _____ ablaze by the rioters.

d. The unions were r _____ powerless by the new laws.

e. These programmes are d _____ unsuitable for screening before 10pm.

f. The robbers b _____ the shopkeeper senseless.

g. His classmates mostly r _____ him as eccentric.

h. The sound of a door banging j _____ me awake.

4.3 For each group, find an adv in the box that collocates with all the adjs in the group.

Dead	Distinctly	Fiercely	Grossly	Painfully	wildly
a.	b.	c.	d.	e.	f.
aware	boring	competitive	different	enthusiastic	inaccurate
honest	funny	independent	odd	inaccurate	inadequate
shy	good	loyal	uncomfortable	optimistic	offensive
slow	right	protective	uneasy	popular	unfair

4.4 Match each bold adj with a suitable adverb. Then use each combination to complete one of the sentences on the right.

blissfully	**absent**	a. I'm not _____ by the largest figures.
conspicuously	**composed**	
eerily	**concerned**	b. She is _____ of her achievements.
notoriously	**familiar**	
oddly	**fickle**	c. He seems_____ of the trouble he

outwardly *unduly* *justly*	**proud** **silent** **unaware**	*caused.* *d. The former chairman* was _____ from the guest list. e. Her voice sounded _____ to me. f. She seemed _____ despite the pressure. g. The street was _____ after the explosion. h. The world of fashion is _____.

5. Collocations with common verbs

do	make	have	take	give
a crossword	an appointment	an accident	action	sb an answer
damage	an attempt	an argument	a bath	sb a chance
a degree	the bed	a bath	a bite	a cry of pain
an exam	a cake	a break	a break	sb help
the food for a party	changes	breakfast	the bus	sb an idea
French at school	a decision	cancer	a decision	the impression that
the garden	dinner	a chat	a deep breath	sb a kiss
your hair	an effort	a cold	sb's details	sb lessons
Hamlet	a film	difficulty	a dislike to sb	sb a lift
your homework	a fuss	a drink	an exam	your opinion
judo	a guess	a feeling	a guess	a party
miles per hour	an impression	fun	a holiday	a performance
nothing	a mark	a guess	an interest in sth	sth a polish

do	make	have	take	give
Paris	a mess	a heart attack	a look	sb a present
a photocopy	a mistake	a holiday	a nap	priority to sth
research	money	an idea	notes	sb a push
a sketch	a noise	an interest	a photo	sb a shock
a translation	peace	a look	size 10	a sigh
the washing	a photocopy	a meeting	a tablet	a speech
some writing	progress	a party	sb's temperature	some thought to
	a promise	a nap	a walk	a welcome to sb
	a sketch	an operation		
	a speech	patience		
	a suggestion	problems		
	your will	a shock		
		a snack		

a. Find the nouns in the lists for tasks and duties (e.g. do the dishes). Which verb is the most often used? Which tasks are exceptions?
b. Find expressions in each column that can be substituted by a single verb (e.g. you can *do damage* to sth or just *damage* sth). Which column has the most?

Appendix 8: Collocations in learners' notebooks

An example of how collocations of the word *Medicine* are taught and kept in students' notebooks.

OCD entry for **Medicine** (noun):

---

medicine *noun*

[1] science of treating/preventing illness

ADJ. **modern** *advances in modern medicine* | **traditional** *qualified in traditional Chinese medicine* | **folk** *Garlic was widely used in folk medicine.* | **conventional, orthodox | alternative, complementary, holistic, homeopathic | preventative, preventive | academic, clinical, forensic, scientific | general** *She gave up general medicine to specialize in geriatric medicine.* | **geriatric, obstetric, paediatric, veterinary, etc. | Chinese, Western | private, public health** *She believed private medicine was a threat to the existence of the National Health Service.*

VERB + MEDICINE **train in | qualify in | practise** *people practising alternative medicine*
PHRASES **a branch of medicine** > Note at SUBJECT(for more verbs and nouns)
[2] substance taken to treat an illness
ADJ. **powerful, strong | cough** *a bottle of cough medicine* | **herbal | prescription**
QUANT. **dose**
VERB + MEDICINE **take | swallow | prescribe (sb) | administer, give sb | treat sb with**
MEDICINE + NOUN **bottle, chest**
PREP. **~ for** *medicine for a chest infection*

Medicine (n): a. (Science of treating an illness)

         Adj: *Conventional / general / traditional*

Ex: *Qualified in traditional medicine*

       She gave up general medicine to specialize in conventional medicine

         b. (substance taken to treat an illness)

       Adj: *powerful / cough*

         V: *take / swallow / treat sb with* (~~drink~~)

Appendix 9: Feedback on writing assignments

Academic writing

Student ID: 1357010232

What causes people with disabilities to be unable to get jobs?

Working is the right and duty of everybody living in society because it is the principle source that maintains the whole community. Therefore, at a certain age, people have to start a career to make a living and to set themselves free from the support and protection of their families. However, not everyone who has become adults already can be employed. Especially with people who were born with disabilities; this situation is even worse due to these three reasons.

First of all, appearances significantly affect the opportunity a disabled person can get a job. People are easily scared by someone with a different look which identifies him or her as alienated or abnormal. Although this is just a metaphor usually used to describe, the influence it brings about can ruin a person's job interview at the very first moment they walk into the room no matter what they dress. Since the staff's appearance become a category to fit into, it is getting harder for people with disabilities to find a job when more and more companies consider the employees as the face of the whole organization. Obviously, employees are the ones mostly seen entering, working in and leaving the building. With a team of all nice-looking workers, the image of the enterprise can be built up.

In addition, lack of work capacity can be counted as one of the main causes making disabled people unemployed. According to research made by the U.S department of Labour, there is only 17,1% of people with disabilities are hired in 2014. Many normal people think of them as the burden of society despite of the fact that not all of them are unable to work. Nowadays, the employers usually look for the working performance judging on physical appearances rather than the actual abilities. According to so many interviewers, a healthier person can definitely earn greater achievement. Besides, they are afraid of the payment they have to spend on health insurance and other treatments if there is any incident. Furthermore, having flaws in the body structures means that the performance of the disabled is low. A wheel chair can cause them move slowly. An arm missing can prevent them from typing fast. Yet none can blame them for what they were born with. Still, it makes the employers frowning to reach a wise decision.

Finally, the preconception about disabled people can be one of the biggest obstacles for those who wish for a job. It is believed that people with disabilities always show their anger at their loss or at anyone without disabilities. For that reason, sometimes they refuse to get help from the society, which turns them into jobless. Regardless of the fact that this number is not very substantial, it is still one essential factor leading to the unemployment. Additionally, people tend to have compassion for the weakness of the disabled rather than considering them as people who have full awareness and productive capacity. As they are giving mercy instead if judging a person on his or her real performance, they can withdraw their favours at any time.

Lastly, the preconceived opinion that the disabled need to be fixed or isolated from the others since they are the faults of nature is why they cannot get on well in that environment. Since they are not accepted mentally, it is not easy for them to be accepted physically.

Although society has some good moves to the disabled people, they still receive so many discriminations. Such main causes include weakness of appearance, the lack of work capacity, and the judgement form the society.

Appendix 10: Second writing

Duration: 45 mins

Participant's number: ☐

Choose <u>one</u> of the two topics below and write an essay of around 350 words.

1. *In order to solve traffic problems, governments should tax private car owners heavily and use the money to improve public transportation. What are the advantages and disadvantages of such a solution?*

2. *Stress is now a major problem in many countries around the world. What are some of the factors in modern society that cause this stress, and how can we reduce it.*

Topic number:

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

Appendix 11: Judgement of combinations from T12A, T22A by the British and American informants

N	Combinations	Ame.			Bri.			Final	Context	American Informant	British Informant
		(+)	(-)	(?)	(+)	(-)	(?)				
I	**TEXT 12A**							**8**			
1	face (a lot of) differences in		(-)			(-)		(-)	They have to face a lot of differences in various fields such as lifestyle …	... have to confront ...	They find differences in various life aspects such as lifestyle
2	means lack of support		(-)		(+)				being far away from home means lack of family support	… means loss of support ...	a lack of family support
3	come back to (their) hometown		(-)		(+)				... after they come back to their hometown.	… return to …	
4	issues get on		(-)			(-)		(-)	financial issues usually get on foreign students' nerves	Foreign students usually face financial issues	Slang not proper English

N	Combinations	Ame.			Bri.			Final	Context	American Informant	British Informant
		(+)	(-)	(?)	(+)	(-)	(?)				
5	daily expense	(+)			(+)			(+)			
6	learning career		(-)			(-)		(-)		Being educated for a career	A career in learning?
7	chief cook	(+)					(?)				Head Chief?
8	bottle washer	(+)			(+)			(+)			
9	students' nerves		(-)			(-)		(-)	financial issues usually get on foreign students' nerves.	financial issues usually present problems	financial issues can be annoying for students
10	expense costs		(-)		(+)				daily expense cost a great deal of money	high daily expenses are great	costs a great deal of money
11	the case of a student	(+)			(+)			(+)	… especially in the case of a student from a developing country come to a developed one		coming to a developed one

N	Combinations	Ame.			Bri.			Final	Context	American Informant	British Informant
		(+)	(-)	(?)	(+)	(-)	(?)				
12	learned abroad	(+)			(+)			(+)	to apply what they have learned abroad to their countries		learnt
13	balance studying									balance studying with working	balance studying with working
	TEXT 22A							10			
1	give students some disadvantages		(-)			(-)		(-)	studying abroad also gives students some disadvantages	students studying abroad present some disadvantages	There are also some disadvantages for students who study abroad
2	adapt to the (new) culture	(+)			(+)			(+)	they cannot adapt to the new culture in foreign countries		
3	put up with (their negative) feelings		(-)			(-)		(-)	they cannot put up with their negative feelings	they cannot live up to	they cannot endure the negative feeling
4	have (many financial) issues	(+)			(+)			(+)			

N	Combinations	Ame.			Bri.			Final	Context	American Informant	British Informant
		(+)	(-)	(?)	(+)	(-)	(?)				
5	neglect (their) study	(+)			(+)			(+)	Some students just focus on earning money and completely neglect their studies		
6	students commit	(+)			(+)			(+)			
7	tuition fee (which may) affect	(+)			(+)			(+)	... in order to pay their high tuition fee, which may affect their study to some extent		fees
8	students …neglect	(+)			(+)			(+)			
9	marks get		(-)		(+)				their marks at school get worse and worse	their marks/grades at school got worse and worse	got worse and worse
10	high tuition fee	(+)			(+)			(+)			fees
11	cost of tuition fee		(-)		(+)					cost of tuition	fees

N	Combinations	Ame.			Bri.			Final	Context	American Informant	British Informant
		(+)	(-)	(?)	(+)	(-)	(?)				
12	completely strange	(+)			(+)			(+)	students will feel isolated in a completely strange country		
	TEXT 12B										
1	bring stress to people		(-)			(-)		(-)	social relationship also brings stress to people	social relationship also causes stress to people	social relationships can also be stressful
2	enduring (extreme) stress	(+)					(?)		Living in modern society means enduring extreme stress		Living in modern society can be extremely stressful
3	minimize effects of stress	(+)			(+)			(+)			the effects
4	pay (a lot of) effort		(-)			(-)		(-)	worker must pay a lot of effort in enhancing productivity	pay a lot of attention	make a lot of effort
5	remain (social) relationship		(-)				(?)		people find hard to remain social relationship.	maintain social relationship	

N	Combinations	Ame.			Bri.			Final	Context	American Informant	British Informant
		(+)	(-)	(?)	(+)	(-)	(?)				
6	result in stress	(+)			(+)			(+)	There are several causes that <u>result in stress</u> …		
7	support the elderly	(+)			(+)			(+)			
8	choose (negative) reaction			?		(-)			they seem to choose negative reaction	to have negative reaction	seem to have a negative reaction
9	adults (have to) support	(+)			(+)			(+)	Adults also have to support the elderly in family		in their family
10	causes (that) result in	(+)			(+)			(+)	There are several causes that result in stress such as ...		
11	employee (have to) keep	(+)			(+)			(+)	employees have to keep improving theirselves	employees have to keep improving **themselves**	themselves
12	relationship brings	(+)			(+)			(+)	social relationship also brings stress to people.		relationships

N	Combinations	Ame.			Bri.			Final	Context	American Informant	British Informant
		(+)	(-)	(?)	(+)	(-)	(?)				
13	swimming … proved	(+)			(+)			(+)	Yoga and swimming are proved to be good activities to relax	Yoga and swimming **have** proved to be good …	swimming are proven to be good
14	Yoga … proved	(+)			(+)			(+)		Yoga has proved to be …	proven
15	tremendous responsibility	(+)			(+)				Adults retain a tremendous responsibility in caring for children	Adults **have** a tremendous …	
16	family's responsibility	(+)			(+)			(+)	they can balance between work and family's responsibilities		family
	TEXT 22B										
1	boost … energy	(+)			(+)			(+)	... because it helps people refresh their minds and boost their energy as well		
2	caused … environment	(+)			(+)			(+)	stress is mainly caused by work pressure and polluted environment.		and the polluted ...

N	Combinations	Ame.			Bri.			Final	Context	American Informant	British Informant
		(+)	(-)	(?)	(+)	(-)	(?)				
3	make use of time	(+)			(+)			(+)	you can make use of time efficiently		
4	provide people with … disadvantages			?		(-)			stress has become more and more common in many nations across the globe and provided people with many disadvantages.	Problem with "provide"	has become increasingly common ... provided people with many challenges
5	put people under pressure		(-)			(-)		(-)	people are put under pressure to earn a living	people are pressured to earn a living	people are under increasing pressure to earn
6	result in stress	(+)			(+)			(+)	People should find ways to prevent noise and air pollution that are taking place in big cities and partly result in stress among people		People should find ways to prevent the noise and air pollution common in big cities which can result in stress.
7	suffer from… traffic jam		(-)		(+)				people may have to suffer terrible traffic jam	people may have to suffer terrible traffic	people might have to suffer a terrible

N	Combinations	Ame.			Bri.			Final	Context	American Informant	British Informant
		(+)	(-)	(?)	(+)	(-)	(?)				
8	spent on relaxation	(+)				(-)			Time spent on relaxation is very pivotal		time spent relaxing
9	environment exert	(+)			(+)			(+)	environment exerts a tremendous influence on people's life		the environment ... peoples' lives
10	measures ... solve	(+)			(+)			(+)	government should take measures to solve environmental issues		
11	overwork … put	(+)				(-)			Overwork just puts you under pressure with lots of negative feelings		overwork leads to negative feelings because of the pressure ?
12	pollution … result in		(+)			(+)		(+)	... air pollution that are taking place in big cities and partly result in stress among people.	… air pollution in big cities results in stress among people	... air pollution in big cities can create stress for citizens
13	pollution … seem	(+)			(+)			(+)	noise pollution and air pollution seem to become more familiar with people and also put people under pressure		seems to ... puts people

N	Combinations	Ame.			Bri.			Final	Context	American Informant	British Informant
		(+)	(-)	(?)	(+)	(-)	(?)				
									nowadays		
14	pollution … take place	(+)			(+)			(+)	People should find ways to prevent noise and air pollution that are taking place in big cities and partly result in stress among people		People should find ways to prevent the noise and air pollution that are taking place in big cities. It creates stress for the citizens.
15	pollution … put	(+)			(+)			(+)	noise pollution and air pollution seem to become more familiar with people and also put people under pressure nowadays		
16	problem … explained	(+)			(+)			(+)	this problem is best explained by work pressure		
17	stress … provided		(-)			(-)		(-)	stress has become more and more common in many nations across the	causes many disadvantages	increasingly common

N	Combinations	Ame.			Bri.			Final	Context	American Informant	British Informant
		(+)	(-)	(?)	(+)	(-)	(?)				
									globe and provided people with many disadvantages.		
18	they (students) … strive	(+)			(+)			(+)	What they (students) are striving for is a better life for their families		
19	alarming issues	(+)			(+)			(+)			
20	appropriate arrangement	(+)			(+)			(+)			
21	terrible traffic jam		(-)		(+)					terrible traffic	
22	victim of stress	(+)			(+)			(+)			
23	arrange appropriately	(+)			(+)			(+)			appropriately arrange

N	Combinations	Ame.			Bri.			Final	Context	American Informant	British Informant
		(+)	(-)	(?)	(+)	(-)	(?)				
24	make use of … efficiently		(-)			(-)		(-)	C	make efficiently use of	efficiently make use of
25	very pivotal	(+)			(+)			(+)			
26	really unpleasant	(+)			(+)			(+)			
27	take the stress		(-)			(-)		(-)	Unfortunately, some people even have to give up their jobs as they cannot take the stress	they cannot withstand the stress	endure the stress
28	people's life									on someone's life/ on people's lives	on someone's life/ on people's lives
29	people's awareness										