LEXICAL COLLOCATIONS

IN A SAMPLE CORPUS OF NURSING RESEARCH

ARTICLES (SCNRA) AND THE EFFECTS OF CORPUS-

BASED INSTRUCTION ON STUDENT'S COLLOCATION

LEARNING AT A THAI UNIVERSITY

Kantapat Trinant

A Thesis Submitted in Partial Fulfillment of the Requirements for

the Degree of Doctor of Philosophy in English Language Studies

Suranaree University of Technology

Academic Year 2017

คำปรากฏร่วมในคลังข้อมูลตัวอย่างภาษาในงานวิจัยตีพิมพ์ด้านพยาบาลศาสตร์
และผลกระทบของการสอนแบบอิงคลังข้อมูลภาษาที่มีต่อการเรียนรู้
คำปรากฏร่วมของนักศึกษาพยาบาลในมหาวิทยาลัยไทย

นายกันตพัฒน์ ตรินันท์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาศิลปศาสตรดุษฎีบัณฑิต
สาขาวิชาภาษาอังกฤษศึกษา
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2560

LEXICAL COLLOCATIONS IN A SAMPLE CORPUS OF NURSING RESEARCH ARTICLES (SCNRA) AND THE EFFECTS OF CORPUS-BASED INSTRUCTION ON STUDENTS' COLLOCATION LEARNING AT A THAI UNIVERSITY

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

Thesis Examining Committee

I. Pramoolsook

(Asst. Prof. Dr. Issra Pramoolsook)

Chairperson

(Dr. Butsakorn Yodkamlue)

Member (Thesis Advisor)

P. Kani

(Asst. Prof. Dr. Kornwipa Poonpon)

Member

(Dr. Adcharawan Buripakdi)

Member

(Dr. Sirinthorn Seepho)

Member

W. Polnigongit

(Assoc. Prof. Dr. Weerapong Polnigongit)

Dean of Institute of Social Technology

_____

(Prof. Dr. Santi Maensiri)

Vice Rector for Academic Affairs

and Internationalization

กันตพัฒน์ ตรินันท์ : คำปรากฏร่วมในคลังข้อมูลตัวอย่างภาษาในงานวิจัยตีพิมพ์ด้าน
พยาบาลศาสตร์และผลกระทบของการสอนแบบอิงคลังข้อมูลภาษาที่มีต่อการเรียนรู้คำ
ปรากฏร่วมของนักศึกษาพยาบาลในมหาวิทยาลัยไทย (LEXICAL COLLOCATIONS IN A
SAMPLE CORPUS OF NURSING RESEARCH ARTICLES – SCNRA – AND THE
EFFECTS OF CORPUS-BASED INSTRUCTION ON STUDENTS' COLLOCATION
LEARNING AT A THAI UNIVERSITY)
อาจารย์ที่ปรึกษา : ดร.บุษกร ยอดคำลือ, 334 หน้า

งานวิจัยนี้มีวัตถุประสงค์ 4 ประการ คือ (1) เพื่อระบุและจำแนกคำศัพท์ (Keywords) ที่พบ
ในคลังข้อมูลตัวอย่างภาษาในงานวิจัยตีพิมพ์ด้านพยาบาลศาสตร์ (SCNRA) ซึ่งได้รับการตีพิมพ์ใน
วารสารวิชาการด้านการพยาบาล (2) เพื่อค้นหาคำปรากฏร่วมในคลังข้อมูลตัวอย่างภาษาในงานวิจัย
ตีพิมพ์ด้านพยาบาลศาสตร์ (SCNRA) โดยใช้คำศัพท์ที่พบเป็นคำหลัก (Nodes) เพื่อหาคำปรากฏ
ร่วมของคำศัพท์เหล่านั้น และจำแนกคำปรากฏร่วมที่ได้ตามชนิดของการปรากฏร่วมที่เกิดขึ้นตาม
ประเภทของคำที่ปรากฏร่วมกัน (3) เพื่อประเมินความรู้ด้านคำปรากฏร่วมที่พบกับนักศึกษา
พยาบาลศาสตร์ ชั้นปีที่ 4 ของมหาวิทยาลัยเทคโนโลยีสุรนารี โดยนำคำปรากฏร่วมที่พบจาก
การศึกษาข้างต้นมาทำแบบทดสอบ และ (4) เพื่อนำคำปรากฏร่วมที่ได้จากการศึกษาไปถ่ายทอดให้
ความรู้แก่นักศึกษาพยาบาลศาสตร์และประเมินผลของการสอนแบบอิงคลังข้อมูลภาษา (Corpus-
based Instruction) ที่มีต่อคะแนนของนักศึกษาเปรียบเทียบระหว่างก่อนและหลังการสอน

งานวิจัยส่วนแรก เริ่มด้วยการรวบรวมบทความงานวิจัยที่ตีพิมพ์ในวารสารวิชาการด้านการ
พยาบาล จำนวน 300 เรื่อง จากวารสารวิชาการจำนวน 10 วารสาร ที่เข้าถึงได้ผ่านบริการของศูนย์
บรรณสาร มหาวิทยาลัยเทคโนโลยีสุรนารี โดยบทความที่รวบรวมแต่ละเรื่องถูกบันทึกเป็นไฟล์
นามสกุล .txt เมื่อรวบรวมได้ครบจำนวนแล้ว ไฟล์เหล่านี้ถูกโหลดเข้าไปในโปรแกรมจัดการ
คลังข้อมูลภาษาชื่อ AntConc เพื่อสร้างคลังข้อมูลตัวอย่างภาษาในงานวิจัยตีพิมพ์ด้านพยาบาล
ศาสตร์ (SCNRA) ซึ่งประกอบด้วยจำนวนคำกว่า 1.25 ล้านคำ จากการประมวลผลภายใต้เกณฑ์ที่
กำหนด คลังข้อมูลตัวอย่างภาษาที่สร้างขึ้นนี้มีคำศัพท์ (Keywords) จำนวน 717 คำ ซึ่งคำศัพท์
เหล่านี้ถูกนำไปใช้เป็น คำหลัก (Nodes) เพื่อหาคำปรากฏร่วม ผลลัพธ์คือ มีคำปรากฏร่วมจำนวน
2,14 8 คู่ แบ่งออกเป็น 14 ประเภทของคำที่ปรากฏร่วมกัน โดยคำปรากฏร่วมที่พบมากที่สุดคือคำ
ปรากฏร่วม นาม+นาม ($N = 889$; 41.39%) ลำดับที่สองคือคำปรากฏร่วม คุณศัพท์+นาม ($N = 610$;
28.4%) ลำดับที่สามคือ คำปรากฏร่วม นาม+กริยา ($N = 240$; 11.17%) ในส่วนของคำศัพท์ ส่วน

ใหญ่เป็นคำนาม (N = 463; 63.51%) ตามมาด้วยคำคุณศัพท์ (N = 157; 21.54%) ลำดับที่สามคือ คำกริยา (N = 98; 13.44%) และคำวิเศษณ์มีจำนวนน้อยที่สุด (N = 11; 1.51%)

        งานวิจัยส่วนที่สอง ว่าด้วยความรู้เกี่ยวกับคำปรากฏร่วมด้านการพยาบาลของนักศึกษา พยาบาลศาสตร์ ซึ่งเริ่มด้วยการสร้างแบบทดสอบจำนวน 60 ข้อ ประกอบด้วย แบบตัวเลือก จำนวน 30 ข้อ แบบเติมคำในช่องว่าง จำนวน 20 ข้อ และแบบเขียนประโยค 10 ข้อ และใช้ทดสอบกับ นักศึกษาพยาบาลศาสตร์ ชั้นปีที่ 4 จำนวน 51 คน โดยมีการทดสอบก่อนเรียน การอบรมให้ความรู้ เกี่ยวกับคำปรากฏร่วมที่พบจากการวิจัย และการทดสอบหลังเรียน ผลการทดสอบก่อนเรียนพบว่า ค่าเฉลี่ยคะแนนรวม = 30.66 โดยมีส่วนเบี่ยงเบนมาตรฐาน = 7.41 ($\bar{x}$ = 30.66; SD = 7.41) ซึ่งบ่งชี้ว่า โดยรวมแล้วนักศึกษามีความรู้เกี่ยวกับคำปรากฏร่วมอยู่ในระดับ "พอใช้" ผลการทดสอบหลังเรียน พบว่า ค่าเฉลี่ยคะแนนรวม = 39.44 มีส่วนเบี่ยงเบนมาตรฐาน = 7.65 ($\bar{x}$ = 39.44; SD = 7.65) ซึ่งบ่งชี้ ว่าโดยรวมแล้วนักศึกษามีความรู้เกี่ยวกับคำปรากฏร่วมอยู่ในระดับ "ดี" ผลคะแนนระหว่างการ ทดสอบก่อนเรียนและหลังเรียนเปรียบเทียบโดยใช้ paired samples t-test แสดงให้เห็นว่านักศึกษา ทำคะแนนได้ดีขึ้นอย่างมีนัยสำคัญ ที่ระดับน้อยกว่า 0.001 ทั้งคะแนนโดยรวมและแต่ละส่วนของ แบบทดสอบ ในส่วนของความสามารถของนักศึกษา พบว่านักศึกษาส่วนใหญ่ทำคะแนนได้ดีขึ้น หนึ่งระดับ จากระดับ "พอใช้" เป็นระดับ "ดี" ยกเว้นเว้นในส่วนที่สาม การเขียนประโยค ที่พบว่า นักศึกษาส่วนใหญ่ทำคะแนนได้ดีขึ้นหนึ่งระดับ จากระดับ "อ่อน" เป็นระดับ "พอใช้"

สาขาวิชาภาษาต่างประเทศ   ลายมือชื่อนักศึกษา _____

ปีการศึกษา 2560   ลายมือชื่ออาจารย์ที่ปรึกษา _____

KANTAPAT TRINANT : LEXICAL COLLOCATIONS IN A SAMPLE CORPUS OF NURSING RESEARCH ARTICLES (SCNRA) AND THE EFFECTS OF CORPUS-BASED INSTRUCTION ON STUDENTS' COLLOCATION LEARNING AT A THAI UNIVERSITY. THESIS ADVISOR : BUTSAKORN YODKAMLUE, Ph.D., 334 PP.

CORPUS-BASED STUDY/ CORPUS-BASED INSTRUCTION/ NURSING COLLOCATIONS/ COLLOCATION TEST/ COLLOCATION KNOWLEDGE

The present study was conducted to (1) identify and classify keywords in the Sample Corpus of Nursing Research Articles (SCNRA); (2) explore lexical collocations found in the SCNRA using keywords extracted as 'nodes' to find their 'collocates', and to classify the collocation pairs according to their combination types; (3) assess collocation knowledge of nursing students at Suranaree University of Technology (SUT) based on the test constructed from collocations extracted from the SCNRA; and (4) provide lessons applying corpus-based instruction on nursing collocations and assess the effects on the students' performance.

The research started with the compilation of 300 research articles from 10 selected academic journals in the field of nursing accessible online via SUT's Library Resources by saving into text files then uploaded onto a corpus tool, the AntConc, to form the SCNRA which comprised over 1.25 million running words. 717 keywords were obtained under the set criteria which then were used as the "nodes" to find their collocates. 2,148 pairs of lexical collocations were obtained with 14 combination types. The majority of them were Noun + Noun (N = 889; 41.39%) followed by

Adjective + Noun (N = 610; 28.4%), and Noun + Verb (N = 240; 11.17%). In terms of the keywords, the majority of them were Nouns (N = 463; 63.51%) followed by Adjectives (N = 157; 21.54%), and Verbs (N = 98; 13.44%) with Adverbs came last (N = 11; 1.51%).

In terms of collocational knowledge of nursing students, Nursing Collocation Test was constructed containing three parts: 30 multiple-choice items; 20 gap-filling items; and 10 items of a sentence writing task. The Test was then administered with 51 fourth year nursing students at SUT before and after a corpus-based instruction on nursing collocations organized for them. The pre-test's mean of the total score revealed that the overall knowledge of collocations of the majority of the students was at a "Fair" level ($\bar{x}$ = 30.66; SD = 7.41). The post-test's mean of the total score showed that the overall performance of the majority of them had improved to a "Good" level ($\bar{x}$ = 39.44; SD = 7.65). To evaluate the effects of corpus-based instruction on the students' performance, the results of the pre-test and the post-test were compared using paired samples t-test. The results showed that there was a statistically significant improvement of the students' performance in the total score ($t(50)$ = (-11.75), p = <0.001) as well as in each of the three parts of the test: part 1 ($t(50)$ = (-7.47), p = <0.001); part 2 ($t(50)$ = (-8.10), p = <0.001); and part 3 ($t(50)$ = (-7.44), p = <0.001). In terms of the students' performance, the majority of the students' scores increased and put them one level higher in all parts from "Fair" to "Good", except for part 3, a sentence writing task, the level of performance was improved from "Poor" to "Fair".

School of Foreign Languages   Student's Signature _K. Toinant_

Academic Year 2017   Advisor's Signature _____

ACKNOWLEDGEMENTS

Fifthly, I would like to thank the Faculty of Liberal Arts and Ubon Ratchathani University for granting me full time study leave permission with full pay. I would also like to thank all my colleagues at the Department of Western Languages and Literature and the Faculty of Liberal Arts for their moral support.

Last but not least, I am thankful for all my teachers who taught me in primary school, middle school, high school, and universities. Without them, I would not be the person who I am today. My family and friends, near and far, you know that you have always been a powerful force for me to keep moving forward. I really appreciate you all.

Kantapat Trinant

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

Page

LIST OF TABLES

LIST OF TABLES (Continued)

LIST OF TABLES (Continued)

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACL   Academic Collocation List

BNC   British National Corpus

CLT   Communicative Language Teaching

COCA   Corpus of Contemporary American English

DDL   data-driven learning

EAP   English for Academic Purpose

EFL   English as a foreign language

ESL   English as a second language

ESP   English for Specific Purposes

ETS   Educational Testing Service

FSs   Formulaic sequences

IMRD   Introduction, Methods, Results, and Discussion

L1   First language

L2   Second language

MFT   Meaning-focused task

MFI   Meaning-focused instruction

MI   Mutual information

MRAs   Mutual Recognition Arrangements

NAWL   Nursing Academic Word List

NRAC   Nursing Research Articles Corpus

SCNRA   Sample Corpus of Nursing Research Articles

SLA   Second language acquisition

SUT   Suranaree University of Technology

CHAPTER 1

INTRODUCTION

This chapter presents background of the present study which leads to the rationale, purposes of the study and the research questions. The significance of the study is also provided afterwards, followed by the scope and limitations of the present study. Then the definitions of key terms are given. The chapter ends with the summary of the chapter.

1.1 Background

Today's era of globalization has led to the need for greater integration among groups of countries. The European Union (EU) is the first of this kind. The ASEAN is currently attempting to follow suit by announcing the economic integration – The ASEAN Economic Community (AEC) – by the end of the year 2015. The 10 member countries have agreed on several matters in order to facilitate the economic integration among them. One important agreement is the mobilization of professional services. This would be proceeded under the Mutual Recognition Arrangements (MRAs). Under the MRAs, there will be the free flow of professional-level labor within the member countries in seven fields, which are engineering, architectural, surveying, medical, dental, nursing, and accounting services (Aldaba, 2013).

Although there seems to be more opportunities for skilled labor to work outside of their countries, what seems to prevent them from doing so, besides other

matters, is the language barrier. Professional knowledge alone will not be sufficient for professionals who seek to pursue their career outside of their motherland where the language used is not their mother tongue. This is particularly a concern for Thai professionals since English skills seem to be the major constraint for them. As clearly stated in article 34 of the ASEAN Charter (2008) that English is the working language of the ASEAN community, it is crucial for these professionals to be fluent in English in order to gain better opportunities in their careers abroad.

In the teaching and learning of English as a foreign language (EFL) or English as a second language (ESL), vocabulary and grammar have always been regarded as essential elements. Nation and Waring (1997) suggest that EFL/ESL learners need to know about 3,000 high frequency words in order to use the language to communicate successfully. Similarly, Sökmen (1997) states that knowing the 2,000 most frequent words in English can be very helpful for the learners. This belief is reflected in several methods of teaching, namely, grammar translation, direct method, and audio lingual. Even in the more recently introduced methods such as communicative language teaching (CLT), though grammar rules may not be strictly in focus, a large repertoire of vocabulary is still emphasized (Richards, 2006). This incident is supported by Nation (2011) as he claims that although the CLT approach "initially had a largely negative effect on the deliberate teaching and learning of vocabulary, teachers have continued to see the importance of giving direct attention to words" (p. 535). Milton (2013) also asserts that vocabulary knowledge is an effective indicator of the four skills of the language.

However, simply knowing a lot of vocabulary and the grammar rules does not seem to be enough to guarantee the correct and appropriate use of the language

(Lewis, 2000). This is because natural language is produced from the 'prefabricated chunks' or fixed terms that the language users have stored in their mental lexicons. Lewis (2002a) also further points out that many learners who know quite a lot of nouns fail to use them effectively. The reason for this is the lack of the awareness on the other words which co-occur with those nouns. Wray (2002) and Hoey (2007) also assert that native speakers of a language subconsciously keep a record of the context and co-context of the words they have encountered, so that they can use later at the right time and right occasions. This causes more trouble for non-native speakers in producing the language to be as naturally as their counterpart native speakers.

Collocation, which is a kind of prefabricated chunks, plays an important role in the production of natural language (Hoey, 2007). Thus, to be fluent in English, EFL/ESL learners should be aware of collocations and how to use them, especially for collocations in their professional field. Communicative competence is resulted directly from the mastery of prefabricated chunks or formulaic sequences (FSs) which collocation is one aspect of them (Henriksen, 2013). In order to master these aspects of the language, the ability to memorize and chunk them into units plays a crucial role. Consequently, there have been research and studies on both FSs and collocations in particular such as Lewis (2000), Schmitt (2004), and Wood (2010).

The term 'collocation', which is derived from the Latin word '*collocare*' means to place together or to assemble, was first introduced by the British contextualist, John R. Firth. He gives the meaning of the term as "collocations of a given word are statements of the habitual and customary places of that word" (1957, p. 181). According to Henriksen (2013), collocations are frequently recurring two-to-three word syntagmatic units. They can be both lexical and grammatical collocations.

Lexical collocations are such as the co-occurring of noun + noun (*air conditioner*), verb + noun (*pay tribute*), and adjective + noun (*hot spice*). Grammatical collocations are such as preposition + noun (*on guard*) and adjective + preposition (*immune to*).

According to Hill (2000), collocation is believed to be an important key to fluency. Collocation, at the same time, has been a huge barrier for EFL/ESL learners to achieve a native-like fluency. A number of studies on collocations have revealed that even high-level learners seem to face problems in using and developing second language (L2) collocational knowledge (Arnaud & Savignon, 1997; Farrokh, 2012; Nesselhauf, 2005). According to the hierarchy of mistakes created by McCretton and Rider[1], collocations are found to be the most frequent mistakes made by L2 learners (James, 1998; Miščin, 2013). Examples of collocational errors given by James (1998) are: 1) semantic error, e.g., '*crooked* stick' is replaced by '*crooked* year'; and 2) statistical weighed preference error, e.g., '*big* losses' is used instead of '*heavy* losses' (p. 152). According to Martin (1984), collocational mismatches are common in the language produced by L2 learners. This is because they hardly encounter a word or combination of words to enable them to determine its range or narrow the item down to its more fixed partnerships resulted from the inadequate exposure to the language. This difficulty particularly occurs with those semantic opaque combinations, when the combination of words leads to the change of the original meaning, or with specific field of the discourse such as engineering or nursing.

---

[1] Hierarchy of mistakes created by McCretton and Rider. (James, 1998, p. 229)

The Most Serious   The Least Serious

Lexis =>spelling =>negation =>word order =>prepositions => verb forms => concord

Although collocation has become a topic of study by scholars since the late 1950s, the study of collocations generally has not been extensively conducted. Most studies have been conducted on collocations in general English (Gledhill, 1996; Granger & Bestgen, 2014; Namvar, Mohd Nor, Ibrahim, & Mustafa, 2012; Wang & Good, 2007; and Webb, Newton, & Chang 2013). For collocations in scientific and professional fields, Mudraya (2006) and Ward (2009) have conducted the studies by looking at engineering textbooks. Menon and Mukundan (2012) explored collocations of high frequency noun keywords in prescribed science textbooks. Takač and Miščin (2013) looked at the collocational competence of non-native users of medical English. Miščin (2013) also investigated verb collocations in medical English.

From the above mentioned studies, it is evident that the study of collocations found in field of nursing, particularly in nursing research articles, is still rare. A nursing academic word list recently created by Yang (2015) is a study devoted to the field of nursing in terms of lexical approach. Since nursing is one of the professions that have been increasingly internationalized along with other health related professions, it is important for nursing students as well as nursing professionals to foster their competence in the use of English language before entering into their professional arena. One step to prepare them for reaching such goal, apart from other skills in English and beyond the knowledge of grammar and vocabulary, is the knowledge of collocations.

As the urge for regional integration is becoming reality with the free flow of skilled workers, preparing eligible professionals for the phenomenon is a wise response to act. For nursing professionals, facilitating them with collocations in their field could bring a great deal of benefits to the concerned parties. With collocations in

their professional field identified, it is believed that they would be able to communicate more effectively with their international colleagues as well as to progress in their career further by conducting research or further their study.

In addition, as it is difficult to judge how much knowledge of collocations nursing professionals have, a test of collocation knowledge has a role to play. Vasiljevic (2014) points out that in order to assess learners' needs and their lexical progress, a test of collocation knowledge is necessary. A test in itself not only measures the knowledge of the collocations learners have, but also can have a positive effect on them by increasing their awareness of collocations. Moreover, lessons on collocations extracted from the study would be designed and provided to the fourth year nursing students at Suranaree University of Technology (SUT). The effect of corpus-based instruction of collocations would then be assessed with the pre-test and the post-test.

Since knowledge of collocations is perceived as essential for effective and natural communication, it is necessary for professionals of all fields to have adequate knowledge of collocations in their fields. The application of corpus-based studies is a useful tool in identifying and studying collocations. A test of collocation knowledge is also a useful tool in assessing knowledge of collocations a person has. Therefore, the present study intended to apply both corpus-based study of collocations and a test of collocation knowledge to explore lexical collocations in a sample corpus of nursing research articles and the knowledge of collocation of the fourth year nursing students at SUT.

Nursing curriculum and nursing students at SUT

From the interview with an informant from the Institute of Nursing who was a senior academic staff, at the academic year 2016, there were 45 fourth year and 65 third year nursing students. In terms of English Courses required for the program, there were five English compulsory courses, which are English I to English V, taught by instructors from the School of Foreign Languages. The focus of these courses is to enable students to use English for communication in all four main skills. The courses also enable the teaching to assigned tasks that suit students' specific field of study.

There was one elective English course taught by the faculty members of the Institute of Nursing. The course was English for Nursing Profession. This elective English course has been designed to prepare nursing students for their nursing profession. The instructors were from the faculty members as they were familiar with the subject matters and working environment of nurses. Furthermore, for each nursing course, particularly in each specialized area of nursing, as textbooks used were usually in English, the first three hours of class was taught using English as a medium. This content-based instruction is believed to encourage students to use English and familiar with using English in their profession. The specialized areas currently being offered at the Institute of Nursing are: (1) Nursing care of families and midwifery, (2) Nursing care of the child and adolescent, (3) Adult and elderly nursing, (4) Community nursing, (5) Psychiatric nursing, and (6) Fundamental nursing.

In terms of academic and research articles, from the third year onwards, nursing students at the Institute of Nursing are required to read academic and research articles relevant to the courses of the specialized areas taken. This aims to encourage students to get familiar with reading academic and research articles in their field to

apply the knowledge gained in their professional practice. The students also knew how to access academic and research articles as they had to take a course in the program called 'Nursing Informatics' in their second year. The SUT library also arranged the workshops for new students on how to access resources available both on traditional paper-based texts and online materials.

Despite the fact that the nursing curriculum at SUT seems to provide necessary courses both in their specialized field and English courses, it is uncertain in terms of collocations used in nursing research articles either being recognized by the students or their knowledge of those collocations. Therefore, the present study intended to investigate into this issue and hope to gain useful information and helpful means in response to it.

1.2 Rationale of the study

It has been suggested by Schmitt (2010) and Milton (2013) that knowing the vocabulary is a crucial component for EFL/ESL learners to achieve certain level of fluency in the language being learned. Having enough of the vocabulary repertoire, learners can predict their other language skills namely reading, writing, speaking, and listening. Thus, to be able to understand the language and produce the language for effective communication, learners need to acquire enough vocabulary. This traditional view is reflected in Lewis' statement that "…grammar is creative, while words are like building bricks, fixed packages of meaning" (2002a, p. 37). This statement of Lewis shows how important vocabulary is in knowing a language. Without enough vocabulary, learners will face difficulty in constructing sentences. This is similar to not having enough bricks which could prevent a building from being built.

Though vocabulary is seen as an essential element in knowing a language, vocabulary alone or even with the grammatical rules is not enough to enable learners to produce natural language or native-like expressions (Nation, 2001). Rather, collocation, which is closely related to vocabulary in the form of formulaic language or chunks, plays an important role in effective and natural communication. Vocabulary knowledge, according to Shokouhi and Mirsalari (2010) and Gaballa and Al-Khayri (2014), involves knowing more than just the meanings or basic meanings of words in isolation, but also the words that tend to co-occur with them. According to Nation (2001), collocations play a very important role in knowing a language with three aspects. Firstly, language knowledge is collocational knowledge. This is because the stored sequences of words are the bases of learning, knowledge and use. Secondly, all fluent and appropriate language use requires collocational knowledge. To produce a native-like language, collocations play the essential role in it. Finally, many words are used in a limited set of collocations and knowing these is a part of what is involved in knowing the words.

Howarth (1998 as cited in Henriksen, 2013) claims that the use of formulaic language as the terms 'formulaic sequences' (FSs), which include collocations, is very genre-specific. Mastery of collocations may be reflected on clarity, precision and lack of ambiguity in language production. As collocations act as central composite syntactic units for clause level production, lack of collocational knowledge may be expected to have a negative effect on L2 performance both productively and receptively. This could lead to the misunderstanding of the message. Lack of collocational competence may also indicate an inadequate academic expertise.

Millar's (2011) study also found that incorrect use of collocations can reduce processing speed even for native speakers.

In professional practice, effective and appropriate use of language is indispensible. As it takes time to master a second language such as English, university students of all professional fields are required to take English courses as part of their curricula. This is also the case in Thai universities. Apart from general English courses, there are also English for Academic Purposes (EAP) and English for Specific Purposes (ESP) courses. In the field of nursing, there is no exception. For nurses, their effective and natural use of English among themselves, with doctors, and with patients is very important. Thus, identifying and classifying collocations found in research articles in the field should make it useful for lexical learning and teaching in the area of nursing collocations. This should also be easier for teachers and students to be explicitly aware of collocations in their field which will certainly benefit them for producing the language in all four skills more effectively and naturally. This can be particularly important for nursing students and professional nurses as they need to read a number of research articles in their specialized areas to apply in their real practice. Additionally, lessons designed based on the collocations gained from the study to teach the fourth year nursing students at SUT should be a great deal of benefit. A pre-test and a post-test administered to compare the results should also reveal the knowledge of the students and the implication for teaching.

1.3 Purposes of the Study

It is apparent that it is inevitable for professional nurses to be efficient in English for both their professional advancement and further study. However, there

seem to be obstacles preventing them from achieving their goals. These obstacles are mainly the inadequate knowledge and lack of awareness of collocations. Besides, there has not been a study of collocations in the field of nursing, particularly from a corpus of nursing research articles. Thus the present study attempted to fill up these gaps by building a Sample Corpus of Nursing Research Article (SCNRA) from which the keyword list and the collocations can be generated. As well, a collocation test based on the collocations generated from the study was built and administered with the fourth year nursing students at SUT to evaluate their collocational knowledge. The lessons on nursing collocations have been provided with the test afterwards to measure their progress.

The aims of the present study were as follows:

1) To identify and classify keywords found in the Sample Corpus of Nursing Research Articles (SCNRA) published in international journals in the field of nursing,

2) To explore lexical collocations found in the SCNRA using keywords found as 'nodes' to find their 'collocates', and to classify collocations found according to their combinations,

3) To assess collocation knowledge of the fourth year nursing students at SUT based on the collocations found from the SCNRA, and

4) To provide lessons on nursing collocations and assess the effect of corpus-based instruction.

1.4 Research Questions

The research questions of the present study are the followings:

1) What are the keywords in the SCNRA based on the frequency of occurrence at ≥ 50 and the keyness value at ≥ 20? What is the proportion according to their parts of speech?

2) What are the lexical collocations of the keywords in the SCNRA? What is the proportion according to each type of combinations?

3) How much collocational knowledge do the fourth year nursing students at Suranaree University of Technology (SUT) have based on a test of collocations extracted from the SCNRA?

4) How much does corpus-based instruction help improve the knowledge of collocations for the fourth year nursing students at SUT?

1.5 Significance of the Study

As collocational knowledge plays an essential part in effective and natural production of language (Nation, 2001), extracting and classifying keywords and collocations from the SCNRA should be beneficial in a number of aspects.

In terms of EFL/ESL learning and teaching, the present study could contribute to the field as follows:

1) With the list of keywords and collocations gathered directly from the SCNRA, it should be more convenient for nursing students as well as nursing professionals to learn and acquire these keywords and lexical collocations.

2) The keyword list and collocations generated from the study should be particularly useful for nursing students and professional nurses who are conducting research in the field as the collocations could help improve their reading and writing skills.

3) It should be less time consuming for the language instructors who teach English for Specific Purposes (ESP) for nursing students to point out to their students if corpus-based instruction of collocations in their field is beneficial.

4) These collocations should also benefit any EFL/ESL learners who are interested in collocations in the field of nursing and learning of English in general.

In terms of corpus-based study, the present study generates a list of specific keywords and collocations, specifically lexical collocations, to the field. This should add new resource and body of knowledge to the field.

A test of collocational knowledge constructed based on lexical collocations identified from the study should also be valuable to the EFL/ESL teaching and learning as follows:

1) The test should reveal how much knowledge nursing students have on collocations found in the SCNRA.

2) This may have the pedagogical implications whether collocations found in the SCNRA should be taught explicitly or not based on the test results.

3) The collocation test created may be a useful tool for assessing the knowledge of collocations in the field of nursing in general.

Lastly, lessons on collocations found in the SCNRA could be directly beneficial to the teaching and learning of collocations in the field.

1.6 Scope and Limitations of the Study

The scope and limitations of the present study were as follows:

1) It aimed to explore keywords and collocations found in selected research articles in the field of nursing.

2) The research articles used to compile into a sample corpus were those selected from 10 international nursing journals which were available online and accessible via SUT's library resources. These 10 journals had been selected based on the fields which were offered at Institute of Nursing at SUT. Therefore, not all of the fields were included.

3) Thirty research articles selected from each selected journals were based on the latest issues as they should represent the up-to-date collocations in the field accessible, which make up the total of 300 research articles to form the SCNRA.

4) As the focus of the present study was on lexical collocations with three-word span on the right side of the nodes, the collocations to be investigated can be those with the word span from two to four words.

5) The test of collocations constructed based on lexical collocations produced from the study aimed to be administered with the fourth year undergraduate students of nursing at SUT to determine the level of their collocational knowledge in the field. This group of students was the target for the test as they were among the target users of the journals and in their final year which would soon enter their professional lives or further their study.

1.7 Definitions of Key Terms

Collocation, in this study, refers to the co-occurrence of words which is predictable as they have tendency to occur together and are naturally co-selected by native speakers.

Lexical collocation, in this study, refers the co-occurrence of two content words within two to four word span found in the SCNRA. The co-occurrence tends to

occur naturally with statistical significance in terms of their association. The co-occurring will be examined based on the nodes which are the content word combinations (*Verb + noun*, *Adj. + noun*, *Noun + verb*, *Noun + noun*, *Adv. + adj.*, and *Verb + adv.*) adapted from the seven types provided by Benson et al. (2010).

Nursing Journals refers to international journals in the field of nursing accessible via Suranaree University of Technology's (SUT) library resources and research articles are taken from the selected journals for the present study.

Nursing research articles refer to research articles taken from 10 purposively selected journals in the field of nursing accessible via SUT's library resources. These journals closely match with the specialized areas of nursing offered at SUT.

Corpus (plural: *corpora*) refers to a collection of texts either written or spoken of a particular genre or variety of genres for a particular purpose of analysis. A corpus is stored in an electronic form to be readable by a computer program. In the present study, a corpus refers to a collection of the selected 300 research articles in the field of nursing which is stored in an electronic form known as the SCNRA to be analyzed by a concordance program for lexical collocations.

Sample Corpus of Nursing Research Articles (SCNRA) refers to the sample corpus built for the present study. The SCNRA comprises of 300 research articles from the selected 10 journals of nursing, 30 latest articles from each.

Collocational knowledge refers to knowledge of collocations a person has which often leads to and affects collocational competence and communicative competence.

Test of Collocation Knowledge refers to a set of test constructed based on lexical collocations identified from the SCNRA created for the present study. It is

meant to evaluate the knowledge of the fourth year nursing students on lexical collocations found in the SCNRA. This test will be applied twice with the same target group of students as a pre-test and a post-test.

1.8 Summary of the Chapter

This chapter presented background of the present study which leads to the rationale, purposes of the study and the research questions. The significance of the study has also been provided. The scope and limitations of the present study has been described. The definitions of key terms in the present study were given. The following chapter provides reviews on related literature concerning the present study.
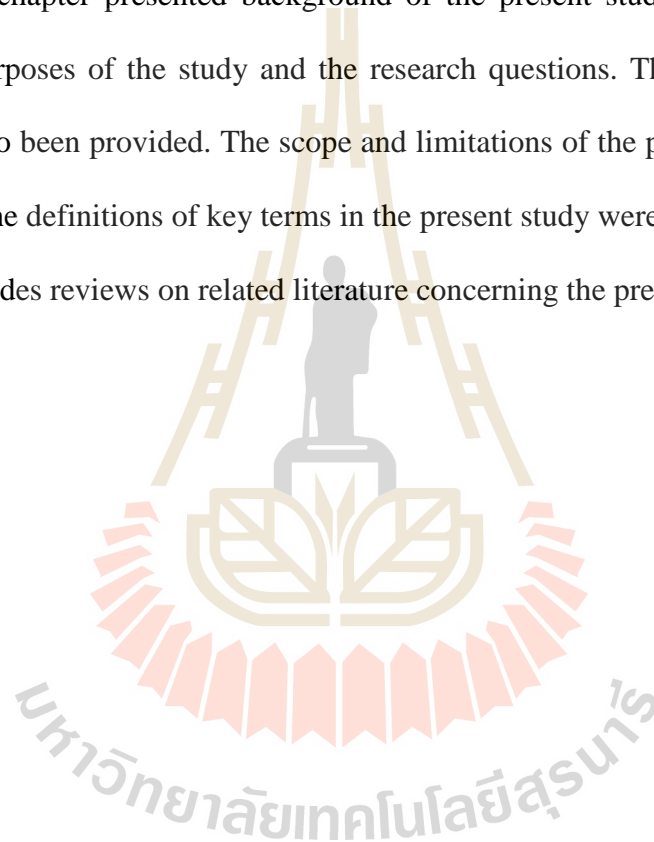
CHAPTER 2

LITERATURE REVIEW

This chapter provides reviews of the literature related to the present study. Before touching on collocations, the chapter begins with exploring the relationship between vocabulary and collocations in second language acquisition (SLA) as well as the learning and teaching vocabulary. The information on lexis and lexical approach, formulaic language, and mental lexicon are included in the first part. The second part of the chapter devotes to collocations which covers the definitions, the classifications, the types, and the criteria for identification. The importance of collocation as well as the teaching and tests of collocation knowledge are also provided. The third part provides information on corpus studies, which includes the definitions, the development, types, and benefits. The information on concordance software, corpus-based lexical analysis as well as corpus-based language teaching and learning are also illustrated. The fourth part gives information on English for Specific Purposes (ESP) in relation to collocations teaching and learning as well as corpus-based instruction. The fifth part introduces pre-experimental research design. The sixth part gives review concerning journals and journal articles as well as nursing journal and journal articles. The last part gives the theoretical framework of the present study.

2.1 Vocabulary and SLA

This part presents the roles vocabulary play in the acquisition of a second language. The views of scholars on how vocabulary and collocation are related and the learning and teaching of vocabulary in connection with SLA are under reviewed. Lexical approach and lexical items, formulaic language, and mental lexicon also are under reviewed as they relate closely to vocabulary as its larger units.

2.1.1 Vocabulary and Collocation

Vocabulary has been categorized into groups by some scholars. Widdowson (1983) classifies words into two groups: schematically bound and words of high indexical or procedural. The schematically bound words can be identified by their use in that they narrow the frames of reference and identify particular fields. They are words that can be used in specific contexts. Indexical or procedural words, on the other hand, can be found in wider range of contexts. They can be used to define and locate within their fields. The example can be seen in the words '*hydrometer*' and '*instrument*'. The former can be found in a more specific area of texts, while the latter can be found in wider circumstances as well as can be used to define the former. Thus '*hydrometer*' is classified as a schematically bound word, while '*instrument*' is a high indexical word.

McCarthy (1990) classifies vocabulary into two distinct groups: core vocabulary and procedural vocabulary. Core vocabulary refers to those words with more influential in terms of meanings and uses. The core vocabulary tends to be usable in various contexts with different word-classes. Procedural vocabulary refers to those that are used to talk about other words, to paraphrase, to define, and to organize them in communication. Comparing the words '*fat*' and '*obese*', for example, are

similar in meaning. However, 'fat' can be found in more places than '*obese*' such as '*fat man*', '*fat chance*', '*fat wallet*'. In this case, '*fat*' has a quality as a core vocabulary, while '*obese*' is a procedural vocabulary.

Carter (1998), meanwhile, categorizes words into two groups: grammatical and lexical words. Grammatical words are those known as 'functional words', functors', or 'empty words'. They do not represent meanings. These grammatical words include pronouns, articles, auxiliary verbs, prepositions, and conjunctions. Lexical words, on the other hand, are those known as 'full words' or 'content words'. They carry meanings or information contents. This group of words includes the nouns, adjectives, verbs, and adverbs. Lexical items are groups of words that go together such as multi-word verbs, phrasal verbs, and idioms. They are sometimes called "vocabulary items" or simply "items". The term is useful and fairly natural hold-all term which captures and helps to overcome instabilities in the term 'word'.

In second language acquisition (SLA), vocabulary repertoire of the target language plays an important part in achieving certain level of fluency. Godwin-Jones (2010) states that "an essential element of language learning is building one's personal store of words and expressions" (p. 4). This will then lead to the competence in all other areas of communication. Tight (2010) also claims that "lexical development is an essential step in second language acquisition" (p. 792). Vocabulary is so important in the acquisition of second language that "no amount of grammatical or other type of linguistic knowledge can be employed in communication or discourse without the mediation of vocabulary" (Richards, 2000, p. xi, as cited in Tight, 2010).

However, Schmitt (2010) points out that the vocabulary is important for language use with the larger is the better. Formulaic language is as well important as

individual words. Schmitt's view is in agreement with a statement made by Nation (2001) that knowing vocabulary involves knowing the other words they typically co-occur with. For example, '*fast food*' cannot be replaced by '*speedy food*' or '*quick food*'. Collocation (Nation, 2001) is the only one relationship that relates to the appropriate interpretation and productive use of vocabulary. Thus, vocabulary and its extension play an important role in the acquisition of a language.

According to Takač (2008), '*lexeme*' or '*lexical unit*' covers a number of word types namely inflections, polysemy, and multi-word items. This leads to an agreement among SLA linguists and theorists that vocabulary is made up of a variety of forms such as morphemes, compounds, idioms and other fixed expressions. Lexical items can hardly be viewed in isolation since their meaning can be studied through componential analysis emerged from the relationship of the combined words, which is called a syntagmatic relationship. This relationship is characterized by restricted linear sequencing of lexemes. These restrictions determine acceptable lexical units to co-occur as finally known as collocations.

From the above mentioned views on vocabulary and its relation to the acquisition of a second language, it is clear that the unit larger than single words is important. Thus, learning and teaching of vocabulary should extend further to formulaic language in which collocation is included as one of its kind. Therefore, in order to develop learners' communicative competence, 'collocation' is one of the most effective techniques (Yoshida, 2013).

2.1.2 Teaching and Learning of Vocabulary

According to Schmitt (2000), vocabulary is not simply learnt receptively at the beginning and productive knowledge can occur automatically later. This is because

words are complex by nature. Nation (1990 as cited in Schmitt, 2010) has listed eight different aspects of knowledge a person must master in order to know a word as follows: its meaning(s), its written form, its spoken form, its grammatical patterns, its collocations, its register, its associations, and its frequency. The different aspects of word knowledge are not necessarily learnt at the same time. Each aspect of word knowledge is likely to be learnt in a gradual manner. Therefore, the acquisition of vocabulary is 'incremental' since it is impossible to learn all of the aspects at the same time.

Nation (1990) points out that there are two approaches in vocabulary learning: direct and indirect. The former occurs when learners do exercises and activities that intend to help them learn particular words. The latter approach is applied when the attention of the learners is not directly on the vocabulary, but more on a broader context such as understanding the passage of the reading instead of some particular words.

The vocabulary teaching approach suggested by Nation (1990), thus, is to apply the two approaches in four ways from the most indirect to the most direct as follows: 1) Use prepared materials that are simplified to suit learners' level; 2) Give attention to the unknown words by giving appropriate explanation; 3) Teach vocabulary in relation to other language activities or recycle the vocabulary in other activities; and 4) Spend time both in class and outside class on learning vocabulary with activities that are not directly linked to language activities.

Schmitt (2000) also states in regard to the teaching and learning of vocabulary that there needs to be the combinations of both explicit teaching and activities that will allow incidental learning to occur. It is also suggested that there is the

relationship between the learners' level of proficiency and the teaching approaches whether it should be explicit or indirect. In this case, the low level learners need more explicit teaching of vocabulary, while the higher level may need less. Both explicit and incidental approaches are important in vocabulary acquisition.

Explicit approach to vocabulary acquisition

Schmitt (2000) suggests three methods of explicit approach to vocabulary acquisition as follows:

1) Integrating new words with old (Sökmen, 1997) is one of the explicit approaches in helping learners to acquire the language. It can be done by grouping similar words together.

2) Teaching the underlying meaning concept of a word. This is because many words are polysemous and some of their different meaning senses have a common underlying trait. By defining the underlying meaning concept of the words, learners can understand the words in a wider variety of contexts.

3) Teaching word families instead of individual word forms. This can be done by mentioning other members of a word family when introducing a new word. Learners may be asked to guess a new word's derivatives as well as including a derivation section as part of assessment.

Incidental learning approach to vocabulary acquisition

Getting maximum exposure to the target language is the most important element for incidental learning of a second language. The best way to do this is joining students' exchange program to spend time in the target language environment (Milton & Meara, 1995; Schmitt, 2000). However, most L2 learners do not have such opportunity to be exposed to the L2 environment by means of spending time at the

place where the L2 is used in daily life. Thus, ways to help learners to get exposed to the target language at their own home country is, according to Schmitt (2000), to read more. By reading authentic materials, learners can greatly increase their vocabulary repertoire.

From the views on the teaching and learning vocabulary above, it is agreeable that there can be either direct/explicit or indirect/incidental approaches to it. One interesting fact found from above is that collocation is included as a part of vocabulary teaching. The empirical evidence is shown in studies by Balcı and Çakır (2012) and Rahimi and Momeni (2012). Both studies revealed that teaching vocabulary through collocations results in a better learning of the words than using traditional techniques. This method of teaching also increases retention of new vocabulary items. Although it is unclear how collocational knowledge is acquired, it is quite certain that it is not easy to achieve. This aspect of collocation, therefore, still clearly distinguishes native speakers from nonnative speakers (Wray, 2002; Hoey, 2007).

2.1.3 Lexical Approach and Lexical Items

Lexical approach emerges with the publication of 'The Lexical Approach' by Lewis (1993). The approach derives from the communicative approach to language learning. The emphasis of the approach is on acquiring extensive words and their combinations. Lewis claims that words do not exist in isolation in that not any words can be placed in any parts of a sentence even though it is grammatically correct. For this point, Lewis claims that "language is grammaticalized lexis, not lexicalized grammar" (1993, p. vi). This means that lexis is more important than grammar in creating meaning. Lexical approach distinguishes between vocabulary and lexis by

which vocabulary refers to a single unit of words whereas lexis refers to word combinations that learners store in their mental lexicon. Lexical approach puts special attention directly to collocations and expressions that include institutionalized utterances and sentence frames and heads. On this point, Lewis states that we deliberatively think of collocations, and to present them in our expressions instead of individual words. This means rather than trying to break things into smaller pieces, we have a conscious effort to see things in larger and more holistic ways (Lewis, 1997).

Research studies by Nattinger and DeCarrico (1992) and Sinclair (1991) suggest that there is more lexical patterning of collocation in language than previously expected. It is found that words act less as individual units and more as part of lexical phrases. This phenomenon reflects how the mind is likely to "chunk" language in order to make it easier to process. Warren (2005) also claims that collocations should be viewed as multiword lexical items with form-meaning pairing in specific contexts that fulfill communicative functions. Gyllstad (2013) states in response to this point that if lexical items such as collocations are accepted as a part of everyone's vocabulary, then we need to start thinking of ways of incorporating lexical items larger than single words into measures of vocabulary size.

According to Jackson and Zé Amvela (2007), '*lexis*' is originated from Greek means '*word*'. This leads to the word '*lexicology*' to mean the study of lexis, which refers to the stock of words in a given language. The words 'vocabulary', 'lexis', and lexicon' may be considered as synonyms. However, each of them conveys slightly different sense. '*Vocabulary*' is more colloquial. '*Lexicon*' is more academic and technical. '*Lexis*' is in between the other two. '*Lexical item*' or '*lexeme*', according to

Crystal (1995), is a "unit of lexical meaning, which exists regardless of any inflectional endings it may have or the number of words it may contain" (p. 74). It is a basic unit of meaning. Their examples can be seen on the dictionary as the headwords in a dictionary are all lexemes. Sinclair, Jones, and Daley (2004) also define a lexical item as a language unit that indicates a specific area of meaning which is unique in terms of co-occurring patterns. A lexeme or lexical item may consist of one or more words. According to Lewis (2002b), lexical items are arbitrary which means they are simply the consensus of what has been institutionalized, the use of the language that has been agreed upon by the group of users, chosen from what could be used, and actual use as opposed to theoretically possible language. Thus, each community group is likely to develop its own way of communication to be understood particularly among the members of the community. For example, people from another part of a country may not understand when they hear people from the other parts speaking.

Hanks (2013) explains that lexis or word may denote any of the following six concepts: (1) a *type* which refers to a unique spelling form; (2) a *token* which is a single occurrence of a lexical type; (3) a *lemma* or *lexeme* which refers to all forms of each word; (4) a *phraseme* or *multiword expression* which has particular meaning; (5) a lexical entry which include lexemes, phrasemes and some partial items; and (6) any of items 1 to 3 including or excluding of proper names. For example, in a sentence below

He came he saw he kicked a ball, I come I see he kicks the bucket.

There are 16 tokens, which are the total number of words in the sentence. There are 12 types, which represent different form of words. There are 9 lemmas, which are root form of each word. There is one phraseme or multiword expression, kick the bucket.

Hank's view on lexis is in agreement with Hill's (2000) which claims that 'a central feature of lexis is collocation" (p. 47).

To this point, another evidence of relationship between vocabulary and collocation is represented. This is because collocation is closely related to '*lexical item*' as well as '*phraseme'* or '*multiword expression'*. In that collocation is one kind of lexical item as well as of phraseme or multiword expression. For example, from the sentence above, the word '*kick'* tends to co-occur with the words '*ball'* and '*bucket'*. Thus collocation has been an integral part of a language all along, only given less attention than that of single unit of vocabulary.

2.1.4 Formulaic Language

Formulaic language has long been recognized mainly in terms of idioms because they have easily noticeable forms and meanings which are not represented by the combinations (Schmitt, 2010). As more research has been conducted on formulaic language, it is found to be the core characteristic of language. Idioms are found in all kinds of texts including children's books. The frequency can be higher in genre specific corpora such as meetings, TV shows, magazines, and news. Apart from idioms, there are many other types of formulaic language. The difference among them is dependent on degree of fixedness, institutionalism/conventionality, and opacity/non-compositionality. The lack of uniformity leads to different terminologies used other than formulaic language. Wray (2002) uses the term '*formulaic sequences'* and found over 50 terms for it such as chunks, formulaic speech, multi-word units, collocations, formulas, prefabricated routines, conventionalized forms, holophrases, and ready-made utterances. According to Henriksen (2013), collocations are "a subset of formulaic sequences (FSs)" (p.29). Many scholars in the field namely Barfield and

Gyllstad (2009); Nation (2001); Schmitt (2004); Wood (2010); Wray, (2002) claim that FSs are the central of communicative competence.

Formulaic language, according to Kuiper (1996), has two underlying properties, which are: 1) the units of formulaic language are not only any sequence of words, but phrases; and 2) they are lexical items exactly like other lexical items such as words, and with the same properties as words would have if they are phrases. The term '*lexical phrases*' is used by Nattinger and DeCarrico (1992) when examining the relationship of formulaic language and their functional usage. When exploring the relationships between two-word pairs, the term '*collocations*' is used. The terms '*prefabricated expressions*' and '*chunks*' are used when focusing on the holistic storage of the forms.

From the above views on formulaic language, it is revealed that collocation is closely related to this type of language as being a part of it. This is especially close with the use as described by Nattinger and DeCarrico (1992) in relation to the present study as it intended to look at formulaic language of two-word pair, which they specify as collocation. Besides, in the linguistics literature, the recurrent combinations of lexical items are often referred to as collocations or formulaic speech (Bonk, 2000).

2.1.5 Mental Lexicon

The term *mental lexicon* or *mental dictionary*, according to Takač (2008), is only a reminiscent of a traditional printed dictionary because it refers to a collection of lexical items. The difference is that a printed dictionary is static, limited, and likely to become outdated, while the mental lexicon is more complex and has more efficient organization. Words in the mind seem to be connected into semantic networks with the strongest links involving coordination and collocation. In Krashen's (1987) term,

mental lexicon is mostly acquired rather than formally learned. Our mental lexicon is larger than previously recognized and very few of lexical items we know were consciously learned (Lewis, 2002b). Mental lexicon, according to Hulstijn (2000), is "a memory system in which a vast number of words, accumulated in the course of time, has been stored" (p. 210).

According to McCarthy (1990), the mind plays a role in language development and general linguistic behavior through the three processes of input, storage, and retrieval. In terms of input, children learn their mother tongue through listening produced by people around them before starting to produce the language. For second language learners, they usually deal with spoken and written form of the language at the same time from the beginning. Through listening, learners' minds familiarize themselves with various aspects of the language. In terms of storage, native speakers are likely to store words in their minds according to the spelling patterns as well as the organization into semantically related families. In terms of retrieval, this involves with receptive knowledge and productive knowledge. Receptive retrieval is related to how the mind manages input to match with the right sound and orthographic patterns and their associated meanings. Productive retrieval, meanwhile, has to be in given forms that match their meanings. These forms can be simple words, derived words and compounds, fixed collocations, and other multi-word units. The retrieval process is believed to be useful for both native-speakers and second language (L2) learners in producing the language with less effort and time.

Skehan (1998) agrees with the processes of language learning and producing provided by McCarthy above. However, he gives emphasis to the importance of noticing. The processes suggested are shown in Figure 2.1 below.
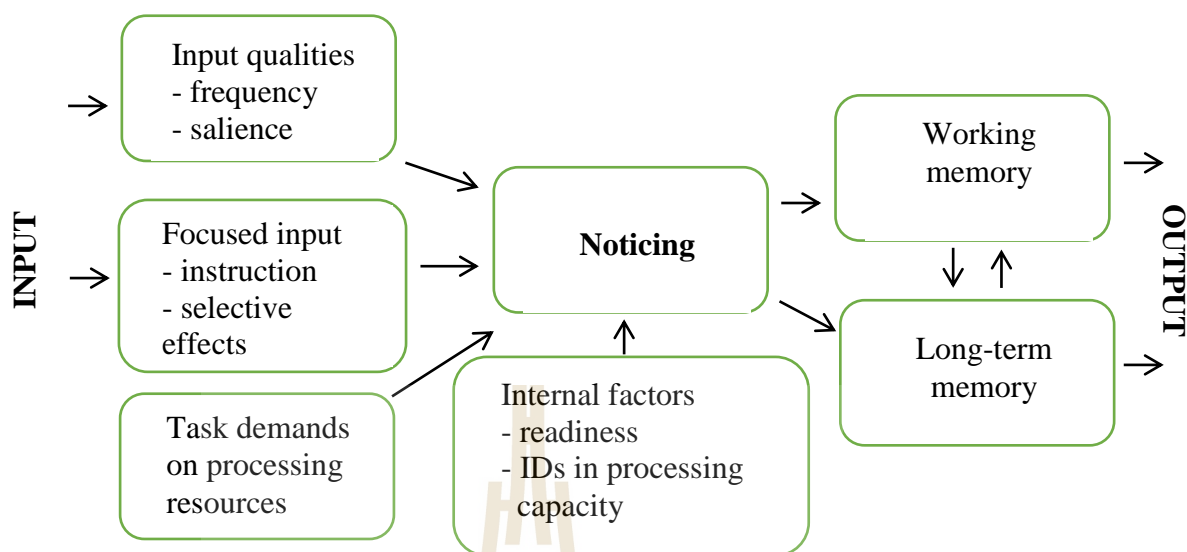
Figure 2.1 Skehan's (1998) language learning processes with emphasis on noticing

It is evident that mental lexicon has a role to play in effective production of a language with less effort. Therefore, raising L2 learners' awareness through the notion of noticing as well as increasing the size of their storage of collocations should bring a positive outcome to their proficiency (Farrokh, 2012). One way of doing this is by applying corpus tools to extract collocations out of a sample corpus of a particular genre which the present study has attempted to do.

2.2 Collocations

This part focuses on collocations which is the important element of the present study. The definitions given by a number of scholars in the field are explored. Then, the review on classifications and types of collocations is provided. The identification of collocations and the importance of collocations are presented as well as the teaching of collocations. The last part depicts the information on the test of collocation knowledge.

2.2.1 Definitions of Collocation

There are various definitions given by various scholars on collocations. Firth (1957) defines collocation as 'the company words keep' − their relationships with other words, and 'the way words combine in predictable ways'. Hill (2000) also defines collocation as 'a predictable combination of words'. If words collocate, they are co-selected by the speaker or writer and they are not a chance co-occurrence (Cheng, 2012). Collocations are sets of two or more words which appear together more frequently than their individual frequencies would lead us to expect (Hoey, 1991; Jones & Sinclair, 1974). Nattinger and DeCarrico (1992) identify collocations as unmarked choices of expression 'co-occurring lexical items that have not been assigned particular pragmatic functions by pragmatic competence'.

Collocation, according to Carter (1998), is "a term used to describe a group of words which occur repeatedly in a language" (p. 51). Collocation can be either grammatical or lexical. Schmitt (2000) gives the definition of collocation as 'the tendency of two or more words to co-occur in discourse' (p. 76). Schmitt (2000) also adds that collocation is one of the eight types of word knowledge introduced by Nation. Later Schmitt (2010) further gives more reflection on collocation that works on collocations mainly look at the relationships between two-word pairs. However, he suggests two key elements to the notion of collocation: words co-occur together and the varying degrees of exclusivity (Schmitt, 2000). The example he gives is the word '*blonde*' which can go exclusively with the word '*hair*', but not with the words like '*paint*' or '*wallpaper*'.

Nation (2001) defines the term 'collocation' as a group of words that belong together, either because they commonly occur together or because the meaning of the

group is not obvious from the meaning of the parts. There are two criteria for identifying a collocation: words frequently co-occur and some degree of semantic unpredictability as a result of the co-occurring. Lewis (2002b) also gives the definition to 'collocations' as "those combinations of words which occur naturally with greater than random frequency" (p. 25). He also adds that collocations co-occur, but not all words which co-occur are collocations. Collocation is about words which co-occur, not ideas or concepts.

Similar to Nation's, Paltridge (2006) defines collocation as associations between vocabulary items which have a tendency to co-occur such as the combinations of adjectives and nouns, verbs and nouns and others. These combinations have expectancy relations in that they tend to be predictable for particular pairs. Jackson and Zé Amvela (2007) define collocation as "a structural or syntagmatic relation, to meaning relations that a word contracts with other words occurring in the same sentence or text" (p. 131). Collocation is closely related to meaning arising from predictable co-occurrence. According to Bennett (2010), collocation is "the statistical tendency of words to co-occur" (p. 8). She explains that when one word is used, there is a high statistical probability that a certain word or words will occur alongside of it.

Apart from definitions given above, Gledhill (2000) defines the term 'collocation' into three ways according to three different perspectives: 1) Halliday's statistical/textual view; 2) the semantic/syntactic tradition in lexicology, and 3) the discoursal/rhetorical model from discourse analysis. The details of each perspective are presented below.

In terms of **the statistical/textual perspective**, collocations have often been defined statistically in corpus-based studies, following M. A. K. Halliday, who frames collocation in terms of statistical probability and co-occurrence. The core focus of this perspective is that it is "the syntagmatic association of lexical items, quantifiable, textually, as the probability that there will occur at *n* removes (a distance of *n* lexical items) from an item *x*, the items a, b, c ..." (Halliday, 1961, p. 276). Therefore, any given item enters into a range of collocation will be ranged from more to less probable.

Gledhill (2000) further explains that a collocate can simply be seen as any word which co-occurs within an arbitrarily determined distance or *span* of a central word or *node*. Collocation is thus considered to be the frequency with which collocates co-occur with one node relative to their frequency of collocation with other nodes. From the point of view of many corpus linguists, all that separates collocation from mere word co-occurrence is the statistical level at which the researcher is happy to say that the co-occurrence is not accidental. This approach is also 'textual' in that it relies solely on the ability of the computer program to analyze large amounts of computer-readable texts.

Gledhill (2000) concludes that this perspective essentially emphasizes collocation as co-occurrence and recurrence of words in a language. The notion of statistical collocation is fundamental to Halliday's theory of discourse. The textual view of collocation also emphasizes the fact that collocations are the result of reformulations and paraphrases which have developed throughout the length of a text. A textual collocation is likely to have a specific textual function or may occur in a rather restricted set of contexts.

In terms of **the semantic/syntactic tradition**, collocation is defined as a more abstract relationship between words, without reference to frequency of occurrence or probability, shifting the emphasis therefore from the textual co-occurrence of an expression to its potential for lexical combinability. The standard definition of collocations in this perspective is given by Benson (1989, p. 85) as "fixed recurrent combinations of words in which each word basically retains its meaning." Mel'čuk (1995), however, defines collocation as "a semantic function operating between two or more words in which one of the words keeps its 'normal' meaning" (p.182). Van der Wouden (1997) claims that idioms and collocations share some properties such as how they are formed and their role in a language. What divides the two, however, is the meaning after the combination. In that the meaning is completely changed is an idiom, while the meaning of individual words remains or is not totally changed is a collocation.

In terms of **the discoursal/rhetorical perspective**, some scholars (Moon, 1987; Fernando, 1996; Fillmore & Atkins 1994; Kay & Fillmore, 1999) believe that collocation can also be determined based on its performance in the language use. For distinguishing collocations from idioms, the notion of marked and unmarked is applied. For example, '*to get sacked*' and '*to be fired*' are more marked than '*to lose one's job*'. Thus, according to this perspective, collocation is those combinations with less marked in comparison to that of idioms.

Although the concept of collocation is very diverse, Gledhill (2000) summarizes that all of the approaches converge on an important and recognizable phenomenon, the 'familiar recurrent expression'. Instead of arguing the case for one specific viewpoint, he sees each as compatible and relevant at different points. Since

the main purpose of his book is to analyze a large corpus of texts, the 'statistical/textual' perspective is the most appropriate approach to be adopted in the first stages of corpus analysis. For the present study in which the combination of two lexical word pairs were the target to be examined based on their mutual information (MI) score with consideration on the frequency of occurrence, the most appropriate approach is the combination of the statistical/textual perspective and semantic/syntactic tradition.

In summary, the definitions of collocation given by scholars are presented in Table 2.1 below.

Table 2.1 Definitions of collocation given by scholars

Scholars	Collocation defined
J. R. Firth (1957)	the company words keep – their relationships with other words and the way words combine in predictable ways
Hoey (1991); Jones & Sinclair (1974)	sets of two or more words which appear together more frequently than their individual frequencies would lead us to expect
Nattinger & DeCarrico (1992)	unmarked choices of expression 'co-occurring lexical items that have not been assigned particular pragmatic functions by pragmatic competence
Carter (1998)	a group of words which occur repeatedly in a language and can be either grammatical or lexical
Schmitt (2000)	the tendency of two or more words to co-occur in discourse; words co-occur together with the varying degrees of exclusivity
Hill (2000)	a predictable combination of words
Nation (2001)	a group of words that belong together, either because they commonly occur together or because the meaning of the group is not obvious from the meaning of the parts
Lewis (2002b)	those combinations of words which occur naturally with greater than random frequency; collocations co-occur, but not all words which co-occur are collocations

Paltridge (2006)	associations between vocabulary items which have a tendency to co-occur such as the combinations of adjectives and nouns, verbs and nouns and others; These combinations have expectancy relations in that they tend to be predictable for particular pairs
Jackson & Zé Amvela (2007)	a structural or syntagmatic relation, to meaning relations that a word contracts with other words occurring in the same sentence or
Bennett (2010)	the statistical tendency of words to co-occur
Cheng (2012)	If words collocate, they are co-selected by the speaker or writer and they are not a chance co-occurrence

From Table 2.1 above showing definitions of collocation given by scholars, there are some similarities and differences among them. The similarities are mainly the combinations of two or more words. The varieties among them are those additional details given such as the co-occurrence happens repeatedly, naturally, predictably, structurally or syntagmatically, statistically, and selectively. However, it may be concluded that collocation is the co-occurrence of words which is predictable as they have tendency to occur together and are naturally co-selected by native speakers. Lexical collocation in the present study then refers to the co-occurrence of two lexical words found in the SCNRA which tend to occur naturally with statistical significance in terms of their association.

2.2.2 Classifications of Collocation

Carter (1998) categorized collocation according to the lexical patterns into three groups: collocational restriction, syntactic structure, and semantic opacity. The details of each group are shown as follows:

1) *Collocational restriction*

This group consists of four sub-groups: unrestricted collocation, semi-restricted collocation, familiar collocation, and restricted collocation.

1.1) Unrestricted collocation: this refers to the capacity of particular lexical items to co-occur with a wide range of items. Most core vocabulary belongs to this category. For example, the adjective '*bright*' can co-occur with a number of other words such as *light*, *color*, *sun*, *future*, and *smile*. The verb '*run*' can co-occur with *business*, *errands*, *shop*, *risk*, and *late*.

1.2) Semi-restricted collocation: this group refers to lexical patterns in which a number of items which can be substituted in different syntactic slots are more determined. For example, the noun '*harbor*' can co-occur with *doubt*, *grudges*, *uncertainty*, and *suspicion*.

1.3) Familiar collocation: this group refers to the combinations between words which keep regular company with each other. For example, *unrequited love*, *unmitigated disaster*, *readily admit*, and *lukewarm reception*.

1.4) Restricted collocation: this group refers to the combinations between words are generally more fixed and closed. For example, *stark naked* and *pitch black*. A range of syntactic patterns are involved in this category. For example, *consider seriously*, *lean meat*, *soft water*, *gin and tonic*, and *accept defeat*. Those combinations which are irreversible binomials such as *cash and carry*, *ups and downs*, *hit and miss*, and *swings and roundabouts* are also included in this category.

As shown above, we can see collocations on the continuum of restrictions of combinations. The less restriction on words combinations resulted in variation of combinations, while the more restriction on words to co-occur limits words to form collocations which resulted in less variation in the combinations.

2) *Syntactic structure*

This group consists of three sub-groups: flexible, regular with certain constraints, and irregular.

2.1) Flexible: this sub-group refers to those flexible combinations between words. For example, *break somebody's heart* and *nice/good/great to see you*.

2.2) Regular with certain constraints: this sub-group refers to those combinations between words with certain limitations. For example, *to drop a brick*, *to smell a rat*, and *we'll now take questions from the floor*.

2.3) Irregular: this sub-group refers to the combinations with no fixed patterns. For example, *to go one better*, *to be good friends with somebody*, *to hold true*, *to go it alone*, and *the more the merrier*.

This group slightly differs from the first group on restriction as it looks at the structure of the combinations. Flexible structure enables the combinations to have high variations. Regular with certain constraints limits words to co-occur. Irregular structure, however, does not necessarily limit words to be combined, only their unpredictable patterns that may cause some confusion.

3) *Semantic opacity*

This group consists of four sub-groups: transparent, semi-idioms/metaphor/ idiomatic similes, semi-transparent, and opaque.

3.1) Transparent: this sub-group refers to those combinations with clear meanings such as *long time, no see*; *when all is said and done*; *honesty is the best policy*; *we're just good friends*.

3.2) Semi-idioms/metaphor/idiomatic similes: this sub-group refers to those combinations which can be used in either direct or indirect meanings such as *we are all in the same boat*; *an open-door policy*; *a fat salary*; *as sober as a judge*.

3.3) Semi-transparent: this sub-group refers to the combinations with meanings that are related to other things such as *the business really took off*; *to get round somebody*; *a watched pot never boils*; *a skyscraper*; *there's more here than meets the eye*; *bumper to bumper traffic*.

3.4) Opaque: this sub-group refers to the combinations with meanings are unclear without contexts and totally unclear. There are two groups under this category: overt and covert. Overt refers to the combinations which meanings can be interpreted with sufficient contextual or cultural knowledge. For example, *O.K.*; *right on*; *yuk*; and *bottoms up*. Covert refers to the combinations which meanings cannot directly interpret from the words. For example, *to be on the wagon*; *to be on the ball*; *to carry the can*; and *to kick the bucket*.

This group can be put on the continuum of meaning from clear to unclear with transparent is placed on the one end and opaque is on the other. The combinations with transparent meaning are likely to produce collocations with clear meaning which can be understood right away with no need of interpretation. For the combinations with opaque meaning, on the other hand, interpretation with the context is needed to understand them.

Nation (2001) suggests that the most effective way classifies collocations using a set of scales into ten scales as follows:

1. *Frequency of co-occurrence*: The scale ranges from 'frequently occurring together' to 'infrequently occurring together'. This can be done by doing computer-based study of corpora.

2. *Adjacency*: Collocates can occur next to each other or separated by variable words or phrases. Thus, the scale ranges from 'next to each other' to 'separated by several items'.

3. *Grammatically connected*: Collocates normally occur within the same sentence as a part of a grammatical construction. Sometimes, however, it is possible for the items within the same text without grammatical connection to each other but in a lexical cohesion relationship as collocates. The scale ranges from 'grammatical connected' to 'grammatical unconnected'. For example, '*silk*' often occurs with a color as in 'Her uniform was of rich raw *silk* in a *shade* which matched with her hair'. According to Kennedy (1998 as cited in Nation, 2001), '*silk'* and '*shade'* is considered as collocates without a strong grammatical connection.

4. *Grammatically structured*: There are some cases that can be classified as 'grammatical connected', but cannot be a collocation that takes account of the major divisions that would be made in analyzing a clause. In this case, the grammatical structure criterion can be applied using a list of permitted structures (Kjellmer, 1982). The scale for this type ranges from 'well structured' to 'loosely structured'. For example, '*although he*', '*of the*', and '*but too*' are not considered as collocations although they are likely to co-occur frequently.

5. *Grammatically uniqueness*: There are some collocations that are grammatically unique. The scale for this type ranges from 'grammatically unique' to 'grammatically regular' with patterned exceptions as the mid-point. For example, '*hell for leather*' is considered grammatically unique, while '*go to bed/town*' is grammatically regular.

6. *Grammatically fossilization*: There are collocations that do not allow any changes in word order, some allow small changes, and some allow substantial changes. For example, '*kick the bucket*' cannot be changed to '*the bucket is kicked*'. The scale for this type ranges from 'no grammatical variation' to 'changes in part of speech', with 'inflectional change' as a mid-point.

7. *Collocational specialization*: Some collocates only occur together. Some consist of one item that only occurs in the presence of the other item, but the other item is not under the same restriction. Some collocations consist of items that can also occur with a range of other collocates. This phenomenon is called by Aisenstadt (1981) as collocational specialization 'restricted connectivity'. For example, '*kith and kin*' the word '*kith*' is limited to this phrase. The scale ranges from 'always mutually co-occurring' to 'all occurring in a range of collocations' with 'one bound item' as the mid-point.

8. *Lexical fossilization*: There are some collocations with collocates that cannot be replaced by other words, but some other can with words of related meaning. Sinclair (1987) calls it 'internal lexical variation'. The fossilization of combinations leads the combinations to be idioms with fixed and unchangeable words to be combined such as '*a bird's eye view*'. Those with 'internal lexical variation' are the combinations that allow some variation of words to be combined such as '*last year*', '*last week*', and '*last night*'. The scale ranges from 'unchangeable' to 'allowing substitution in all parts' with 'allowing substitution in one part' as the mid-point.

9. *Semantic opaqueness*: The most idiomatic collocations are those the meaning of the whole cannot be drawn from their combinations. For example, '*kick*

the bucket' and '*for good'*. The scale for this type ranges from 'semantically opaque' to 'semantically transparent'.

10. *Uniqueness of meaning*: Some collocations have only one meaning, while some others have more than one. For example, '*answer the door'* and '*keep promise'*. The scale for this type ranges from 'only one meaning' to 'several meanings' with 'related meanings' as the mid-point.

In addition to the above mentioned, Handl (2008) develops a multi-dimensional classification of collocations consists of semantic dimension, lexical dimension, and statistical dimension.

1) Semantic dimension considers the meaning of the combination as a criterion. If the meaning inside the combination is the same as the meaning outside, the expression is maximally transparent and is positioned towards the free-combination endpoint of the dimension. If the meaning outside the combination does not help in understanding the expression, it is a semantically opaque idiom.

2) Lexical dimension considers the size of the collocational range. The range of a node word can be obtained from the list of all the co-occurring lexical items from its concordance. A typical collocation may consist of elements chosen from a restricted set of lexical items. A node with restricted range is likely to be an idiom and a compound, while a node with a large range tends to be a collocation.

3) The statistical dimension considers the statistical scores as high scores show the tendency to be a collocation and low scores combination is likely to be an idiom.

The above classifications of collocation show that there are variations in doing so. However, the classification given by Handl (2008) seems to be the most comprehensive by covering almost all of the others. The semantic dimension can

cover semantic opacity given by Carter and semantic opaqueness and uniqueness of meaning given by Nation. The lexical dimension can cover collocation restrictions and syntactic structure given by Carter and those given by Nation such as adjacency, grammatically connected, grammatically structured, grammatically uniqueness, grammatically fossilization, collocational specialization, and lexical fossilization. The statistical dimension also covers frequency of co-occurrence as given by Nation.

2.2.3 Types of Collocation

According to Bahns (1993) and Schmitt (2000), there are two basic types of collocations: grammatical/syntactic collocations and semantic/lexical collocations. Grammatical collocations are those combinations between a dominant word with a grammatical word. The examples are *abide by*, *access to*, and *acquainted with*. Lexical collocation, however, usually are those combinations between two equal words such as Noun + Verb, (*e.g., ball bounces*), Verb + Noun (e.g., *spend money*), and Adjective + Noun (*e.g., cheerful expression*). Besides these two basic types of collocations, Allerton (1984), suggests another type where the combinations are not based neither on grammatical nor lexical patterning. This type of collocation focuses on the relatively arbitrary propositions attached to time since there are no logical reasons on the use of *at*, *on*, or *in* before time. For example, we use *at* noon and *at* night, but *on* Monday, and *in* June. This type of collocation, however, may be considered as a kind of grammatical collocation as prepositions are the part of the combination.

Hausmann (1990 as cited in Bartsch, 2004) classifies collocations into six types with the emphasis on the combinations of content words. These six types of collocations are presented in Table 2.2 below.

Table 2.2 Types of collocations as categorized by Hausmann (1990)

Types	Combinations	Examples
1	Verb + Substantive (Object)	to tackle a problem
2	Adjective + Substantive	weak tea
3	Substantive (Subject) + Verb	the heart palpitates/throbs
4	Substantive + Substantive	a pack of dogs, a pride of lions
5	Adverb + Adjective	keenly aware
6	Verb + Adverb	hurt badly/seriously/deeply/slightly

Hill (2000) also suggests seven types of collocations as a guide for teachers to introduce them to their students. These seven types of collocations are shown in Table 2.3 below.

Table 2.3 Types of collocations as categorized by Hill (2000)

Types	Combinations	Examples
1	Adjective + Noun	a huge profit
2	Noun + Noun	a pocket calculator
3	Verb + Adjective + Noun	learn a foreign language
4	Verb + Adverb	live dangerously
5	Adverb + Verb	half understand
6	Adverb + Adjective	completely soaked
7	Verb + Preposition + Noun	speak through an interpreter

In the BBI Combinatory Dictionary of English, Benson et al. (2010) view collocations as "fixed, identifiable, non-idiomatic phrases and constructions" (p. xix). Similar to Bahns (1993) and Schmitt (2000), they categorize collocations into two major groups: grammatical and lexical collocations. Grammatical collocations are divided into 8 categories as shown in Table 2.4 as follows:

Table 2.4 Grammatical collocations as categorized by Benson et al. (2010)

Types	Combinations	Examples
G1	Noun + preposition combinations	music in the park, breakfast in bed
G2	Noun + to + infinitive	pleasure to do, right to do, need to do, promise to do
G3	Noun + that clause	an agreement that she would represent us in court
G4	Preposition + noun combinations	by accident, in advance
G5	Adjective + preposition combinations	angry at everyone
G6	predicate adjectives + to + infinitive	necessary to work, ready to go
G7	Adjective + that clause	afraid that, necessary that
G8	19 verb patterns	

  1 = svo *to* o (or) svoo   11 = sv possessive v-ing
  2 = svo *to* o   12 = sv(o) *that*-clause
  3 = svo *for* o (or) svoo   13 = svo *to* be c
  4 = sv prep. o (or) svo prep. o   14 = svoc
  5 = sv *to* inf.   15 = svoo
  6 = sv inf.   16 = sv(o)a
  7 = svv-ing   17 = sv(o) wh-word
  8 = svo *to* inf.   18 = s(*it*)vo *to* inf. (or) s(*it*)vo *that*-clause
  9 = svo inf.   19 = svc (adjective or noun)
  10 = svov-ing

Notes: s = subject; v = verb; o = object (direct or indirect); c = complement;
a = adverbial (when obligatory); v‑ing = verb form in -ing.

In terms of lexical collocations, there are seven categories provided as shown

in Table 2.5 below.

Table 2.5 Lexical collocations as categorized by Benson et al. (2010)

Types	Combinations	Examples
L1	Verb + Noun Verbs denote *creation* and/or *action*	compose music, commit suicide

L2	Verb + Noun	dispel fear, ease tension
	Verbs mean *eradication* and/or *nullification*	
L3	Adjective + Noun	reckless abandon, rough estimate
L4	Noun + Verb	bombs explode, blood circulates
L5	Noun + *of* Noun	school of whales, act of violence
L6	Adverb + Adjective	deeply absorbed, hopelessly addicted
L7	Verb + Adverb	affect deeply, argue heatedly

In the present study, the researcher follows the classification suggested by Benson et al. (2010) as they provide clear distinctions between the two groups of collocations. For lexical collocations, they cover all major combinations of the parts of speech. However, the types of combinations will be reduced from seven to six by merging Benson et al.'s L1 and L2 and making a change in L5 from Noun + *of* Noun to simply Noun + Noun to better suit for the study as its focus is on the lexical collocations which arise from the combinations of content words. The adjusted lexical collocations to be investigated in the present study are shown in Table 2.6 below.

Table 2.6 Lexical collocations adapted from Benson et al. (2010)

Types	Combinations	Examples
L1	Verb + Noun	compose music, ease tension
L2	Adjective + Noun	reckless abandon, rough estimate
L3	Noun + Verb	bombs explode, blood circulates
L4	Noun + Noun	wound dressing, road accident
L5	Adverb + Adjective	deeply absorbed, hopelessly addicted
L6	Verb + Adverb	affect deeply, argue heatedly

2.2.4 Criteria for Identifying Collocation

According to Cowie and Howarth (1995), lexical phrases or collocations can be put on a 4-level scale of complexity: idiom, invariable collocation, collocation with limited choice at one point, and collocation with limited choice at two points.

At level 1, idioms represent multiword lexemes that consist of frozen collocation. This means that the combination is fixed and cannot be replaced by other words. This level has the least complexity and variation since there is no variation allowed. Therefore, if any variation is inserted into the idiom, the existence as a unit will be collapsed. For example, the idiom '*kick the bucket*' which means '*to die*' will no longer be an idiom if any part of it has been changed such as '*kick the pail*', '*boot the bucket*', or '*kick a bucket*'.

In level 2, invariable collocation, variation is still not allowed. However, the meaning is different from that of idioms. The meaning of each collocation comes from the words that are combined which cause this level to be more complex. These combinations are such as '*smart phone*' and '*outer space*'.

For level 3, collocation with limited choice at one point, there is a slot for a limited list of words with mostly similar in meaning to be filled. This increases in both variation and complexity to the combinations. These combinations are such as '*science*' can be with other words like '*computer science*', '*life science*', and '*health science*'.

Level 4, collocation with limited choice at two points, is similar to that of level 3 only with two slots to be filled instead of one. This level has the most variation and complexity. For example, within the combination of '*urban legend*', each word can be

combined with other words such as '*urban living*' or '*urban area*' and '*untold legend*' or '*well-known legend*'.

From the criteria presented above, although both idioms and collocations lexically belong to similar group as lexical phrases or multiword lexemes, there is a clear distinction between them. The fixedness of combinations and their meanings make the combinations become idioms, while the variation allowed although in different degrees of the combinations makes collocations.

2.2.5 Importance of Collocation

According to Nation (2001), collocations play very important roles in knowing a language with three aspects. Firstly, language knowledge is collocational knowledge. This is because the stored sequences of words are the bases of learning, knowledge and use. Secondly, all fluent and appropriate language use requires collocational knowledge. To produce a native-like language, collocations play the essential role in it. Finally, many words are used in a limited set of collocations and knowing these is part of what is involved in knowing the words.

Nation (2001) also gives three major types of evidence to support the issue that collocations may be important building blocks in language use and language learning as follows:

1) There is the intuitive feeling that certain phrases seem to act as units.

2) There is the evidence from corpus studies that certain groups of words recur.

3) There is evidence from studies of learning and knowledge.

Kozlowski and Seymour (2003) provide a number of aspects of collocation to show its importance to the EFL/ESL learners as follows:

1) Collocation constitutes a large part of our mental lexicons. They explain that first language (L1) learners have the ability to store in their heads huge amounts of memorized text, addresses, telephone numbers, dates, poetry, proverbs, idioms, names and many others. The memorized information can be quickly and easily retrieved when required. In case of second language (L2) learners, they also have the ability to store language in chunks. Repetition plays the important role in memorization. Therefore, language patterns need to be heard, written, spoken and read a number of times in order for them to become fixed. As a result, recycling of language patterns through listening, speaking, reading and writing activities should be encouraged.

2) Collocation enables students to express ideas clearly and accurately. Students need to be shown examples of how words are used and what words they collocate with. Teachers must instruct students that words cannot be learned in isolation. For example, one of the meanings of the verb *to diet* is *to lose weight*. In order for students to use the word effectively, they need to know that a person can */diet/go on a diet/be on a diet/start a diet/*, but not *make a diet*.

3) Collocation improves writing. L2 learners often produce awkward and unnatural language. Get learners to focus on awkward language and to replace it with appropriate collocations is one way of helping them to improve their writing. Writing can be improved by consistently looking for, and recording, language patterns in context to increase the chances of acquiring meaningful language.

4) Collocation can help students surpass that intermediate plateau. Students often become discouraged at the upper-intermediate level. This is because grammatical structure becomes less problematic for them but collocation competence

deters them from getting ahead. The ability to produce accurate and natural language makes them excited and eager to learn as well as allows them to think quickly and communicate effectively.

5) Collocation improves rhythm and stress. Once students are able to memorize longer collocation patterns, their stress and intonation become more natural. This is because the ability to produce a large number of collocations and longer patterns enables students to learn the stress patterns of whole phrases, and leads to better stress and intonation.

Bennett (2010) points out that studying collocation is important as it should give a deeper understanding of the meaning and use of a word than simply studying a single word alone. Collocations can also provide better understanding of particular words which are used in a certain phrase. She gives an example of the use of '*between*' and '*through*' studied by Kennedy, which found that '*between*' is usually used after nouns, whereas '*through*' is more frequently found after verbs (1991, p. 107, as cited in Bennett, 2010). Duan and Qin (2012) also assert that collocation is an important aspect in vocabulary acquisition as it is a universal linguistic phenomenon. Collocations as well enhance the ability of the learners to memorize new words. Farrokh (2012) stresses the importance of developing collocation knowledge of language learners in three aspects. Firstly, language knowledge requires collocational knowledge. As collocations are everywhere, a word-by-word approach is no longer effective. Secondly, efficient language acquisition requires collocational knowledge. As the human minds tend to chunk language to make it easier to process, collocations have the important role to play. Finally, fluent language use requires collocational knowledge.

From the views presented above, one quality that makes collocation important is its large coverage in a language which makes it the integral part of the language. Knowing collocation, therefore, helps shorten the process of producing the language as they are stored in the mental lexicon and are readily accessible. Another quality of collocation is that, as it is an integral part of a language, it helps in natural and appropriate production of the language. Furthermore, collocation acts as important building blocks in language use and language learning. This is because they are prevalent in the language and can be found in all genres of language use. Additionally, knowing collocation enables deeper understanding of the language more than that of knowing single words. This can also improve the main four skills of that language.

2.2.5.1 Previous Studies on Importance of Collocation

Previous research studies have confirmed that collocation knowledge plays an important role in effective use of a language of all skills. The examples of these studies are presented as follows:

Hsu (2007) investigated the use of English lexical collocations and their relation to the online writing of Taiwanese college English majors and non-English majors. 41 English major and 21 non-English major students at a national university of science and technology in southern Taiwan were the participants of the study. Each student was asked to take a 45-minute online English writing test, administered by the web-based writing program, Criterion Version 7.1 to examine the use of lexical collocations (i.e. frequency and variety). The test was also used to measure writing scores of the two groups. The findings showed that there was a significant correlation between all the students' tokens of lexical collocations and their online writing scores.

For the same category of correlation, no obvious relation was found if the English and non-English majors were considered separately. However, the significant correlation was found between the English majors' types of lexical collocations and their online writing scores in that the higher the writing scores gained indicates the more types of lexical collocations used in the writing. Thus collocational competence is important in effective communication, as writing is one of its forms.

The effect of different levels of lexical collocational density on EFL learners' reading comprehension was conducted by Sadighi and Sahragard (2013). In the study, 80 sophomore students with different levels of proficiency studying at Zand Institute of Higher Education in Shiraz, Iran were chosen based on their score distribution on a reduced TOEFL (Test of English as a Foreign Language) test constructed by Educational Testing Service (ETS, 1998). Forty participants were randomly assigned to the control group, while the other 40 were in the experimental group. Another instrument used in this study was a lexical collocation test containing two texts (as pre- and post-tests): A high and a low lexical collocational density tests designed by the researchers. A few paired/independent sample t-tests, and a two-way repeated measure were used to answer the five research questions. Results indicated that texts with high lexical collocational density influenced learners' comprehension positively. Although the instruction of lexical collocation did not have any effects on answering the vocabulary items significantly, teaching lexical collocations affected learners' reading skills positively.

The relationship between the knowledge and use of collocations and speaking proficiency was investigated by Mohajeri and Ketabi (2013). The participants were 20 Iranian EFL learners from Sadr Institute of higher Education who had passed

placement test and were preparing for IELTS (The International English Language Testing System) test. The students took two tests: a lexical collocational test and an IELTS speaking test. Data analysis for correlations between the participants' knowledge of lexical collocations and their speaking proficiency showed a significant positive correlation between their knowledge of lexical collocation and their speaking proficiency in IELTS test. The researchers concluded that knowledge of lexical collocation play a significant role in developing speaking proficiency.

The relationship between knowledge of collocation and reading, writing, speaking and listening proficiency was also explored by Yazdandoost, AmalSaleh, and Kafipour (2014). The participants were 50 Iranian EFL graduate students of different fields of study in an English language institute in Shiraz, Iran. They took a test of both lexical and grammatical collocations to measure their collocational knowledge. Then, an IELTS sample test (Version two) was administered to find the students' reading, writing, speaking and listening proficiency. Pearson correlation coefficient illustrated a significant correlation between knowledge of collocation and reading (P = 0.724), writing (P = 0.724), listening (P = 0.706) and speaking (P = 0.885) proficiency. Regression model was conducted to find the exact contribution between variables. It indicated that knowledge of collocation can be a predicator for all four language skills. The researchers concluded that knowledge of collocation proved to be a prerequisite for successful language learning. This research confirmed the influential role of collocation knowledge in essential language learning.

From the previous studies presented above, it is evident that lexical collocations play an important role in all skills of the language. Lexical collocations can enhance writing, reading, speaking, and even listening ability of language

learners. This can lead to effective communication in general which should facilitate them with the ability to handle communication tasks of all forms effectively and confidently.

2.2.6 Teaching of Collocation

According to Nesselhauf (2005), many types of prefabricated chunks, including collocations, have not been paid adequate attention in English language teaching. However, as teaching words in chunks by means of lexical approach can enhance learners' understanding through the meaning negotiation process (Nattinger & DeCarrico, 1992; Willis & Willis, 2006), it is essential to help learners be aware of chunks, enable them to identify, organize, and record them. This is when corpus-based teaching and learning of a language has a role to play (Sinclair, 1997).

When it comes to teaching of collocation, there are some suggestions from scholars in the field who value the teaching of collocation. Hill (2000) does not perceive collocation as merely an additional aspect of language learning. Thus, it deserves to be paid attention from lesson one. He suggests four ways of teaching collocation as follows:

1) Teaching individual collocations: this is to treat collocations the same way as individual words. For example, instead paying attention to single words such as *bath*, *friends*, *belief*, the attention can be shifted to *take a bath*, *make friends*, *strong belief*, *belief in God*, respectively. There words such as *speak*, *say*, *tell* where the different among them only be made by the knowledge of their collocations. This point is supported by the studies by Bahns and Eldaw (1993); Hashemi, Azizinezhad, and Dravish (2011); Rahimi and Momeni (2012); Balcı and Çakır (2012).

2) Making students aware of collocation: this is to help learners become more independent on learning collocations. One way to do this is to encourage learners to think bigger than the word. For example, words that occur with *speak* may be *a foreign language*, *in public*, *fluently*, *English*. This point is supported by the study by Farrokh (2012).

3) Extending what students already know: this is to expand the knowledge of 2,000 words of vocabulary to cover their co-occurring words. For example, students who know a single word '*make*' plus its collocations such as *make a mistake/ a meal/ a complaint/ friends* will use the language much better. This point is supported by the studies by Willis (1998); Balcı and Çakır (2012).

4) Storing collocations: this is to make effort in keeping locations being learned. Writing down and organizing them in the way that is easy to look back is recommended. This point is supported by the study by Akpınar and Bardakçı (2015).

Similarly, Nation (2001) recommends how to teach collocation according to three points on a scale (shown below in Table 2.7): idioms, at one end; allow some substitution, at the middle; and allow a lot of substitution and grammatical change and are transparent, at the other end. For the first group, idioms, they need to be dealt with as if they were single words as they are fixed in their combinations. The explanation on their history and analysis of their parts, as well as how they function in discourse should be taken into account when teaching. For the second group, allow some substitution, it is necessary to look for any patterning that occurs. Their frequency of occurrence is the starting point for dealing with the range of related collocates. The predictable collocations should be treated as part of the enrichment of the individual collocates that make them up. Some very frequent collocations can simply be

memorized and used, and later be analyzed when the learners' level of proficiency is more advanced. For the final group, allow a lot of substitutions and grammatical changes with transparency, the learning burden is dependent on the words that are combined. The learning burden is high when its form, meaning, and use are not predictable. On the other hand, the learning burden is low when its form, meaning, and use are in predictable patterns.

Table 2.7 Methods of teaching collocation on the scale of combination (Nation, 2001)

Idioms	Allow some substitution	Allow a lot of substitution
Teach as a single word, explain history, analyze the combination, and function in discourse	Look for pattern of occurrence based on frequency and predictability	Look for pattern of occurrence, the combination of words plays a major role in the learning and teaching

Vasiljevic (2014) also suggests two main ways of how collocations should be taught: dictionary training and teaching activities. She points out that dictionaries are useful means for acquiring collocations as they often include common and useful collocations that can help learners improve their fluency, precision, and naturalness of expression. These dictionaries are such as the Cambridge Advanced Learner's Dictionary (CALD), the Oxford Collocations Dictionary for Students of English (OCDSE), and the Macmillan Collocations Dictionary. In terms of teaching activities, collocations must be a part of the planned language input. As there are a large number of lexical chunks that exist, teachers must be highly selective in their choice of target phrases, which the frequency of use is one of the most commonly criteria. Teachers must also be selective with regard to the collocation patterns to be taught. Teachers should also try to help learners remember common word combinations by making the learners aware of the linguistic motivation of multiword chunks. Review is crucial for

the acquisition of multiword chunks. Thus learners must be given opportunities to encounter the target collocations, in different contexts and in a relatively short period of time, so that memory traces can be formed. Vocabulary substitution exercises can also be useful to consolidate learners' collocation knowledge.

2.2.6.1 Previous Studies on Teaching of Collocation

There a number of previous research studies on how teaching of collocation helps EFL/ESL learners to improve their effective use of the language. These studies are conducted by scholars such as Pirmoradian and Tabatabaei (2012); Szudarski (2012); Eidian, Gorjian, and Aghvami (2013); Ördem (2013); Shooshtari and Karami (2013); and Khonamri and Roostaee (2014). The details of their studies are presented as follows:

Pirmoradian and Tabatabaei (2012) examined the effect of applying Collins Collocation Dictionary as a concordancing tool on learning lexical collocations of Iranian EFL university students. 30 students were randomly selected and then divided into two groups: experimental group and control group. A pretest was taken from both groups at the same time. In the following week concordancing practice (task1) was given to the students in the experimental group and they were asked to work with 10 lexical collocations and identify (mis)collocations. At the same time students in control group received some texts and they were asked to notice the lexical collocations. Then, posttest 1 was given to the students in both groups. In the following week, the same procedure with 10 other lexical collocations was repeated and posttest 2 was administered. The results showed that the experimental group performed better on lexical collocations than the control group and their results were significant.

The effect of meaning-focused and form-focused instructions on the acquisition of collocations was studied by Szudarski (2012). Forty-three L1 Polish EFL learners were divided into three groups: meaning-focused instruction plus focus-on-forms (MFI plus), meaning-focused instruction (MFI only) and a control group. During a three-week treatment, the two experimental groups were provided with two different types of instruction. The MFI plus group read stories that contained target collocations and additionally completed explicit exercises focused on collocational patterns, while the MFI only group read the same stories but no mention of collocations was made. The target collocations were verb-noun combinations with frequent delexical English verbs (e.g. 'give birth' or 'take a step') likely to be known by participants receptively but causing difficulty in language production. Three tests used to assess collocational competence at different levels of vocabulary mastery revealed that MFI followed by Focus on Forms (FonFs) was an effective way of enhancing learners' collocational knowledge at both the productive and receptive level, whereas MFI only does not seem to lead to much improvement.

The impact of lexical collocation instruction on learners' writing proficiency was conducted by Eidian et al. (2013). The study was conducted with 50 pre-intermediate Iranian language learners studying English at Ahvaz Islamic Azad University in Iran majoring in EFL teaching. They were selected through non-random convenient sampling procedure. They were randomly divided into experimental and control groups. The control group was taught based on conventional methods of writing instruction and the experimental group received treatment based on lexical collocation instruction in writing one paragraph essays. The design of the research was based on pre and post-test method. Pre-test was a lexical collocation test included

35 items administered before the treatment period to make the researchers sure that the groups' homogeneity on lexical knowledge in writing paragraphs. During the treatment period, five topics were administered to the participants to write one paragraph essay for each topic. A post-test on lexical collocation consisting of 35 items of multiple-choice, matching, and cloze task dealing with lexical collocations acquired through the treatment was administered. The results showed that there was a significant difference between the scores of the participants in the control and experimental groups. In addition, all the one paragraph essays of the study were analyzed through analyzing the components of writing. The results also showed that there was a significant difference between the mean scores of control and experimental groups in writing these components. Lexical collocation instruction developed the writing components of vocabulary and mechanics rather than grammar, relevance, and fluency in writing one paragraph essays.

A study to investigate whether teaching vocabulary via collocations would contribute to retention and use of English was conducted by Ördem (2013). A quasi-experimental design was formed to see whether there would be a significant difference between the treatment and control groups. Three instruments developed were conducted to 60 participants. The experimental group was taught collocations through lexical approach by means of ten different kinds of activities for ten weeks. On the other hand, the control group was taught in a traditional way, only focusing on word definitions from dictionary, antonyms, synonyms and guessing from the text. The results showed that the participants in the experimental group outperformed the ones in the control group in all of the three instruments. The study also indicated that a period of treatment and exposure to lexical collocations led the treatment group to

remember and produce the collocations in the reading courses more appropriately and less deviantly than the control group. This result showed that teaching collocations in the class systematically week by week and scaffolding learners' progress could lead to better learners who can remember and use collocations in their reading comprehension in English.

A study to explore whether receiving treatment on the use of lexical collocations affects EFL students speaking proficiency was conducted by Shooshtari and Karami (2013). In this study, 50 pre-intermediate students of Iran Language Institute, Ahvaz branch were chosen and divided into two groups. In the instruction period of ten sessions, the experimental group received instructions on five common lexical collocation patterns such as verb-adverb, noun-verb, verb-noun, adverb-adjective, and adjective-noun. Both groups took the same test before and after the treatment to measure their knowledge of collocation patterning. They also participated in a speaking task to assess their use of lexical collocation and overall oral proficiency. Results showed that the instruction of lexical collocation had a positive effect on the learners speaking proficiency and a moderate effect on their use of lexical collocations. This suggests that receiving instruction on the use of lexical collocation patterning can be effective in the enhancement of EFL students' language skills, specifically, their oral proficiency.

The effect of form-focused versus meaning-focused tasks added to an Extensive Reading (ER) program on the development of lexical collocations was conducted by Khonamri and Roostaee (2014). 41 Iranian intermediate EFL students of English language and literature at the department of foreign languages in Mazandaran University participated in this study. A reading comprehension test taken

from TOEFL was used to measure candidates' reading ability to homogenize them in terms of their entry behavior. Moreover, Word Associates Test (WAT) developed by Read (1993, 1998) was administered to examine the participants' depth of vocabulary knowledge. Participants were divided into two experimental groups: Both groups were assigned to read extensively and do some after reading tasks; the first group was given a form-focused task (FFT) while the second group worked on a meaning-focused task (MFT). The results of paired and independent sample t-tests revealed the fact that both FFT and MFT groups progressed in the interval between the pre- and post-test, but, there was not a significant difference between the effects of form-focused and meaning-focused task.

2.2.7 Tests of Collocation Knowledge

According to Bachman and Palmer (1996), the primary purpose of a language test is to provide a measure to interpret as an indicator of a test-taker's language ability. Reliability and construct validity are essential elements for this. Reliability refers to consistency of measurement. A reliable test score will be consistent across different characteristics of the test situation or the test format. Construct validity, refers to the extent to which the test scores can be interpreted as an indicator of the ability or construct being measured. To obtain construct validity, the tasks or test formats employed have to be appropriate in measuring the ability the test intends to measure. This is when test construct has a role to play. Test construct, (McNamara, 2000), refers to aspects of knowledge or skill possessed by the test takers that the test aims to measure.

As language knowledge can be viewed as a domain of information in memory available for use by the metacognitive strategies in creating and interpreting discourse

in language use (Bachman & Palmer, 1996), collocation knowledge, therefore, can also refer to a domain of information in memory in part of collocations available to use similar to other aspects of language. This includes receptive and productive knowledge. According to Nation (2001), receptive knowledge refers to ability to comprehend a language when the input is either listening or reading. Productive knowledge, on the other hand, is the ability to produce a language by means of speaking or writing. In terms of collocations, therefore, receptive knowledge is the ability to know which pair or set of words are likely to co-occur, as for using them appropriately in speaking or writing to show the productive ability.

In order to assess learners' needs and their lexical progress, Vasiljevic (2014) suggests that a test of collocation knowledge is necessary. A test, in general, can basically have a positive effect on the learners as it should help increase their awareness concerning collocations. Popular testing formats of collocation knowledge are:

1) L1-L2 translation;

2) L2 sentence cloze items;

3) sentence generation tasks; and

4) discrete tests where a node-word is provided and test takers are required to select or give one or more of its collocates such as in a matching test and a multiple choice test.

Jaén (2007) states that, similar to vocabulary tests, tests of collocation knowledge can be divided into two categories: the ones which are designed to test productive knowledge and those to test receptive knowledge. The tests of productive knowledge are those tests that intend to measure whether the learners are able to use

the collocations appropriately in given situations. The test formats are namely translation tasks where test takers have to translate the given collocations from L1 to L2 or vice versa. Another test format used is gap filling tasks where test takers are required to fill the missing pairs of the given collocations. The tests of receptive knowledge, meanwhile, are those tests that intend to explore if the test takers are able to select the most appropriate collocates of the nodes given in certain situations. The test formats are those in which alternatives are provided such as in a multiple choice format.

Hargreaves (2000) points out that tests that include the measurement of appropriate use of collocations are a standard part of University of Cambridge Local Examinations Syndicate's (UCLES) examinations. UCLES is responsible for the five Cambridge Main Suite Examinations which are the Key English Test (KET), the Preliminary English Test (PET), the First Certificate in English (FCE), the Certificate in Advanced English (CAE), and the Certificate of Proficiency in English (CPE). The examples from CPE, which tests the proficiency of the use of English, are as follows:

1. **breach**~code (of ethics)

Any doctors who …………. the medical profession's code of ethics is severely reprimanded.
A fractures   **B** cracks   **C** ruptures   **D** breaches

2. **pursue**~point

She obviously didn't want to discuss the matter so I didn't ………… the point.
A maintain   **B** follow   **C** pursue   **D** chase

In order to select the appropriate collocation, the learner has to know both what is possible and what is not.

Leśniewska (2006) also claims that there are possible ways of examining the learners' use of collocations. The language they produced either written or spoken may be used to look for collocations. Another way could be done by using test items either by gap-filling or multiple choice tests. Psycholinguistic tests based on word association patterns such as identifying correct pairs of collocations could also be applied to indirectly investigate the structure of the mental lexicon of learners.

According to Siyanova and Schmitt (2008), there are three general types of elicitation tools for measuring and assessing collocational competence and development of learners. The first type is in the form of written online tasks. This may be done by assigning learners to write essays or other forms of writing. The second type is off-line elicitation tools in the form of productive translation tasks, cloze format tasks, and association tasks as well as receptive multiple-choice and judgement tasks. The third type is on-line reaction tasks which are useful in tapping into the processing of collocations in language use.

Despite various formats applicable in the tests of collocation knowledge, it is agreeable that there are two categories of the test: to assess receptive and productive knowledge. A multiple-choice format seems to be used the most in testing receptive knowledge, while a gap-filling or cloze test format and a sentence generation are used in testing productive knowledge. Additionally, there has not been any standardized test designed specifically to assess the knowledge of collocations.

2.2.7.1 Previous Studies on Tests of Collocation

A translation task and a gap filling task were applied by Bahns and Eldaw (1993). Open-ended and multiple-choice cloze tests were used in a research study by Keshavarz and Salimi (2007). Ali Zarei and Baniesmaili (2010) used a fill-in the

blank test. A cloze test and a C-test were used in a research study conducted by Ebrahimi-Bazzaz, Samad, bin Ismail, and Noordin (2012). Multiple choice, gap-filling and translation of sentences containing collocations were used in the study by Miščin (2013). A gap-filling, a collocation selection test, and descriptive writing were used in a study by Suwitchanphan and Phoocharoensil (2013). A productive collocational test and a receptive collocational test were applied in the study by Torabian, Maros, and Subakir (2014). A specially designed test was used in El-Dakhs'(2015) study. The details of each study are presented as follows:

Bahns and Eldaw (1993) examined German advanced EFL students' productive knowledge of English collocations applying translation task and a gap filling task. In the translation task, it was found that more than half of the unacceptably translated lexical words were collocates. The participants did not express the collocational phrases significantly better in the translation task, where it was possible to paraphrase, than in the cloze task. The researchers concluded that one cannot easily paraphrase one's way around collocations in order to avoid the problem which they present. It was also found that some collocations in the translation task were successfully paraphrased by many students while others were rarely successfully paraphrased. Thus suggestion was that EFL teaching should concentrate on those collocations which cannot readily be paraphrased.

The relationship between the collocational competence and overall language performance was investigated by Keshavarz and Salimi (2007). The study was conducted with Iranian EFL learners using open-ended and multiple-choice cloze tests. 50-item test consisting of grammatical and lexical collocations were applied. The results of the statistical analyses showed that there was a significant relationship

between collocational competence and performance on cloze tests. Assuming the cloze test to be an effective measure, the findings point to the importance of improving EFL/ESL learners' collocational knowledge to enhance their proficiency level in the target language.

The effect of different patterns of lexical collocations on the recognition and production was examined by Ali Zarei and Baniesmaili (2010). The study was conducted with Iranian upper-intermediate learners of English. A fill-in the blank test, consisting of 111 items, was given to the participants after the recognition test in order to measure their production of the collocational patterns. Results indicated that the participants performed relatively better on adjective-noun and noun-noun combinations of lexical collocations, although the differences among the scores of the participants on the different patterns were not statistically significant. Analysis showed that the participants' performance on the noun-verb pattern was significantly poorer than the other patterns.

The relationship between the language proficiency and the knowledge of verb-noun collocations was conducted by Ebrahimi-Bazzaz et al. (2012) with Iranian EFL learners applying a cloze test and a c-test. The language proficiency level of the participants was scrutinized through their performance on a cloze test. A 50-item c-test was used consisting of verb noun collocations in which the verb missing but the first letter/phoneme was provided. The results of the statistical analyses demonstrate that there was a high positive relationship between collocational competence and general language proficiency of learners. Therefore, it can be concluded that the c-test is an effective measurement to assess learners' collocational competence.

Miščin (2013) used multiple choice, gap-filling and translation of sentences containing collocations to measure collocational knowledge of the first-year students of medical English. Concordance software such as Simple Concordance, Collocation Extract, TermeX were used. The corpus analysis established that the nouns "function" and "infection" occur with most verbs (30) followed by "pain" (28 verbs), "muscle" (24 verbs). 362 verbs occurred with nouns and among them the most frequent were "cause", "have", "develop", "treat", "prevent", and "produce". After that, the test was devised to examine which collocations students use with the most competence. 297 first-year students of School of Medicine in Zagreb were tested. The average result in multiple choice was 9.8 with the s = 2.0 as a standard deviation, in gap-fill 5.0 with s = 2.17, in translation into Croatian 6.7 with s = 2.10, and in translation into English 5.2 with s = 2.53. Then, glossary was made which should help future users of medical English.

Suwitchanphan and Phoocharoensil (2014) investigated how Thai EFL students studying in the regular and English programs use adjective-noun collocations as well as to find out the relationship between school curricula and collocational competence of adjective-noun using three tests: gap-filling, collocation selection test, and descriptive writing. The participants were 30 regular program students and another 30 English program students from a private secondary school in Bangkok. The main findings revealed that, in gap-filling test, the regular program participants (69.33%) scored higher than the English program participants (57.67%). The collocation selection test revealed that there was no significant difference between the curricula. The descriptive writing task showed that the regular program participants used more adjective-noun collocations (279 tokens) than did the English program

ones (211 tokens). The researchers suggested enhancing learners' collocational competence, especially adjective-noun collocations.

The possible relationship between collocational competence and vocabulary knowledge was also investigated by Torabian et al. (2014). The study was conducted with Iranian undergraduate learners. The participants were given a vocabulary test to identify their basic knowledge of common word meanings and to show how they can identify the meanings of words at different levels. Then, productive collocational test and a receptive collocational test were given to the participants to reveal the possible difference between the participants' receptive and productive knowledge. The results revealed that there was a significant difference between the receptive and productive lexical knowledge of the undergraduate learners. It was also revealed that there was a significant relationship between the collocational knowledge and the vocabulary knowledge of the learners.

The collocational competence of Arab undergraduate EFL students was examined by El-Dakhs (2015) using a specially designed test in a gap-fill and translation formats. Two exercises were used. The first exercise comprised 15 sentences with a missing verb where participants were required to fill in the missing verb per sentence in addition to an Arabic equivalent provided for the missing verb. The second exercise also consisted of 15 sentences, each with a missing adjective. Similar to the first exercise, the participants had to write the missing adjectives based on their understanding of the English sentences and the provided Arabic equivalents of the adjectives. The test was conducted to measure collocational competence of 90 Arab undergraduate learners at three academic levels in a private Saudi university was assessed. Findings showed that the collocational competence of learners was notably

unsatisfactory despite the fact that English is the medium of instruction at the University. It was also found that collocational competence improves with increased language exposure but at a slow rate, and that learners were more confident in their use of verb-noun collocations than adjective-noun collocations. The study also revealed that learners produce more intralingual than interlingual errors of collocations.

It is evident that the test of collocation can be used to measure the learners' collocational knowledge. The scores gained by the test takers should indicate their knowledge of the collocations. This will also lead to the step to be taken in response to the test results. To measure the knowledge of the learners, certain test formats have been used with the most popular are namely gap-filling, multiple-choice, and translation. In the present study, the test of collocation knowledge on collocations found in the SCNRA will be an off-line test in the form of multiple-choice, gap-filling, and a sentence writing task. These test formats are widely used by scholars in the field as seen in the previous studies. With these test formats, both receptive and productive knowledge of the test takers can be tested. Therefore, they are considered appropriate to be used to measure knowledge of nursing students on their collocations of their field extracted from the SCNRA.

2.2.8 Collocation Study in Thailand

Studies on collocations have been recently conducted in Thailand. These studies are such as by Khittikote (2011), Yumanee and Phoocharoensil (2013), Bueraheng and Laohawiriyanon (2014), Suwitchanphan and Phoocharoensil (2014), and Usen and Musigrungsi (2015).

Khittikote (2011) studied the ability to use collocations for business purposes by 50 Thai EFL learners as well as the relationship between the frequency of exposure to English and the ability to use collocations. The tests consisted of 15 multiple-choice and 10 blank-filling verb-noun collocations. A questionnaire was also used to investigate the frequency of exposure to English. The results revealed that the learners performed better in receptive test than productive one. In terms of the exposure to English, the majority of the participants had moderate level of exposure and no relationship was found with their collocation knowledge.

Yumanee and Phoocharoensil (2013) investigated collocational errors produced by 60 Thai EFL students. Two collocational tests, a 45-item multiple-choice test and an 18-item Thai-English translation test, were employed. The results showed that the students' performance in both the receptive test and the productive test appeared to be influenced by mother-tongue transfer. Additionally, it was found that the synonymy strategy, the learners' creative invention and the strategy of analogy, the paraphrasing strategy, and low knowledge of collocational skills were potential factors contributed to the high degree of collocational errors.

Bueraheng and Laohawiriyanon (2014) explored the degree of exposure to English language in relation to learners' collocational knowledge. COLLMATCH 3 receptive collocation test and productive collocation test were administered with two groups of 196 students. The results revealed that both groups of students had higher score on receptive test and International program students outperformed English major students. The researchers recommended that in order to elevate learners' productive collocational knowledge, a substantial amount of time should be devoted to learning activities such essay writing and conversation.

Suwitchanphan and Phoocharoensil (2014) investigated how Thai EFL students studying in the regular and English programs use adjective-noun collocations as well as to find out the relationship between school curricula and collocational competence of adjective-noun using three tests: gap-filling, collocation selection test, and descriptive writing. The participants were 30 regular program students and another 30 English program students from a private secondary school in Bangkok. The main findings revealed that, in gap-filling test, the regular program participants scored higher than the English program participants. The collocation selection test revealed that there was no significant difference between the curricula. The descriptive writing task showed that the regular program participants used more adjective-noun collocations than did the English program ones. The researchers suggested enhancing learners' collocational competence, especially adjective-noun collocations.

Usen and Musigrungsi (2015) explored the effectiveness of teaching collocations to grade 6 primary school students. Twelve lesson plans and a collocation test with 32 items were employed. The results revealed that students performed significantly better in the posttest. The vocabulary retention rate was 15.44 and significant. Verb-noun collocations were scored the highest, while Noun-preposition collocations were scored the lowest.

From the above samples of studies on collocations in the Thailand's context, it is evident that collocations in the specialized field have not been widely investigated, not to mention the field of nursing. The present study, therefore, should add up to what have already existed, particularly in terms of the knowledge of nursing students

on collocations in their professional field found in a sample corpus of research articles.

2.3 Corpus Studies

Since the present study involves a compilation of texts to be then analyzed using a computer software program, this section provides reviews of the literature on corpus studies. This will include definitions of corpus studies, the development of corpora and corpus studies, types of corpora, benefits of corpus studies, concordance software, corpus-based lexical analysis, identifying collocations from a corpus, as well as corpus-based language teaching and learning.

2.3.1 Definitions of Corpus and Corpus Studies

A corpus (plural: *corpora*), according to Cheng (2012), is a collection of texts that has been compiled for a particular reason. The collection of texts is based on a set of design criteria which the corpus aims to be representative. Bennett (2010) also states that a corpus is "a principled collection of authentic texts stored electronically that can be used to discover information about language that may not have been noticed through intuition alone" (p. 12). A corpus, according to Gries (2009), is a machine-readable collection of texts in either written or spoken that were produced in natural communicative settings. The collection of these texts is for two purposes. One is to be representative and balanced with respect to a particular linguistic variety or register or genre. The other is to be analyzed linguistically. A corpus can also be a large collection or database of language, incorporating stretches of discourse ranging from a few words to entire books. The applications of a corpus, according to Schmitt (2000) are: it can provide the frequency of occurrence of the words it contains; it can

reveal which words tend to co-occur; and it can illustrate the structure of the language.

It may be summarized that a corpus can be referred to a collection of texts either written or spoken of a particular genre or variety of genres for a particular purpose of analysis. A corpus is stored in an electronic form to be readable by a computer program. Corpus studies, therefore, is a study of the data in the forms and manners mentioned above. For the present study, a corpus should refer to a collection of the Sample Corpus of Nursing Research Articles (SCNRA) accessible online via SUT's library resources which is stored in an electronic form to be analyzed for identifying lexical collocations using a concordance program.

2.3.2 Development of Corpus Studies

The corpora have come to existence since the early 1900s. They were created manually with the hard work for a long period of time. The early computer based corpora were created in the 1960s. According to Schmitt (2000) and Dash (2009), the Brown University Corpus of American English (known as Brown Corpus) was the first of its kind created by Francis and Kučera at Brown University. The Brown Corpus was compiled during 1963-1964. It contains over 1 million words compiled from 500 samples of running text of edited English prose printed in the United States during the year 1961. Another corpus being created at the relatively the same time on the European counterpart was the Lancaster-Oslo/Bergen Corpus (known as the LOB Corpus). It was created by a group of scholars from the University of Lancaster, the University of Oslo, and the Norwegian Computer Centre for the Humanities. The project was based at the Department of Linguistics and Modern English Language at the Lancaster University in the United Kingdom. The LOB Corpus contained a

million words compiled from 500 British English texts from 15 categories originally published in 1961. The texts were selected by stratified random sampling. These two corpora were created by means of the written texts had to be manually typed in.

The advancement of computer technology has positive effect on the development of corpora. Once texts can be scanned and turned into text files and put into the computer, corpora can be created much easier and faster. With currently available technology, a corpus can contain hundreds of millions of words. The examples of these corpora are those of the well-known namely the COBUILD Bank of English, the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA).

COBUILD, an acronym for Collins Birmingham University International Language Database, is a British research facility set up at the University of Birmingham in 1980 funded by Collins publishers. The COBUILD Bank of English is an international English language project conducted by the COBUILD team at the University of Birmingham, UK. The text bank comprises 200 million words of both written and spoken English. The whole 200 million word corpus was annotated morphologically and syntactically during 1993-94 at the Research Unit for Computational Linguistics (RUCL), University of Helsinki, using the English morphological analyzer (ENGTWOL) and English Constraint Grammar (ENGCG) parser. The project was led by Prof. John Sinclair in Birmingham, and Prof. Fred Karlsson in Helsinki. By 1997, the corpus grew to comprise over 300 million words.

The British National Corpus (BNC) is a 100 million word collection of samples of contemporary written and spoken British English from various sources. The building of the corpus began in 1991, and was completed in 1994. The latest

edition is the BNC XML Edition, released in 2007. The written part of the BNC, which cover 90% of the entire corpus, includes extracts from regional and national newspapers, specialist periodicals and journals, academic books and popular fiction, published and unpublished letters and memoranda, school and university essays, and many other kinds of text. The spoken part, cover 10% of the corpus, consists of orthographic transcriptions of unscripted informal conversations which were recorded by volunteers selected from different age, region and social classes in a demographically balanced way. The spoken language was also collected in different contexts, ranging from formal business or government meetings to radio shows and phone-ins. The BNC is distributed in a format which makes possible almost any kind of computer-based research on the nature of the language. The application areas include lexicography, natural language understanding (NLP) systems, and all branches of applied and theoretical linguistics.

The Corpus of Contemporary American English (COCA) is the largest freely-available corpus of English. The corpus was created by Mark Davies of Brigham Young University. It is used by tens of thousands of users namely linguists, teachers, translators, and other researchers every month. The corpus contains more than 450 million words of American English texts and is equally divided among spoken, fiction, popular magazines, newspapers, and academic texts. The corpus added up 20 million words each year from 1990-2012 and is also updated regularly with the most up to date texts are from summer 2012. Because of its design, it is perhaps the only corpus of English that is suitable for looking at current, ongoing changes in the English language.

2.3.3 Types of Corpora

According to Gries (2009), there are two basic types of corpora: general corpora and specific corpora. Similarly, McEnery, Xiao, and Tono (2006), also state that there are two broad types of corpora in terms of the range and the text categories compiled in the corpus: general and specialized corpora. The former are those intend to be representative and balanced for an overall language description in general. The examples of this type of corpora are such as the British National Corpus (BNC), the COBUILD Bank of English Corpus, and the Corpus of Contemporary American English (COCA). The latter are restricted to a particular variety such as written or spoken, register such as medicine or law, or genre such as newspaper or journal article.

According to Bennett (2010), there are eight types of corpora which are: generalized, specialized, learner, pedagogic, historical, parallel, comparable, and monitor. The application of each type depends on the purpose of the study. She provides more information on four types of corpora as follows:

Generalized corpora: This is the broadest type of corpus, which is often large, consists of over 10 million words, and contains a variety of language. Generalized corpora provide a whole picture of a language. Examples of this type of corpora are the British National Corpus (BNC), the American National Corpus (ANC), and the Corpus of Contemporary American English (COCA). These corpora contain written texts from various sources such as newspapers and magazines articles, works of fiction and nonfiction, as well as scholarly written journals. These corpora also contain transcripts of spoken language from both formal and informal conversations, speeches, and meetings.

Specialized corpora: This type of corpora contains texts of a certain type and aims to be a representative of the language of this type. The size can be large or small depending on the questions to be answered. Examples of specialized corpora are the Michigan Corpus of Academic Spoken English (MICASE), which contains spoken language from a university setting; the CHILDES Corpus, which compiles language used by children; the Michigan Corpus of Upper-level Student Papers (MICUSP), which is a collection of papers from various disciplines; and a medical corpus which contains language used by nurses and hospital staff. This type of corpora is used in ESP setting.

Learner corpora: This type of corpora is considered specialized corpora, only the difference is that the texts are from those produced by students who are learning the language. Learner corpora are usually tagged and can be examined to see common errors made by students. This type of corpora can be useful for the teaching and learning of the language.

Pedagogic corpora: This type of corpora contains language used in classroom settings. They can contain academic textbooks, transcripts of classroom interactions, or any other written text or spoken transcripts in an educational setting. They can be used to ensure students are learning useful language, to examine teacher-student dynamics, or as a self-reflective tool for teacher development.

Tognini Bonelli (2010), however, classifies corpora as proposed in the course of a European Union (EU) project as follows:

1) Sample corpora: As most corpora cannot represent the entire language of the all periods of time and all genres, they are often referred to as 'sample corpora'. This type of corpus is like a 'snapshot' of particular types of text at a particular time.

The aim of this type of corpora is to present the normal linguistic features of a language or variety in approximately the proportion occurred in general use. The Brown Corpus is the example of this kind.

2) *Corpora for comparison*: Under the same design criteria and of the equal size, two or more corpora can be compared. The areas comparable are in terms of geographical and historical differences, topics, and contrastive features of the language use. In terms of geographic comparison, the examples are that of the Brown Corpus and the LOB corpus. The former represents the American English and the latter represents the British English with the texts collected in the same year of publication, 1961. In terms of historical comparison, the corpora are designed to be compared along a time dimension. This type of corpora is not very common and there are two kinds: diachronic and monitor corpora. The former represents 'snapshots' at intervals of time such as for a generation, while the latter can reveal the language change. In terms of topic, the corpora are created with texts organized by topic such as documents or reports of a particular persons or things. This type of corpora tends to be small in size. For contrastive corpora, their main components have been chosen to facilitate the study of variety. Each component, however, is designed without reference to the others. The example of this type of corpora is the Longman Grammar of Spoken and Written English compiled by Biber and his team.

3) *Special corpora*: these corpora are those compilations of texts which are not designed to be representative of a language or variety. The focus is on their extraordinariness and uniqueness of language choices. The examples of this type of corpora are the works of Shakespeare and Goethe. The first project of special corpora in turning text into electronic format is the works of St Thomas Aquinas. One of the

earliest corpora available is the Leuven Drama Corpus and Frequency List created in 1975 by L. K. Engels and colleagues at the University of Leuven.

4) Corpora along the time dimension: This type of corpora is similar to that of the above mentioned historical comparison corpora. There are two of them: diachronic corpora and monitor corpora. The former present 'snapshots' at intervals of time, at least a generation, the latter are compiled to investigate changes in the language of different times. The Helsinki corpus, which is a collection of English texts from c.750 to c. 1700, created at University of Helsinki, Finland, is the example of the diachronic corpus. The example of monitor corpora is the AVIATOR (Analysis of Verbal Interaction and Automatic Text Retrieval) project, which consists of an annual ten million words of *The Times* newspaper with software to detect innovations of various kinds, created at Birmingham University.

5) Bilingual and multilingual corpora: Bilingual corpora have been inspired by the bilingual nature of Canada with the need to have all information available in both French and English by means of electronic translation. Multilingual corpora are emerged from a relationship of translation among the constituent texts, which is also called parallel corpora. The first kind of this corpora is the Canadian Parliament's Proceedings which available in electronic format with both English and French. The example of the multilingual comparable corpus is that of the PAROLE Corpus which comprises all the official languages of the European Union. For the contrastive corpus, the corpus consists of two sub-corpora of the same language with one being translated and the other is not.

6) Normativeness: The Birmingham Collection of English Texts was set out to be normative by means of having adult native speakers as the originators of the texts

and having foreign learners as the recipients of the outcomes in the forms of dictionaries, grammars, and other publications. However, there has been possibility of corpus building to reassess the standards, targets, and models for non-native users and ex-colonial varieties.

7) Non-native speaker corpora: The language of learners can be explored much further by comparing with normative model corpora. The example of this kind of corpora is a project conducted at the Centre for English Corpus Linguistics (CECL) at the University of Louvain in Belgium.

8) Spoken corpora: Speech corpora are part of special corpora. The MICASE (Michigan Corpus of Academic Spoken English) corpus of academic American English is an example of this type of corpora. MICASE is created by English Language Institute at the University of Michigan and is freely accessible online.

From the reviews above, corpora can be classified into a number of types. The corpus in the present study, as it will be created solely from the selected journal articles in the field of nursing, can be well put under a specialized corpus as classification given by Bennett (2010) that it contains texts of a certain type and aims to be a representative of the language of this type, which is texts of journal articles in the field of nursing. The corpus in the present study also can be put under a sample corpus as classification given by Tognini Bonelli (2010) as it cannot represent the entire language of the all periods of time and all genres.

2.3.4 Benefits of Corpus Studies

McCarthy and O'Keeffe (2010) point out a number of benefits of corpus linguistics as follows:

1) Language teaching and learning

The corpus becomes the center of knowledge, the students take on the role of questioner and the teacher is challenged to hand over control and facilitate learning. Students' writing can be improved through the use of error tagging and follow-up student corpus investigation (Chambers & O'Sullivan, 2004). The development of learner corpora enables the learners to create and work with their own language. The studies in this area are such as by Vannestal and Lindquist (2007) entitled "Learning English Grammar with a Corpus: Experimenting with Concordancing in a University Grammar Course"; by Belz and Vytkina (2008) entitled "The Pedagogical Mediation of a Developmental Learner Corpus for Classroom-Based Language Instruction"; and by Breyer (2009) entitled "Learning and Teaching with Corpora: Reflections by Student Teachers". The use of corpora also covers testing and teacher education. For testing, corpora can facilitate the issues of key standards and rating. In this area there are research studies by scholars such as Barker (2001) and Hasselgren (2002). Corpora can also be useful for teachers in terms of practice and professional development. A study that touches on this issue is by McCarthy (2008) entitled "Accessing and interpreting corpus information in the teacher education context".

2) Discourse analysis

Corpus linguistics enables the analysis of above-sentence discourse such as Conversation Analysis (CA), Discourse Analysis (DA), and Critical Discourse Analysis (CDA). With the application of a corpus in their study using available features such as wordlists, concordances, and key word searches, the researchers can draw on theories and applications of either CA, DA, or CDA. For example, a corpus is used to compare the turn sequence of an opening of a telephone call to a radio

station. Research studies in this area are such as by Conrad (2002), Baker and McEnery (2005), Mautner (2007), and Kim (2014).

3) Literary studies and translation studies

Corpora can be used to compare two volumes of poetry as well as the texts of movie scripts. In this manner, the researcher can assign semantic categories to key words in the corpora which are being compared to explore their stylistics. Studies in this area are such as by Bettina (2009) and Biber (2011). For the area of translation, corpora enable the comparison of patterns across languages by comparing source and target texts. Research studies in this area are such as by Baker (1999), Olohan and Baker (2000), and Kenny (2006).

4) Forensic linguistics

Corpus linguistics can be a useful tool in authenticating authorship. Cotterill (2010) notes that forensic linguists tend to refer to corpus linguistics as a tool or a resource since there is no other methods of analysis can guarantee the identification or elimination of author. Research studies in this area are such as by Coulthard (1994) and Grant (2010).

5) Pragmatics

The application of corpora in the area of pragmatics is quite slow since it usually employs data from role-plays, interviews, and Discourse Completion Tasks (DCTs). However, there has been successful use of corpora in insightful pragmatic studies. These studies are such as the investigation of individual pragmatic features such as pragmatic markers, hedging and politeness, irony, and humor. Research studies in this area are such as by Flowerdew (2002, 2004), Massimo (2011), and Vaughan and Clancy (2013).

6) Sociolinguistics, media discourses and political discourse

For sociolinguistics, it is not sufficient to work with a purely textual transcript. Thus, speaker's information such as age, gender, educational background, and geographical origin become integral parts of the analytical process. Research studies in this area are such as by Barbieri (2005, 2007, 2009), Fraser (2009), Blackwell and Fox Tree (2012), Kendall (2011), and Kendall and van Herk (2011). Media discourse, similar to CDA, tends to expose the ideologies that inform and underlie texts. Corpora can be a useful tool for working on this type of discourse. This area has been studied widely such as by Grundmann and Krishnamurthy (2010), Jaworska (2012). Tobina and Lyddy (2014), and Ju and Yeon (2015).

2.3.5 Concordance Software

Concordance software or a concordance is a computer program which is an important tool for analyzing corpora. The concordancer is capable of identifying all the instances of a target word/string in the corpus being used and show the lines of text in which they occur. The output of the target word/string, also called *node* or *keyword* is displayed at the center of the screen to make it easier to read and analyze. The output which displays the results on the screen is called a *concordance*.

According to Stubbs (1995), the concordances reveal that words may habitually collocate with other words from a definable semantic set, which may carry either positive or negative connotations. The example is the word '*cause*' would typically co-occur with unpleasant things such as *problems*, *trouble*, *damage*, *death*, *pain*, and *disease*. The word '*provide*', on the other hand, usually goes with positive words such as *facilities*, *information*, *services*, *aid*, *assistance*, and *money*. This phenomenon is described by the term '*collocational prosody*'.

There are several concordance tools available. The following are some of them.

1) WordSmith Tools: It is a paid software package developed by the British linguist Mike Scott at the University of Liverpool. It was first released as version 1.0 in 1996. The current version 6.0 (Scott, 2012) was released in 2012 by the Lexical Analysis Software Limited. It is an integrated suite of programs for looking at how words behave in texts. The core areas of the software package include three modules: Concord, WordLisit, and KeyWords. The WordList tool can generate a list of all the words or word-clusters in a text, set out in alphabetical or frequency order. The concordancer, Concord, can extract any word or phrase in context and reveal what sort of company it keeps. KeyWords enable users to find the key words in a text. Each of the modules is offering a number of features available to certain other features of the analyzed text corpus. WordSmith is found to be an effective tool for research studies in the corpus studies. To mention some of them here are such as by Ruth, (2007), Ahour, Rasoulizadeh, and Behnam (2013), and Shou, Wang, and Wang (2014).

2) AntConc: It is a computer-based freeware corpus analysis toolkit for concordancing and text analysis developed by Laurence Anthony, a professor in the Faculty of Science and Engineering at Waseda University, Japan. The current version is 3.4.4 (Anthony, 2014). AntConc can operate on Windows, Macintosh OS X, and Linux. The software includes seven tools namely Concordance Tool, Concordance Plot Tool, File View Tool, Cluster/N-Grams, Collocates, Word List, and Keyword List.

Concordance Tool enables users to search results in a 'KWIC' (Key word in context) format. This feature allows users to explore how the search words are used in the language. *Concordance Plot Tool* provides search results plotted as a 'barcode' format. This provides the search results appear in target texts. *File View Tool* enables users to see the text of individual files. This allows more detailed investigations of the results generated in other tools of AntConc. *Clusters/N-Grams* displays clusters based on the search condition. In effect it summarizes the results generated in the Concordance Tool or Concordance Plot Tool. The *N-Grams Tool*, on the other hand, scans the entire corpus for '*N*' (e.g. 1 word, 2 words, …) length clusters. This allows users to find common expressions in a corpus. *Collocates* shows the collocates of a search term. This enables users to investigate non-sequential patterns in language. *Word List* counts all the words in the corpus and presents them in an ordered list. This helps users to easily identify which words are the most frequent in a corpus. *Keyword List* shows which words are unusually frequent (or infrequent) in the corpus in comparison with the words in a reference corpus. This allows users to examine characteristic words in the corpus, for example, as part of a genre or ESP study. AntCont has been widely used as a tool in corpus studies. The examples of its use are such as by Jablonkai (2009), Bal (2010), Yang (2012), and Hou (2014)

3) MonoConc Pro: This is another commercially available concordance program developed by Michael Barlow with the current version 2.2 (Barlow, 2000). It is used in the analysis of English or texts in other languages such as Spanish, French, Japanese, Chinese, etc. for linguistic or language teaching and language learning purposes. As well as providing KWIC concordance results, the software also produces wordlists and collocation information. The program is easy to use and comes with a

range of powerful features such as Context Search, Regular Expression search, Part-of-Speech Tag Search, Collocations, and Corpus Comparison. Research studies that employed MonoConc Pro are such as by Koo (2006) and Steuber (2011).

4) NooJ: This open-source developed by Max Silberztein (2003), is freely available at www.nooj4nlp.net. NooJ provides users with regular grammars, context-free grammars, context-sensitive grammars, unrestricted grammars as well as their graphical equivalent (finite-state, recursive, and contextual graphs) to facilitate the description of each phenomenon. NooJ's multi-layer approach allows linguists to gather elementary descriptions and describe phenomena that cross linguistic levels. NooJ can even allow users to apply sophisticated linguistic queries to large corpora in real time, in order to construct indices and concordances, annotate texts automatically, and perform semantic and statistical analyses. NooJ was used in research studies by scholars such as Cheikhrouhou (2014), Salza (2014), and Sidhom and Lambert (2014).

5) Sketch Engine: This concordance tools developed by Adam Kilgarriff (2015) is a commercially available program for anyone wanting to research how words behave such as how a particular word occur in a sentence or with which other words it usually co-occurs. It is a corpus software interface which works online and offers many corpora in many languages. It is a Corpus Query System incorporating word sketches, one-page, automatic, corpus-derived summaries of a word's grammatical and collocational behavior. Sketch Engine was used in research studies by scholars such as Deroey (2011), Gerow and Keane (2011), and Luzón-Marco (2011).

2.3.5.1 Previous Studies Conducted Using Concordance Software

A number of studies have been conducted using concordance software as the main tool. These studies are such as by Yang (2012); Ackermann and Chen (2013); Molavi, Koosha, and Hosseini (2014); and Gulec and Arif Gulec (2015). The details of their studies are presented as follows:

Yang (2012) investigated gender representation in an English textbook series used in Hong Kong schools. The corpus software AntConc was used to analyze the collocations of gendered terms *He/he, She/she, Man/man, Woman/woman, women, Boy/boy, Boys/boys, Girl,* and *Girls/girls* in the textbook series to examine gender stereotype. The results revealed that females were no longer regarded as delicate or weak but stronger than males. In occupational roles, females were no longer portrayed only as housewives in the family. On the other hand, the stereotyped images of males wearing shorts, jeans or shirts and females putting on skirts or dresses still exist. In addition, it was found that the male terms have more collocates and negative adjectives were only used to describe males and males were never described in terms of their physical attractiveness. The researcher gave the reason for the application of the AntConc because it is a freeware and is simple and easy to use.

MonoConc Pro 2.2 and human judgement were used as the analysis tool in a study conducted by Ackermann and Chen (2013) in developing the Academic Collocation List (ACL) from the written curricular component of the Pearson International Corpus of Academic English (PICAE) which comprises over 25 million words. The results yielded 2,468 most frequent and pedagogically relevant entries of lexical collocations. The ACL can help learners increase their collocational competence and thus their proficiency in academic English. The ACL can also

support EAP teachers in their lesson planning and provide a research tool for investigating academic language development.

The AntConc 3.2.1 concordancer program coupled with the Open American National Corpus (OANC) available online were employed to compare lexical collocations from textbook to their real use by native-speakers. This study was conducted by Molavi et al. (2014) to examine the distribution of lexical collocations in three selected series of general English textbooks through analyzing, face to face and telephone conversation scripts collected from intermediate and upper-intermediate levels. The findings suggested that a special attention has been paid to collocations noun-verb and adjective-noun combinations while the frequency of collocations in series could not be affective on learners' collocations learning. On the other hand, comparing textbooks collocations to reference corpus (OANC) showed choice of collocations in these series did not have big refers to real use of language by native speakers.

The software applications Concordance, AntConc, and Wordsmith were used in a study by Gulec and Arif Gulec (2015) to investigate verb-noun lexical collocations across the health, physical and social sciences in the written academic journals and analyzed these lexical collocations through the frequency and chi-square analysis. The study aimed to find similarities and differences between the verbs with their collocations. The results showed that there were more similarities and relationship between the health and physical sciences, while the social sciences indicated a significant difference compared to the other two. The study found 165 common verbs used across the three sciences. 12 verbs among the 165 verbs were found to be candidates verb-noun lexical collocations as prototypes. To write better,

the researchers suggested that learners need to be aware of the collocates of the verbs they know.

Although most researchers did not give reasons for selecting particular concordance software applied in their studies, it is assumable that with their compatible capacity, functions and features available, the researchers can choose the software that they see appropriate to their research. For the present study, AntConc will be used as the instrument in the present study as it is readily available online with free of charge. Most importantly the program is equipped with all the features needed for the study namely *Word List*, *Keyword List*, and *Collocates*. According to Kezhen (2015), AntConc has "an easy-to-use, intuitive graphical user interface and offers a powerful concordance, word and keyword frequency generator, tools for clustering and lexical bundle analysis, and a word distribution plot" (p. 62). The program also has been constantly developed by the developer with the current version 3.4.4 and version 3.5.0 is under the development.

2.3.6 Corpus-based Lexical Analysis

The corpus-based lexical analysis is the analysis of lexical words within a corpus with the help of a concordance tool. It involves the identification and analysis of "association patterns" in language use. The association patterns refer to the systematic arrangements in which linguistic features are organized in association with other linguistic and non-linguistic features (Biber, Conrad, & Reppen, 2006). The linguistic associations consist of two main categories: lexical associations and grammatical associations. The former involves the investigation of the way the linguistic feature is systematically associated with particular words. The latter

involves the investigation on how the linguistic feature is systematically associated with grammatical features in the immediate context.

There are four essential characteristics of corpus-based analysis given by Biber et al. (2006), which influence a scope and reliability of analysis. The first of these characteristics is that it is empirical. This involves the analyzing of the actual patterns of use in natural texts. The second characteristic is that it utilizes a large and principled collection of natural texts, which is known as a "corpus," as the basis for analysis. The third characteristic is that a computer program plays a major role in the analysis, using both automatic and interactive techniques. The last characteristic of corpus-based analysis is that both quantitative and qualitative analytical techniques are used.

In corpus-based lexical analysis, there are basic descriptive statistics such as frequency, token and type, word list, keyword list, and keyness. Frequency refers to the number of occurrences of each element that can be counted. For example, the number of '*token*' is the frequency count from each individual word occurs in the entire corpus; while '*type*' refers to the number of each type (similar or repeated) of the words that occur in the corpus. Word list refers to the list of words in a corpus which may be displayed according to the rank of frequency or alphabetically. Keyword list refers to words which are unusually frequent (or infrequent) in the corpus in comparison with the words in a reference corpus. The keywords listed from a particular corpus are likely to be good representatives of it. Keyness refers to a statistical value of the keywords measured on a basis of the value of log-likelihood or Chi-square statistics calculated from the frequency in the study corpus compared with the frequency in the reference corpus (Scott, 2012).

Another important aspect to take into consideration in corpus-based lexical analysis, particularly with the study of collocations, is statistical measures of lexical association. Hunston (2002) and Biber et al. (2006) suggest two statistical measures: the Mutual Information (MI) score and the t-score.

MI score: the MI score or MI index, according to Biber et al. (2006), indicates the strength of association between two words. The measurement is based on the likelihood of two words co-occurring within a specific span of words. The MI score is calculated by comparing the probability of ability of observing the two words together with the probability of observing each word independently based on the words' frequency. A score 0 indicates no relationship between the words. The higher the score, the stronger the association strength between the words. Normally collocate pairs with high frequency tend to have a high MI score. However, this is not always the case.

T-score: according to Biber et al. (2006), it indicates how the collocate pairs are used differently instead of how they associate. T-score measures the degrees of words that are more likely to co-occur with one particular word rather than with another. The differences in the collocates can indicate the differences in the use of the words. To find out the t-score, the pair words have to be identified. Then the frequencies of the collocates for each word are analyzed and compared to each other. Collocates with large positive t-scores tend to appear with the first word, and collocates with a large negative t-scores tend to appear with the second word.

Hunston (2002) notes the important differences between MI and T-scores as follows:

1) MI score is a measure of strength of collocation, while t-score is a measure of certainty of collocation.

2) The value of an MI score does not depend on the corpus size, while the t-score does. The larger the corpus size, the more significant is for the t-score value.

3) MI scores can be compared across corpora, while t-scores cannot since the size of the corpus can affect the t-score.

4) The top collocates measured by t-score are likely to comprise information about the grammatical behavior of a word, while MI score tends to give information about lexical behavior of a word.

5) The collocates with highest t-scores tend to be frequent words that collocate with various other words, while the collocates with the highest MI scores tend to be less frequent words with restricted words to co-occur with.

2.3.6.1 Previous Studies on Corpus-Based Lexical Analysis

A number of studies have been conducted applying corpus-based lexical analysis. These studies are such as those conducted by Biber, Gray, and Poonpon (2011); Yang (2012); Ackermann and Chen (2013); and Parkinson (2015). The details of their studies are presented as follows:

Biber et al. (2011) conducted a critical evaluation of T-units and clausal subordination as measures of writing development arguing that these measures have not proven to be effective discriminators of language proficiency differences. They investigated 28 grammatical features in research articles in comparison with that of in real conversations. The surprising results showed that most clausal subordination measures were actually more common in conversation than academic writing. In contrast, fundamentally different kinds of grammatical complexity were common in

academic writing: complex noun phrase constituents (rather than clause constituents) and complex phrases (rather than clauses). Based on the findings, the authors hypothesize a sequence of developmental stages for student writing, proposing a radically new approach for the study of complexity in student writing development.

A study by Yang (2012), as mentioned in the previous section (2.3.5) above, investigated gender representation in an English textbook series used in Hong Kong schools. The corpus software, AntConc, was used to analyze the collocations of gendered terms *He/he, She/she, Man/man, Woman/woman, women, Boy/boy, Boys/boys, Girl,* and *Girls/girls* in the textbook series to examine gender stereotype. In terms of the association strength between words under investigation, MI scores were used as the measurement. The reason was in accordance with Hunston's (2002) view that MI scores are more appropriate as they give information about words' lexical behavior. Additionally, the size of the corpus does not affect the MI values. The results of the study revealed that females were no longer regarded as delicate or weak but stronger than males. In occupational roles, females were no longer portrayed only as housewives in the family. On the other hand, the stereotyped images of males wearing shorts, jeans or shirts and females putting on skirts or dresses still exist. In addition, it was found that the male terms have more collocates and negative adjectives were only used to describe males and males were never described in terms of their physical attractiveness. The researcher gave the reason for the application of the AntConc because it is a freeware and is simple and easy to use.

In a study by Ackermann and Chen (2013) in which they developed the Academic Collocation List (ACL) from the written curricular component of the Pearson International Corpus of Academic English (PICAE) which comprises over 25

million words, the computational analysis using MonoConc Pro 2.2 and human judgement were applied. In the analysis procedures of the study, MI score and t-score as well as the frequency were taken into account with the MI score of at least 3 and the t-score of at least 2. The results yielded 2,468 most frequent and pedagogically relevant entries of lexical collocations. The ACL can help learners increase their collocational competence and thus their proficiency in academic English. The ACL can also support EAP teachers in their lesson planning and provide a research tool for investigating academic language development.

In a study conducted by Parkinson (2015) on the use of noun-noun collocations by learners in their academic writing, three sub-corpora were created: essays written by L1 speakers of Mandarin, Spanish, and Tswana. Production of noun-noun phrases in written English by L1 Mandarin writers (a language that permits noun-noun phrases) was compared to writing by L1 Spanish writers (a language that does not allow noun-noun phrases). The MI score of 3 or more was applied in the analysis process as a significant collocation threshold. The noun-noun combinations with an MI score less than three were referred to as noun-noun phrases. The findings showed that learners whose L1 permits noun-noun phrases produced significantly more of them in English than learners whose L1 does not.

From the above mentioned review concerning corpus-based lexical analysis, apart from the four essential characteristics described by Biber et al. (2006), measurements for association strength are the indispensable. The MI and t-score are being the most used measurements in this respect with different purposes. In that the MI score tends to closely relate to lexical behavior of words, while the t-score is likely to relate to the grammatical behavior of words being investigated.

2.3.7 Identifying Collocations from a Corpus

According to Nattinger and DeCarrico (1992), a corpus consists of authentic material, full of unexpected and diverse constructions which are often treated as too peripheral or ill-formed to be of much interest of theoretical grammars. Thus, they often require unconventional categories of description. A concordancer, a computer program, can scan all these data for collocations.

Unlike syntax which deals with general classes of words and their combinations, collocations, on the other hand, describe specific lexical items and the frequency with which these items occur with other lexical items. A collocational unit consists of a '*node*' that occurs with a '*span*' of words on either side. The span consists of particular word classes filled by specific lexical items. If the node word occurs with a span of particular words at a frequency greater than chance would expect, the result is a collocation. The more certain words in the span are to co-occur with the node means the more fixed and idiomatic the collocation. If collocations become less fixed or more variations are allowed, this would lead to predictability lessens and meaning increases.

The identification of collocations in a corpus starts with listing all the node words with the respective spans. The next step is to delete words in the spans which occur only once, function words, and repeated but rare words that appear only because of the subject of the text. The possible collocations are those remaining words. The deleting of articles and other function words makes the resulting collocations to be only those of combinations of the four major syntactic classes such as *N+N, Adj+N, Adv+Adj, Adv+V.*

Church and Hanks (1989) introduce an objective measure for identifying collocation based on the mutual information (MI). By examining the 'association ratio' based on 'mutual information' gained from the probability of observing X and Y together, with the probability of observing X and Y separately. If there is a genuine association between them, the joint probability will be greater than chance. Word probabilities are calculated by counting the number of Xs and Ys in a corpus and normalizing by the size of the corpus. Joint probabilities can be calculated by counting the number of times that X is followed by Y within a span of W words and normalizing by N.

In terms of measuring the significance of collocation, McEnery et al. (2006) and Cheng (2012) suggest that apart from z-score and log likelihood, t-score and mutual information (MI) value are the two most commonly used indicators of association strength. Most concordancing software programs are also equipped with the feature for calculating the value of these two measures. Significance for the t-score is ≥ 2 and for the MI value is ≥ 3. The difference between these two measures, according to Stubbs (1995), is that a collocate list based on the t-score is more likely to include function words than that of based on MI value. The collocations generated based on the MI value, on the other hand, is more likely to consist of lexical collocates with less function or grammatical words.

As the present study is set to examine lexical collocations in a sample corpus, MI value will be applied as a measure for the significance of the co-occurring pairs. The keywords generated from the sample corpus of each part of speech will be treated as '*nodes*'. Each '*node*' will be examined for its collocates with three-word '*span*' on

its right side for each type of combinations. The MI value of ≥ 3 and the frequency of occurrence of ≥ 10 will be acceptable collocational strength for this study.

2.3.8 Corpus-based Language Teaching and Learning

In EFL/ESL teaching and learning, corpus linguistics with the assistance of computer technology has been increasingly interested by scholars. This part presents information concerns corpus-based language teaching and learning which includes the use of corpora in language teaching and learning and the advantages as well as disadvantages of using them in various aspects namely in teaching English for Academic Purposes (EAP), teaching and learning grammar, writing, listening, and reading.

The use of corpora, corpus-analytic tools, and corpus evidence in English language teaching and learning have been increased for the last two decades (McEnery et al., 2006; Cheng, 2010). Corpus-based language teaching is advocated by Sinclair (2004) as a new revolution in language teaching. Fligelstone (1993) indicates that there are three aims of corpus-based linguistics in teaching: teaching about, teaching to exploit, and exploiting to teach. Teaching about covers the principles and theory concerning the use of corpora. Teaching to exploit focuses on the practical, methodological aspects of corpora. Exploiting to teach involves using corpora to derive or drive teaching materials. The fourth aim is added by Renouf (1997), which is teaching to establish resources. This involves the learners in data collection, corpus design, and corpus compilation. However, with this approach of teaching, according to McEnery et al. (2006), the traditional 'three Ps' (Presentation, Practice, and Production) may no longer be applicable. The better suited approach may be the 'three Is' (Illustration, Interaction, and Induction). Illustration refers to

looking at real data. Interaction refers to discussing and sharing opinions and observations. Induction refers to making one's own rule for a particular feature, which will be refined and polished as the amount and size of the data increase.

In terms of the advantages of using corpora in language teaching and learning, according to Gavioli and Aston (2001), corpora can capture reality and are able to provide valid models for learners as they represent authentic language. They are useful to test claims based purely on intuition and motive the decisions for teaching particular linguistic features. Corpora are also a useful tool for engaging learners in the interpretive process to create models of their own (Leech, 1986). As the learners are able to access the corpus data, they become more active learners. This aspect of corpus in language learning is advocated by Johns' (1991) 'data-driven learning' or DDL. According to Samburskiy (2014), a corpus reveals register variation of a language and a complex relationship between lexicon and grammar. It also allows learners to investigate the frequency of formulaic lexical bundles in any register. However, some disadvantages are raised from the applications of corpora in the learning and teaching language. Liu (2011) found that it took times for his students to get used to the use of the corpus and how to operate it. Samburskiy (2014) found similar problem with his students. He also found that in a search beyond specific words or phrases, students have to learn special commands.

English for Academic Purposes (EAP), according to Coxhead (2010), also has benefited a great deal from corpus linguistics. Corpus tools show students the frequency of particular features of the language. As EAP is for students whose first language is not English, it is very important to them to know more about academic language in use. This is because its distinctive and highly routinized natures that can

be problematic even for native writers, let alone non-native writers (Gilquin, Granger, & Paquot, 2007). There are a number of research studies on incorporating corpora in EAP such as by Lee and Swales (2006) and Charles (2012, 2014) as shown in Table 2.8 below.

Table 2.8 Corpus-based studies in EAP

Researcher(s)	The study	Findings
Lee & Swales (2006)	Participants were given access to specialized corpora of academic writing and speaking. Participants compiled two written corpora: own writing and expert writing to make comparison.	The participants developed more control over their own learning. The use of corpora to compare the writings helped raise the awareness of learners on how to write better.
Charles (2012)	50 students constructed and examined their own individual, discipline-specific corpora. Questionnaires were used to collect the data.	Over 90% of students found it easy to build their own corpora. Most of them were enthusiastic about working with their own corpora and agreed that their corpus helped them improve their writing and intended to use it in the future.
Charles (2014)	40 international graduate students in an EAP course built and examined their own corpora of research articles in their field. One year after the course, they completed an email questionnaire.	70% of the respondents had used their corpus with 38% used regularly (once per week or more). Most users consulted the corpus for checking grammar and lexis and 93% of them considered that corpus use had improved their academic writing.

In terms of teaching and learning grammar, Biber and Conrad (2010) provide three reasons why corpus-based teaching and learning of grammar is encouraged. The first reason concerns frequency information. This means that a corpus can reveal grammatical features that are common or rare. An example corpus-based research found progressive aspect is more common in conversation than in other written

registers. The second reason concerns the associations between grammar and words. This means that corpus-based research has found that not every word is equally likely to occur in a given grammatical structure. The third reason is a corpus enables register comparisons. It is found that strong patterns of use in one register do not necessarily mean to occur similarly in other registers. For example, in fictions and newspapers, the verb *say* is more frequent than any other lexical verb; in conversation, the verbs *go* and *know* are as frequent as *say*; while in academic writing, the only especially frequent verb is *BE*.

Hughes (2010) has examined the use of corpus in grammar teaching. She claims that using corpus-based materials enables teachers to highlight the differences between assumptions about language structure in the abstract and what is found in the real use. There are also commercial grammar books based on corpus data available. However, with the more technological advancement, the teaching and learning in classroom in general is likely to be more teacher-led and technology-led than depending on commercially available textbooks. This also implies the increasing role of corpus-based in language teaching and learning. Research studies in this area are such as by Vannestål and Lindquist (2007), Liu (2011), Phoocharoensil (2012), Hanafiyeh and Keshi (2013), and Rapti (2013) as shown in Table 2.9 below.

Table 2.9 Corpus-based studies in teaching and learning grammar

Researcher(s)	The study	Findings
Vannestål & Lindquist (2007)	Students worked with problem-solving assignments that involved formulating their own grammar rules based on the examples they found in the corpus. A system of peer teaching was applied.	Using corpora with students requires a large amount of introduction and support as some students appreciated working with corpora, while others, especially weak students, found it difficult or boring.

Liu (2011)	The use of corpora for problem-based learning/teaching of lexicogrammar in a college English grammar course was investigated through students' individual and group corpus research projects, reflection papers on corpus use, and responses to a post-study survey.	Four themes found: (1) critical understanding about lexicogrammatical and broader language use issues, (2) awareness of the dynamic nature of language, (3) appreciation for the context/register-appropriate use of lexicogrammar, and (4) grasping of the nuances of lexicogrammatical usages.
Phoocharoensil (2012)	17 Thai graduate students' attitudes towards corpus-based grammar teaching were explored through a questionnaire and an interview.	The participants mainly perceived the benefits of using corpus data to learn grammar. Most of them had a very positive attitude towards this concordance-based information.
Hanafiyeh & Keshi (2013)	60 students were selected and randomly assigned to an experimental group of concordance (n=30) and a control group of thesaurus (n=30). Seven writing tasks were assigned using either the concordance or the thesaurus to help their writing.	There were recognizable differences in the EFL writing quality between the groups. There were significant differences in that the concordance group gained more grammatical knowledge than the thesaurus group (p< .05).
Rapti (2013)	14 young learners in Greece were studied to investigate the impact of DDL on motivation and the learning of grammar using concordance-based tasks.	Some students were motivated to study grammar after the completion of the study. The majority of the participants had acknowledged the contribution and potential of corpora but also pointed out the difficulties when involved in DDL.

In teaching and learning of writing, Flowerdew (2010) explores the use of corpora for enhancing students' writing. Apart from the ability of the concordance output to show grammatical and lexical features, she claims that the even more useful function of corpora in writing is they can reveal phraseological patterning such as collocations, colligations, and semantic preferences and prosodies. She states further

that the use of corpora is an ideal tool for helping learners master those phraseological patterns as they are not easily found in either dictionaries or grammars. Research studies in this issue has been conducted by scholars such as Gaskell and Cobb (2004), Yoon and Hirvela (2004), Hegelheimer (2006), and Friginal (2013), as shown in Table 2.10 below.

Table 2.10 Corpus-based studies in teaching and learning of writing

Researcher(s)	The study	Findings
Gaskell & Cobb (2004)	20 adult Chinese EFL learners assigned 10 writing assignments. Feedbacks were given with online concordance links for five typical errors. They were required to revise the text for final submission.	The participants felt their English writing skills had improved. Eight of them would continue to use concordancing as a learning tool in future.
Yoon & Hirvela (2004)	23 ESL writing students participated in the study to explore their corpus use behavior and their perceptions of the strengths and weaknesses of corpora as a second language writing tool.	Overall, the students perceived the corpus approach as beneficial to the development of L2 writing skill and increased confidence toward L2 writing.
Hegelheimer (2006)	The use of an online resource to improve advanced- level ESL learners' writing by increasing their grammatical awareness and ability to correct their grammatical errors.	Learners were satisfied with the system applied. They also exhibited greater awareness of grammar as well as of their own mistakes.
Friginal (2013)	An exploratory study investigated the use of corpora to develop the research report writing skills of college-level students.	Corpus instruction contributed positively to the patterning of the frequencies and distributional data of linking adverbials, reporting verbs, and verb tenses in the students' research reports relative to the professional corpus.

In terms of speaking and listening, Walsh (2010) states that a collection of students' spoken work can be useful to help them understand their own problems and improve their speaking skills. Other corpora of spoken language available such as MICASE are also helpful as they represent authentic use of the language. The spoken corpora show patterns and structures which are commonly found in the genre. The corpora of spoken language can also help the learners to recognize fixed expressions which are commonly used. Research studies into this area are such as by Izumi, Uchimoto, and Isahara (2004); Furui, Nakamura, Ichiba, and Iwano (2005); and Rashtchi and Afzali (2011) as shown in Table 2.11 below.

Table 2.11 Corpus-based studies in teaching and learning of speaking and listening

Researcher(s)	The study	Findings
Izumi et al. (2004)	A compilation of a large-scale speech corpus called "The SST Corpus", which based entirely upon the audio-recordings of an English oral proficiency interview test called the Standard Speaking Test (SST). This corpus can be exploited for automatic detection of learners' errors with a machine learning technique.	By using the corpus in the experiment, the recall of article errors was 35% and the precision was 48%. By adding corrected sentences and artificially-made errors, recall and precision improved to 43% and 68% respectively.
Furui et al. (2005)	The analysis and recognition of spontaneous speech using a large-scale spontaneous speech database "Corpus of Spontaneous Japanese (CSJ)".	Recognition accuracy significantly increased as a function of the size of acoustic as well as language model training data and the improvement levels off at approximately 7M words of training data.

Rashtchi & Afzali (2011)	45 university students participated in a listening course employing corpus-based materials to explore whether awareness of spoken grammar features could affect learners' comprehension of real-life conversations.	The students in the experimental group who had exposed to spoken grammar through awareness raising tasks, comprehended everyday conversations much better. They also had highly positive views of spoken grammar.

In terms of reading skills, as a corpus comprises authentic language from the real use, corpus-based language learning is also likely to be beneficial for the improving of the reading skills. Walsh (2010) emphasizes that with a large resource offered by a corpus, teachers can select texts to suit particular groups of learners according to their level or content. As a result, the learning potential of materials is maximized. In a study conducted by Wang, Zheng, and Cai (2015) with students in advanced English reading course. They found that students' textual analysis skills and reading comprehension ability have improved through the student's active involvement in the construction of a textbook-related corpus and the use of corpus analysis methods in the course. Apart from the study by Wang et al. (2015), there is also a study earlier by Kırkgöz (2006) as shown in Table 2.12 below.

Table 2.12 Corpus-based studies in teaching and learning of reading

Researcher(s)	The study	Findings
Kırkgöz (2006)	The compilation of a corpus of academic texts from the disciplines of economics and business administration, as the basis for designing a lexical component of the EAP reading course and developing teaching materials.	Students became familiarized with as well as gained more lexical competence in subject-specific lexical items. The course also enabled the students to have easy access to the written academic texts in their disciplines.

Wang et al. (2015)	The experimental study with Chinese EFL students enrolled in Advanced English Reading course to investigate learners' textual analysis skills through compiling their own corpora and conducting textual analysis using a corpus tool.	In the experimental group, students' textual analysis skills and reading comprehension ability have improved significantly.

From the above mentioned concerning corpus-based language teaching and learning, it is evident that there are a wide range of corpus-based applications in the field which cover various aspects of EFL/ESL teaching and learning. The present study, where a corpus analysis tool will be used to identify lexical collocations in the Sample Corpus of Nursing Research Articles (SCNRA), should also contribute to the teaching and learning of English particularly the lexical collocations in the field of nursing. This could also further be useful as a building block for other four skills needed for effective communication.

2.4 English for Specific Purposes (ESP)

According to Robinson (1991), ESP is normally goal directed as the learners need English for their study or work not because they want to learn the language. Therefore, ESP course is arranged based on a needs analysis to meet learners' use of the language. As learners of ESP tend to be adults with some professional knowledge, the course is organized to help them communicate with others involved in their work situations.

Dudley-Evans and St John (1998) indicate that ESP puts the emphasis of English language teaching on practical outcomes. The main purpose of ESP is to enable the learners to communicate effectively in their specific situations in which

they have to be involved in either in the daily life or professional settings. They give the definition of ESP based on its two characteristics: absolute characteristics and variable characteristics as follows:

ESP defined based on absolute characteristics:

1) ESP is designed to meet specific needs of the learners;

2) ESP makes use of the underlying methodology and activities of the disciplines it serves;

3) ESP is centred on the language (grammar, lexis, register), skills, discourse and genres appropriate to these activities.

ESP defined based on variable characteristics:

1) ESP may be related to or designed for specific disciplines;

2) ESP may use, in specific teaching situations, a different methodology from that of general English;

3) ESP is likely to be designed for adult learners, either at a tertiary level institution or in a professional work situation. However, it could also be used for learners at secondary school level;

4) ESP is generally designed for intermediate or advanced students. Most ESP courses assume basic knowledge of the language system, but it can be used with beginners.

2.4.1 Vocabulary teaching in ESP

According to Dudley-Evans & St John (1998), point out that the teaching vocabulary of ESP follows similar principles as applied in that of general English which concerns two groups of vocabulary: vocabulary needed for comprehension and vocabulary for production. For comprehension, learners need to be able to deduce the

meaning of vocabulary from the context and the structure of the actual word. For production, learners need to find ways to store vocabulary in their mental lexicon for use when they need them. They suggest three ways to teach vocabulary in ESP in the way that should help facilitate learners' cognitive process and remember the vocabulary, which are:

1) Situational, semantic and metaphor sets

The retrieval of a vocabulary item from memory is aided by the grouping of words according to their meaning. This may be according to topic (situational sets). For example, the word 'library' associates with words such as 'book', 'shelf, 'borrow', 'read', and so on. Or may be according to chains of association (semantic sets) so that synonyms, antonyms, superordinate and subordinate terms can be taught. Teaching of metaphor also can activate learners' cognitive process.

2) Collocation and the use of corpora

Corpora enable learners to witness not only single vocabulary, but also words that frequently co-occur with other words as well as the contexts in which they occur. This is beneficial to the learners as they should understand the words better than when looking at the words without the contexts.

3) Lexical phrases

Learning vocabulary in the form of lexical phrases or chuck is another useful way help learners to better remember the words. By identifying frequently occurred phrases and their use and let learners practice the use of those chunks, the learners should improve their proficiency quickly.

2.4.2 Corpus-based Instruction and ESP

Corpus-based instruction, according to Jabbour (2001), is also referred to as corpus-based linguistics. It is the study of genre texts for the production of materials that fit a specific group of learners. It provides a rich learning environment to the learners. Corpus-based instruction is useful in three ways. Firstly, a large number of texts can be examined as one unit under the same conditions, which make it possible for generalizations. Secondly, it enables to look at language in a range of contexts, from single words to phrases, and larger units. Finally, a corpus enables to determine not only the immediate environment of a word, but also its larger context.

According to Gavioli (2005), since ESP requires teachers to equip with knowledge of English as well as other specialized disciplines, corpora of specialized texts are useful for isolating and providing indications about key lexical, grammatical or textual issues in dealing with ESP classes. She points out that applying corpus-based instruction in ESP is helpful for teachers in the way that it enables them to teach those items which do not seem adequately dealt with in traditional teaching materials.

A number of studies have been conducted by investigating corpus-based instruction (Farr, 2010; Huang, 2012; Ashouri, Arjmandi, & Rahimi, 2014; and Bardovi-Harlig, Mossman, & Su, 2017). The details of these studies are as follows:

Farr (2010) investigated the ways in which corpora have been incorporated into a language systems module on an MA in ELT program. The survey results uncovered 25 student teachers' perspectives on their experiences of using corpora as well as the potentials and problems foreseen in relation to using the approach in their careers. The findings showed there was generally a positive predisposition towards the use of corpora. These attitudes varied in relation to the projected adaptation in

ELT, and the results also showed that the real teaching scenario often did not permit the ideal of full application.

A study conducted by Huang (2012) examined whether a corpus-based instruction could deepen EFL learners' knowledge of periphrastic causatives: make, cause, and let. The participants were 47 Taiwanese undergraduates from two intact classes. One class as the experimental group received a three-month corpus-based instruction; the other as the control group had no instruction on English causatives. A pre-test was first administered to measure participants' knowledge of periphrastic causatives. Following a data-driven model of illustration – interaction – induction, the researcher as instructor conducted the instruction and took notes on students' performance. After the instruction, a post-test was given to both groups with a questionnaire on learning effects and students' feelings for corpus-based activities distributed to the experimental group. The results revealed that the experimental group improved and outperformed the control group significantly in the post-test. The questionnaire results confirmed that the instruction was effective in increasing students' knowledge of the three causatives. However, the field notes revealed learners' difficulties in using certain causatives. The author suggests that learners should attend to semantic distinctions more than syntactic structures. Clear guidance on data search and data interpretation should also be provided.

Similarly, Ashouri et al. (2014) investigated the impact of corpus-based collocation on EFL learners' collocation learning and awareness with 60 Iranian EFL learners who were chosen randomly based on their scores in an OPT exam. There were two groups, experimental and control ones. The study examined the effects of direct corpus-based collocation instruction on EFL learners' collocation learning. For

15 sessions the control group received single-item vocabulary or, the usual work of their class and the experimental group received lexical collocation instruction as treatment. The same test as post-test was given to the learners when the treatment accomplished, and after that a t-test and kolmogorov-smirnov test between the pre-tests and post-tests were calculated. The results demonstrated the effectiveness of the treatment. The study suggests that direct corpus-based collocation instruction can be a worthy alternative. It demonstrated that the learners, who were in the experimental group, got aware of the existence of collocations, used them and learnt them properly, and they also started to find the collocation of every other word by themselves because the treatment appealed to them.

Bardovi-Harlig et al. (2017) also compared the effect of using corpus-based materials and activities for the instruction of pragmatic routines under two conditions: implementing direct corpus searches by learners during classroom instruction and working with teacher-developed corpus-based materials. The outcome was compared to a repeated-test control group. Pragmatic routines used for agreement, disagreement, and clarification in academic English discussion are targeted. 54 students in seven intact communication classes participated. 43 students received instruction in four 50-minute lessons across two to three weeks applying input from MICASE with noticing and production activities. The corpus-materials group (N = 26) received corpus excerpts and the corpus-search group (N = 17) conducted equivalent searches. The pre- and post-tests were administered through a computer-delivered oral-production task that simulated group discussion and included 30 items: 10 agreement, 10 disagreement, and 10 clarification scenarios. The results showed a significant increase in the oral production of pragmatic routines. The corpus-materials group additionally

showed an increase in the clarity of speech acts. The corpus-search group reported engagement in self-directed searches outside the classroom, captured by a post-test questionnaire.

From the above mentioned studies concerning corpus-based instruction applied in various aspects of EFL/ESL teaching and learning, all of them yield positive and satisfactory results. Therefore, it is convincing that the present study should apply this this method of instruction since the present study is a corpus-based study from the beginning.

2.5 Pre-Experimental Research Design

According to Ary, Jacops, and Sorensen (2010), there are two research designs which are classified as pre-experimental: One-group Pretest-Posttest Design and Static Group Comparison. However, some other scholars such as DePoy and Gitlin (2011) include One-Shot Case study into the category.

One-Shot Case Study, according to DePoy and Gitlin (2011), refers to an experiment where the independent variable is introduced followed by the dependent variable is measured in only one group. This type of design is useful in answering descriptive questions such as "what happened after a phenomenon occurred. The example can be seen in the case of a course enrolled by students where there is no a pre-test administered and at the end of the course the students have to take the examination. In such case, students' scores on the examination can only tell what they learned, but cannot attribute their learning to the course.

One-Group Pretest-Posttest Design, according to Gall, Gall, and Borg (2007) and Ary et al. (2010), involves three steps: (1) administration of a pretest; (2)

implementation of the experimental treatment for participants; and (3) administration of a posttest. The effects of the experimental treatment are determined by comparing the pretest and posttest scores. This research design is appropriate when trying to change a characteristic that is very stable or resistant to change. It is justified when extraneous factors can be estimated with a high degree of certainty or assumed to be minimal or nonexistent. In this research design, history and maturation are two extraneous variables that are not able to control and may affect the treatment. The longer period of time taken between pre-test and post-test could threaten internal validity of this design.

Static Group Comparison, according to Ary et al. (2010), is a research design where two or more preexisting or intact groups are used and only one group is exposed to the experimental treatment. The flaw of this research design is that the subjects are not randomly assigned to the groups coupled with no pre-test administered. To assess the effects of the treatment, the test scores are compared between the groups. This design, as commented by Ary et al. (2010), is basically worthless since there is no randomization or matching on a pre-test applied. This leads to a doubt in the outcome whether it is a result of the experimental treatment.

From the descriptions of the three types of the pre-experimental research design given above, the second part of the present study complies with One-Group Pretest-Posttest Design.

2.6 Academic Journals and Journal Articles

This section provides information on academic journals and journal articles. Special attention is given to those in the field of nursing. The definitions are given

including the information on journals' impact factor as well as the historical background of nursing journals and characteristics of nursing journal articles.

2.6.1 Definitions of Academic Journals and Journal Articles

An academic journal, according to Svensson and Wood (2007), serves as a communicative interface between scholars in the field of a research discipline. It is also designed as a communicative channel for scholars to reach practitioners such as executives, managers and consultants or vice versa. According to Jerz (1999), an academic journal is a form of a publication which publishes *scholarly*, *peer-reviewed articles* that are *written by experts* in the field. The main function of a journal is to distribute knowledge among scholars in the field and those who might be interested. '*Scholarly*' denotes that each fact or opinion is documented with the exact source for the information from the outside by means of a reference is provided in the agreed manner. The article will probably be long, complex, and possibly difficult for a non-expert to understand right away. '*Peer-reviewed*' referred to selected and approved by a panel of experts. Each academic journal has a peer review board or a panel of experts that decides which submission is acceptable for publication. The review board may send a paper back to the author with suggestions for improvement. '*Written by experts*' means that academic journals typically identify their contributors as professors, graduate students, or others with first-hand experience with the subject matter.

A journal article, according to University of Toronto's libraries (2015), is sometimes called a 'scientific article', a 'peer-reviewed article', or a 'scholarly research article'. Journal articles in a particular field are often referred to as 'the literature'. Journal articles are most often primary research articles. However, they

can also be review articles which have different aims and requirements. Sometimes, an article describes a new tool or method. Without background knowledge in the field, journal articles may be hard to understand; however, the readers do not need to understand an entire article to be able to get valuable information from it. Reading a journal article may also lead to a number of other journal articles on closely related topics. In terms of the structures of a journal article, according to Cargill and O'Connor (2013), the conventional format is in AIMRaD (Abstract, Introduction, Materials and Methods, Results, and Discussion), though there may be some variations.

2.6.2 Nursing Journals

According to Mason, Kennedy, Schorr, and Flanagin (2006), the first journal in the field of nursing was launched in 1953 entitled '*Nursing Research*'. The journal was founded by the AJN Company and the American Association of Collegiate Schools of Nursing and its five sister organizations (the American Nurses Association, National League for Nursing Education, National Organization of Public Health Nursing, National Association of Colored Graduate Nurses, and American Association of Industrial Nurses). '*Nursing Research*' became the gold standard for the profession's research journals and continues a mandate to educate nurses about research through its publication of articles, announcements of research conferences, and opportunities for research funding. In the 1970s, there were more research in the field of nursing came out. Consequently, a number of new nursing journals were launched.

According to Binger (1981), there were 40 new nursing journals launched from the mid-1960s to the end of the 1970s. The titles of these journals reflected the

specialization in the nursing profession such as *Journal of Emergency Nursing*, J*ournal of Gerontological Nursing*, *Journal of Neurosurgical Nursing*, *Cardiovascular Nursing*, J*ournal of Enterostomal Therapy*, and *Nephrology Nurse*. According to Mason et al. (2006), most of the journals were affiliated with a specialty nursing professional society. For nursing societies, journals act as an arena for nurses to publish specialty-specific research and clinical review articles. They also provide forums for communication on related issues among members. On the European side, In England, for example, *Nursing Standard* is the leading general nursing journal as it is the official journal of the Royal College of Nursing (RCN). Other nursing journals are such as the *Journal of Advanced Nursing*, the *Journal of Nursing Scholarship*, and *International Nursing Review*.

From a study conducted by Díaz-Membrives, Farrero-Muñoz, and Lluch-Canut (2012), there were 74 nursing journals with an Impact Factor (IF) in 2009 and increased to 91 in the following year. In 2010, 93.5% were published in English, mostly bimonthly journals (43%) and for specialties, maternity, and paediatrics were the most frequent (25%). Almost three-quarters (72.8%) of the original articles were quantitative studies performed mostly in hospitals (42%) and with patient samples (34.6%). The most frequently studied topics were "evidence-based care" (23.5%), "measuring quality care" (18.52%), and "effectiveness of nursing interventions" (14.81%). Authors were mostly from Europe and United States and the most common workplace was a university.

2.6.3 Nursing Journal Articles

According to Alexander (2011) and Oermann and Hays (2011), the most common types of nursing journal articles are: research articles, quantitative and

qualitative articles, evidence-based practice articles, quality improvement articles, clinical articles, literature reviews, case studies, and nursing narratives and exemplars.

Research articles are reports of original data, findings, and results. They summarize a study, its purpose, methods, and findings. The typical format of this type of research is the IMRD (Introduction, Methods, Results, and Discussion) format.

Quantitative and qualitative articles can be either quantitative or qualitative reports. For the former, they usually follow the IMRD format with the results and discussion parts organized according to the purposes, research questions, or hypotheses of the study. The latter are normally used to report on the investigation on patterns of particular situations or behaviors of particular population.

Evidence-based articles are used to report on the practice and evaluate the effectiveness of new approaches in patient care. In evidence-based practice, nurses identify a clinical question or problem and then search for evidence to answer that question coupled with critically appraise studies and assess the quality of the evidence.

Quality improvement articles report on the problem that led to the need for the study, population, setting, intervention, outcomes of the study, and local conditions. An accurate and complete report is essential for this type of article since its goal is to report the situation and ways to improve it.

Clinical articles are those reports that address topics in clinical practice. They may be written for nurses across specialties or for nurses practicing in a particular clinical area. The writing format varies but normally includes a description of the patient problems and nursing interventions by means of the article is presented.

Literature reviews are articles that summarize and evaluate the previous and recent research on particular topics. This type of articles may serve as a preview for the primary research articles. Accurate and complete citations are important in this type of article.

Case studies are articles that present case reports to provide new information on nursing practice or care of patients with particular health problems through the presentation of an actual case. The articles usually begin with the reason for reporting the case and its significance for nursing practice.

Nursing narratives or Exemplars are recommended for a novice author as they are reports on a basis of a personal account that describes outstanding examples of the actions of individuals in clinical settings that benefit patient care.

The various types of nursing journal articles presented above reflect the vast area of specialties and responsibilities professional nurses have. For the present study, these types of articles will not be taken into account as a criterion for the selection as the emphasis will be put on their format. However, the well balance in terms of the specialties of the journals where the journal articles will be taken from is one of the selection criteria.

2.6.3.1 Previous Studies Conducted with Nursing Journal Articles

A study conducted by Forbes (2009) explored clinical intervention research in nursing journal articles. From the total of 517 articles, 88% (n = 455) were classified as non-clinical intervention and 12% (n = 62) as clinical intervention studies. Characteristics of intervention studies were identified which include the underpinning theory, linkage to previous (pre-clinical) work, evidence of granularity, protocol

clarity (generalizable and parsimonious), the phase of knowledge development, and evidence of safety.

Palese, Coletti, and Dante (2013) conducted a retrospective study to examine publication efficiency of nursing journals through the time occurring between data collection and manuscript publication. The articles published in the selected journals, from 1st January to 31st December 2009 were used based on the 2009 Impact Factor List of Nursing Journals published by the ISI web of Knowledge. 1152 articles were included. From the end of data collection to manuscript publication online/on paper it took an average of 981 days [CI95% 929–1032] (2.5–3 years). Meta-analysis and systematic reviews have demonstrated the fastest process, requiring an average 1.3 years and 1.9 years respectively. Case–control, cohort and quasi-experimental studies have required more time to enjoy publication in nursing journals, 4 years, 3.5 years and 3.2 years respectively. The results implied that the production time of an article from its data collection involves significant processes and skills.

Gaskin and Happell (2014) investigated 10 highest ranked nursing journal published in the 2011 based on their 5-year impact factors to (a) assess the statistical power of nursing research to detect small, medium, and large effect sizes; (b) estimate the experiment-wise Type I error rate in these studies; and (c) assess the extent to which (i) a priori power analyses, (ii) effect sizes, and (iii) confidence intervals were reported. The results suggested that the use, reporting, and interpretation of inferential statistics in nursing research needed substantial improvement. Researchers should also abandon the misleading practice of interpreting the results from inferential tests based solely on whether they are statistically significant (or not) and, instead, focus on reporting and interpreting effect sizes, confidence intervals, and significance levels.

Nursing researchers also need to conduct and report a priori power analyses, and to address the issue of Type I experiment-wise error inflation in their studies.

Currie and Chipps (2015) conducted a study to identify military nursing papers in the main databases and to describe the field of military nursing research for the period 1990-2013 in terms of research productivity, trends in topic focus, trends in authorship, and country of publication. 237 research articles were examined through Bibliometric review together with SPSS version 21. Findings revealed that the majority of publications were from America (n = 175, 73.8%) and the number of papers has increased significantly since the beginning of the second Gulf War in Iraq from 2003 onwards (n = 156, 65.8%). It also was found that there has been a shift in topic focus from professional (n = 16, 20.3%) and occupational issues (n = 17, 21.5%) pre 2003, to clinical (n = 48, 30.4%) and an increase in multidisciplinary research from 4% in 1990-1994 to 29% in 2010-2013. The researchers concluded that the military nursing research field appeared stronger than it had been in the past twenty years and had demonstrated increased transferability to other fields.

Yang (2015) has created a Nursing Academic Word List (NAWL) from a 1,006,934-word corpus called the Nursing Research Articles Corpus (NRAC), which contains 252 English nursing research articles from online resources. The study aimed to create a nursing academic word list to serve as a guide for English for Specific Purposes (ESP) instructors and material designers in English for Nursing purposes (ENP) curriculum preparation and English for Academic Purposes (EAP) textbook design, to provide further evidence for EAP researchers who are interested in producing field-specific academic word lists and to facilitate nursing students' acquisition of academic vocabulary. The 252 nursing research articles were compiled

from 21 subject areas of nursing being taught at the researcher's university. There were four criteria for selecting the journal articles: 1) in the IMRD format; 2) written by native English speakers; 3) the articles have been published between 1995 and 2011; and 4) the articles with the length between 2000 and 10,000 words. 12 articles were selected from each of 21 subject areas. The articles were collected in their electronic version with their reference lists, appendices, captions, footnotes, and acknowledgments removed. Range, a software program, was used to analyze the corpus data. Three criteria used in creating a nursing academic word list include range, frequency, and word family by which members of a word family to be included in the NAWL had to occur in at least 11 subject areas, at least 33 times in the corpus and outside the first 2,000 most frequently-occurring word families of English. Findings showed that a frequency and range-based nursing academic word list generated consists of 676 word families, which accounts for 13.64% of the coverage in the NRAC under study. The NAWL generated was believed to provide a useful academic word pool for non-native English learners who need to read and publish nursing articles in English.

The previous studies concerning journal articles in the field of nursing as mentioned above showed that there has not been studies that focus on collocation. The closest study is a Nursing Academic Word List (NAWL) conducted by Yang (2015). By focus further on collocation, as the present study intends to do, should give more insight as well as broaden the study in both fields of nursing profession and the EFL/ESL learning and teaching.

2.7 Framework of the Present Study

The present study is originated from the idea that collocation is essential for natural and appropriate use of a language and an effective communication as well as the learning of a language in general as supported by a number of scholars namely Bennett (2010), Duan and Qin (2012), Farrokh (2012), Henriksen (2013), Hill (2000), Hoey (2007), Kozlowski and Seymour (2003), Lewis (2000), Nation (2001), and Wray (2002). A corpus-based lexical analysis is a great method of identifying and extracting collocations from a sample corpus of a particular field of specialization. The case of the present study is the Sample Corpus of Nursing Research Articles (SCNRA). With the identified and listed collocations, it is believed that it would greatly benefit the teaching and learning of the collocations in the field as supported by the lexical approach (Lewis, 1993; 1997) in teaching and learning of a language. The focus of the present study is on lexical collocations adapted from the combination patterns as provided by Benson et al. (2010). Since collocation, especially lexical collocation, is closely related to vocabulary as its larger units, the present study will look at the combination of lexical words within the three-word span from the right side of the node. A pre-test of collocational knowledge constructed based on the collocations gained from the study will be administered with the fourth year nursing students at SUT. This will be followed by lessons on collocations by means of a corpus-based instruction. The post-test will be then administered to determine the effect of teaching on students' performance. The framework for the present study can be presented in diagram as shown in Figure 2.2 below.
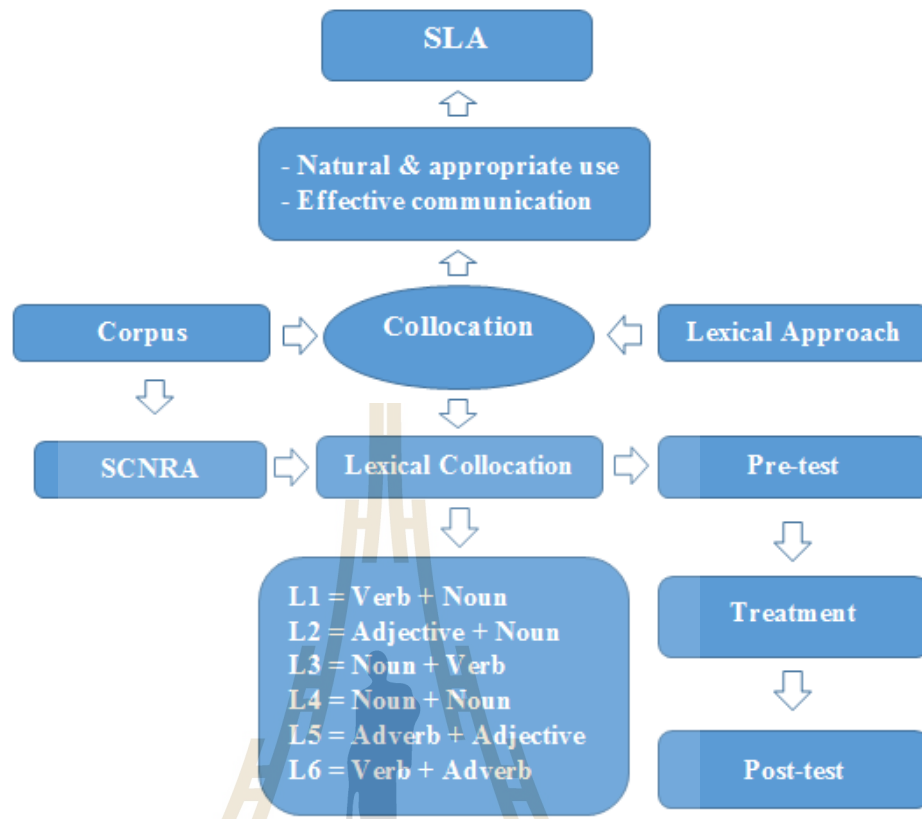
Figure 2.2 Framework for the present study

From Figure 2.2 above, the center is collocation as it is the central of the present study which is believed to lead further to natural and appropriate use of a language as well as effective communication as they are the indicators of the ultimate goal of second language acquisition (SLA). With this in mind, the present study aims to explore lexical collocations from a Sample Corpus of Nursing Research Articles (SCNRA) by means of corpus linguistics. The lexical collocations will be those adapted from that of provided by Benson et al. (2010). Following the extraction and classification of the collocations, a test of collocation knowledge will be constructed based on lexical collocations found from the study. The test will be conducted with the fourth year nursing students as a pre-test in order to evaluate their knowledge of

lexical collocations in their field. This will be followed by the lessons on collocations using a corpus-based instruction. The post-test will be conducted once the collocation lessons have completed to determine the effect of teaching on the test takers' performance.

2.8 Summary of the Chapter

This chapter presented reviews of the literature related to the present study. It started by exploring the relationship between vocabulary and SLA which involves how vocabulary and collocation are related, teaching and learning of vocabulary, lexical approach and lexical items, formulaic language, and mental lexicon. The following part was the reviews on collocations which covered the definitions, the classifications, the types, and the criteria for identification. The importance of collocation as well as the teaching, tests of collocation knowledge, and collocation study in Thailand context were also provided. The third part provided information on corpus studies which includes the definitions, the development, types, and benefits. The information on concordance software, corpus-based lexical analysis as well as corpus-based language teaching and learning were also illustrated. The fourth part provided information on ESP in relation to collocations and corpus-based instruction. The fifth part introduced pre-experimental research design. The sixth part reviewed journals and journal articles as well as nursing journal and journal articles. The last part described the framework of the present study which crystalized from the reviews of literature and the rationale of the study.

CHAPTER 3

RESEARCH METHODOLOGY

This chapter presents research methodology to be applied in this study. It starts with a research design which consists of two main parts. Then it provides in detail regarding the instruments, data collection, and data analysis for each part of the research involved in the study. The data analysis which describes how the collected data are analyzed is provided. The summary of the chapter is also given.

3.1 Research Design

According to Biber et al. (2006), a corpus-based study is a method applied for investigating "association patterns" of language in a corpus. These association patterns are either linguistic features or their varieties or texts. In terms of linguistic association, there are either lexical or grammatical associations. Lexical associations can be investigated by looking at particular words for their association with other words. Grammatical associations can be examined by looking at the grammatical constructions of the language. In terms of varieties or texts, the investigation can be on linguistic association patterns such as individual linguistic features or classes of features or co-occurrence patterns of linguistic features. Since the present study started with the investigation of lexical collocation, which is considered as a part of linguistic association patterns of a language under lexical association patterns, a corpus-based design is appropriate for this study.

As the objectives of this study were: (1) to identify and classify keywords found in the Sample Corpus of Nursing Research Articles (SCNRA) published in international journals in the field of nursing; (2) to identify and classify lexical collocations using keywords generated from the SCNRA; (3) to assess lexical collocation knowledge of nursing students based on the collocations found from the SCNRA; and 4) to provide lessons on nursing collocations and assess the effects of corpus-based instruction, this study thus consists of two main parts: the identification and classification of lexical collocations and the tests of nursing collocation knowledge of nursing students. With these two parts, the research part I provides adequate information for the first two objectives, while the research part II gives a clear conclusion regarding collocational knowledge of nursing students under investigation and the effects of corpus-based instruction on their performance by which a pre-experimental research design with One-Group-Pretest-Posttest Design was employed.

3.2 Research Part I: Identifying Lexical Collocations

This part provides the answers to Research Questions 1 and 2.

RQ 1: What are the keywords in the SCNRA based on the frequency of occurrence at ≥ 50 and the keyness value at ≥ 20? What is the proportion according to their parts of speech?

RQ 2: What are the lexical collocations of the keywords in the SCNRA? What is the proportion according to each type of combinations?

3.2.1 Research Instruments

The instruments applied in this part were research articles published in the selected nursing journals and a corpus analysis tool: a concordance software.

Nursing Journals and Nursing Research Articles

In terms of the research articles used as the data for the present study, they were those research articles published in the academic journals in the field of nursing accessible via SUT's library resources. This is important as these journals are meant to serve both academic staff and students of the Institute of Nursing at SUT. By identifying lexical collocations from the Sample Corpus of Nursing Research Articles (SCNRA), it should be directly beneficial for both academic staff and students of nursing at SUT. Regarding the availability of the journals, most of the journals accessible are as up to date as 2016. However, a few of them can provide the access of the latest issues with a few years back. From a questionnaire sent out and replied by the nursing students, these journals are recognized by the majority of the students.

A Corpus Analysis Tool

As this part of the study involves the examination of lexical collocations from the SCNRA, a concordance software is necessary. The AntConc version 3.4.4 (Anthony, 2014), which is the latest version, was used. This program has been developed by Laurence Anthony, a professor in the Faculty of Science and Engineering at Waseda University, Japan. The reasons why the AntConc has been chosen as a corpus analysis tool for the present study are:

1) AntConc is a computer-based freeware corpus analysis toolkit for concordancing and text analysis. This makes it economical and practical for anyone working with a corpus.

2) AntConc can operate on Windows, Macintosh OS X, and Linux.

3) AntConc is equipped with seven tools namely Concordance Tool, Concordance Plot Tool, File View Tool, Cluster/N-Grams, Collocates, Word List, and Keyword List. With all these available features of the program, it serves as the appropriate tool for the present study.

4) A number of previous studies had employed AntConc successfully as a tool. These studies are; for example, by Yang (2012) in investigating gender representation in an English textbook series used in Hong Kong schools; by Stvan (2013) in examining of the term 'stress' in naturally occurring vernacular prose; by Molavi et al. (2014) in investigating the distribution of lexical collocations in three selected series of general English textbooks; by Gulec and Arif Gulec (2015) in exploring verb-noun lexical collocations across the health, physical and social sciences in the written academic journals and analyzed these lexical collocations through the frequency and chi-square analysis; and by Getkham (2016) in investigating how linguistic devices are used to convey authorial stance in 36 Introduction sections and 36 Discussion sections of doctoral dissertations written in English by Thai students graduated in language education from different universities in the United States during the period 2008 to 2013.

From the accessibility and features the AntConc has, it can efficiently serve the purpose of this study as it can identify keywords of the sample corpus and further identify collocations of those keywords. According to Wilkinson (2012), in comparison to other tools available, AntConc is quite efficient. It is quite compatible with the best available commercial programs such as WordSmith and MonoConc Pro. The interface of the AntConc version 3.4.4 is shown in Figure 3.1 below.
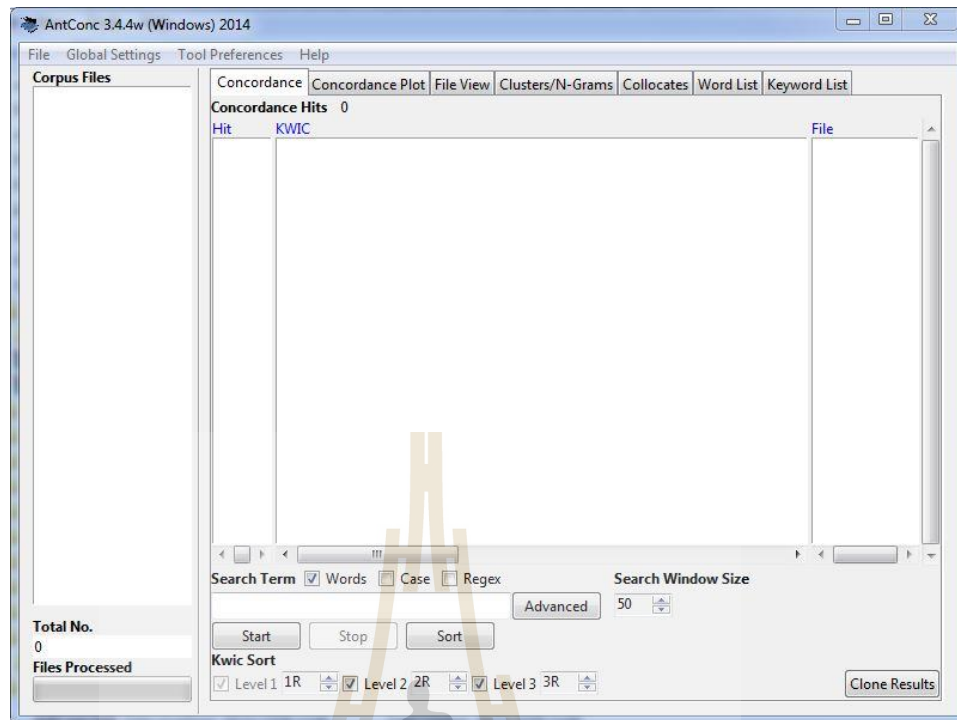
Figure 3.1 The Interface of the AntConc version 3.4.4

3.2.2 Data Collection

The data collection procedures in the first part of the present study involves the selection of nursing journals and nursing research articles as well as the compilation of the selected research articles for data analysis.

3.2.2.1 Selecting Nursing Journals and Nursing Research Articles

Since this study focuses on lexical collocations in nursing research articles, the process of selecting the research articles started with identifying journals in the field of nursing accessible via SUT's library resources. The targets of this study were nursing journals with international recognition and internationally serve researchers in the field to publish their research work. Once the journals had been selected, the articles from the latest publication available were obtained and compiled into text files ready for use with the concordance software. The latest issues of the journals

collected were as up to date as 2016. Only some of them were a few years back. The procedures of selecting the nursing journal are shown in Figure 3.2 below.

```
┌─────────────────────────────────────────┐
│         SUT's Library Resources         │
└─────────────────────────────────────────┘
                    ⇩
┌─────────────────────────────────────────┐
│            Academic Journals            │
└─────────────────────────────────────────┘
                    ⇩
┌─────────────────────────────────────────┐
│            Nursing Journals             │
└─────────────────────────────────────────┘
                    ⇩
┌─────────────────────────────────────────────────────┐
│  32 International Nursing Journals available Online  │
└─────────────────────────────────────────────────────┘
                    ⇩
┌─────────────────────────────────────────────────────┐
│    10 Nursing Journals matched with areas offered at │
└─────────────────────────────────────────────────────┘
                    ⇩
┌─────────────────────────────────────────────────────┐
│  30 articles in IMRD format from Latest issues of each selected │
└─────────────────────────────────────────────────────┘
```
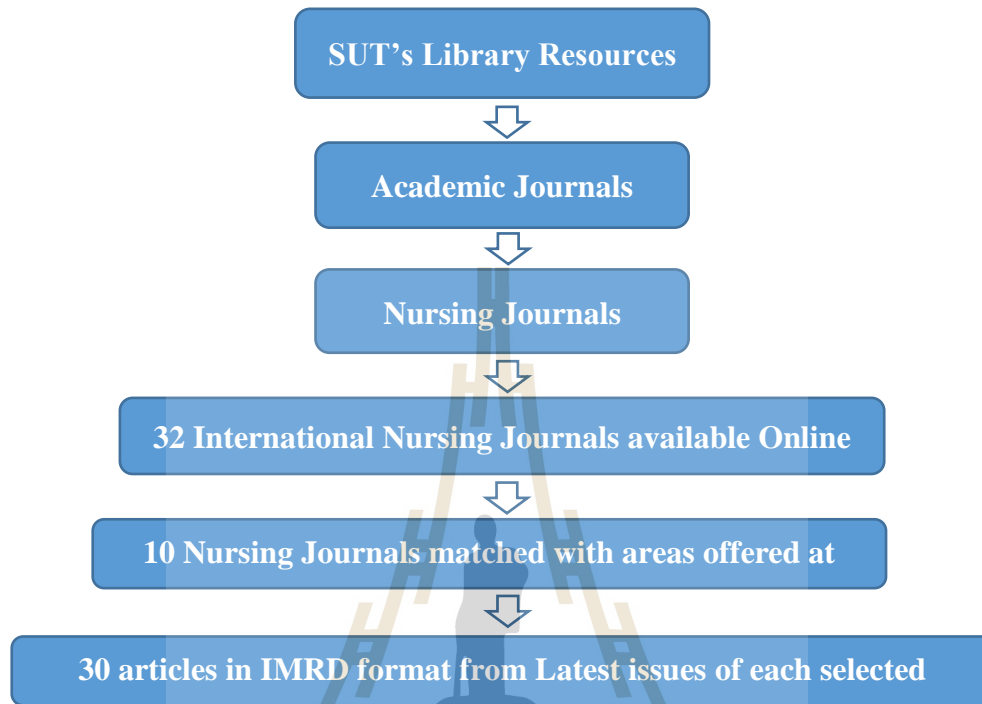
Figure 3.2 The process of selecting nursing journals from SUT's Library resources

Criteria for selecting nursing journals

As there were altogether 32 international journals accessible online via SUT's library resources (See Appendix A), 10 journals were selected for the present study. These 10 journals are purposively selected for the study based on the following criteria:

1) The titles which indicate the specialized areas that match with those were on offer by Institute of Nursing at SUT, which were six areas as follows:

(1) Nursing care of families and midwifery,

(2) Nursing Care of the Child and Adolescent,

(3) Adult and elderly nursing or Gerontological nursing,

(4) Community nursing,

(5) Psychiatric nursing, and

(6) Fundamental nursing

2) These journals were accessible online with full research articles in the IMRD format available either in HTML or PDF files. This is to prove the ease of assess of the journals and the research articles they contain for the academic staff and students of SUT as well as the researcher.

3) The selected journals were those with the Impact Factor (IF), since it indicates the popularity and credibility of the journals.

4) From the information gained from the fourth year nursing students of Academic year 2015 through an online questionnaire, the students recognized most of the selected journals and had read research articles from these journals.

Based on the criteria above mentioned, the selected nursing journals for the present study are listed in Table 3.1 below with their impact factors.

Table 3.1 The selected Nursing Journals with their Impact Factors

No.	Nursing Journals Selected	IF	Areas offered at SUT
1	Journal of Epidemiology & Community Health	3.501	Community health nursing
2	International Journal of Mental Health Nursing	1.95	Psychiatric nursing
3	Journal of Nursing Management	1.5	Fundamental nursing and others
4	Nursing Inquiry	1.439	Fundamental nursing and others
5	Journal of Family Nursing	1.342	Nursing care of families and midwifery
6	Clinical Nursing Research	1.278	Fundamental nursing and others
7	Journal of Clinical Nursing	1.255	Fundamental nursing and others
8	Journal of Pediatric Oncology Nursing	0.903	Nursing Care of the Child and Adolescent, Adult and elderly

| 9 | Journal of Psychiatric and Mental Health Nursing | 0.844 | Psychiatric nursing |
| 10 | International Journal of Nursing | 0.98 | Fundamental nursing and others |

As shown above in Table 3.1, it is noticeable that among the 10 journals selected, there are five journals that matched with the area of Fundamental nursing and others and the other five journals matched with each of the five areas offered at the Institute of Nursing at SUT. Since six areas of nursing fields were offered with 10 journals to be selected, this was not possible to make equal representative of each area. Therefore, it is justifiable to select one journal of each of the six areas with the rest four journals selected were those journals belong to the area of Fundamental nursing and others since this area does not represent specific areas as the other five journals do.

Criteria for selecting nursing research articles

Once the journals had been selected, the research articles from those selected journals were compiled according to the following criteria:

1) The selected research articles were those which had been written in the appropriate proportions of IMRD (Introduction, Methods, Results, and Discussion) format with length not less than 3,000 words. The IMRD format is a format commonly used in quantitative and experiment-based research (Englander, 2014). This format of writing can also be found in the journals in the field of nursing.

2) From each journal, 30 research articles were selected based on the latest issues available at the time of the study being conducted. The selection started from the latest available issue of each journal. Each journal was examined for their accessibility following with their format of writing where the IMRD was the target. If

the article met the criteria, it was saved in a text file format and named for the ease of management. For the articles that did not meet with the criteria set, they were not chosen for this study. The process was repeated until 30 articles were collected for each journal. This made up the total of 300 research articles compiled for the present study.

In terms of corpus size, McEnery et al. (2006) explain that the size of the corpus needed depends on the purpose for which it is intended. Quesada (2011) points out that a corpus is big enough when, for any new learning experience added, the probability of adding a new type is so low that is negligible. This means that adding more to the size of a corpus will have little effect on the new data to be found. For the present study, as research articles have a typical length ranging from 3,000 to 10,000 words (Björk, Roos, & Lauri, 2009), the SCNRA compiled from 300 research articles comprised about 1,500,000 words. This is efficient for the investigation of lexical collocations the present study intends to achieve. With this size of the SCNRA, it is large enough to provide acceptable frequency of occurrence to the keywords as well as the co-occurrence of the collocations under investigation. The files of the research articles compiled into the SCNRA were manageable and analyzed using a concordance program, the AntConc. The details of the sample corpus in terms of the number of files are shown in Table 3.2 below.

Table 3.2 Nursing Journals with the number of research articles used in the study

No.	Nursing Research Articles taken from	Quantity
1	Journal of Epidemiology & Community Health	30
2	International Journal of Mental Health Nursing	30
3	Journal of Nursing Management	30
4	Nursing Inquiry	30

5	Journal of Family Nursing	30
6	Clinical Nursing Research	30
7	Journal of Clinical Nursing	30
8	Journal of Pediatric Oncology Nursing	30
9	Journal of Psychiatric and Mental Health Nursing	30
10	International Journal of Nursing Practice	30
	Total	300

3.2.3 Data Analysis

The data analysis in this part involves two parts: the identifying keywords and lexical collocations generated from the keywords of the SCNRA.

3.2.3.1 Procedures in Identifying Keywords

For the present study, keywords according to the keyness value were listed for further use as nodes to identify their collocates. The number of keywords for the present study had been determined by two criteria: the keyness value and the frequency of occurrence. The keyness value set for the present study for the keywords is ≥ 20, while the frequency of occurrence is ≥ 50. According to Baker (2006), the higher the keyness score, the stronger the keyness of that word. The frequency of occurrence indicates how common the word is in the corpus. The identification of the keywords in the SCNRA followed the steps below:

1) Upload the text files onto the analysis tool, the AntConc, to form the SCNRA.

2) Set preferences in 'Global Settings'.

3) Set preferences in 'Tool Preferences'. At this stage, under '*Collocates*', MI value is selected as a statistic measure for the association strength of collocational pairs. Also, at this stage, under '*Keyword List*', Reference Corpus is uploaded. For the

present study, the British National Corpus (BNC) is used as the reference corpus as it is one of the largest corpora of general English and commonly used as a reference corpus in corpus-based studies as well as the word list is readily available online.

4) Once all the necessary preferences are set, the analysis can be commenced. By clicking on the '*Word List*' tab followed by clicking on '*Start*' button, the results will reveal the numbers of types and token the SCNRA contains. The concordance software also displays the list of the words by ranking according to the frequency of occurrence.

5) By clicking on a '*Keyword List*' tab followed by the '*Start*' button, the list of keywords of the SCNRA will be revealed.

As not every word displayed on the keyword list is acceptable as keywords, for example, '*et*', '*al*', proper nouns, abbreviations, and acronyms can be on the keyword list, they were excluded from the list. To create the keyword list according to the set criteria, the researcher looked through the list according to their keyness ranking and manually deleted unacceptable keywords that leave the rest to become the keyword list under the set criteria which are words with the frequency of occurrence at ≥ 50 and the keyness score at ≥ 20. Once the list of keywords had been made, they were ready for the next stage of the study, identifying collocations. These keywords were also classified and grouped according to their part of speech.

3.2.3.2 Procedures in Identifying Collocations

Once the keywords had been identified, as they were then used as the 'nodes' to further identify their collocates, the next stage of the study could be proceeded. The processes of identifying collocations in the SCNRA were in steps as follows:

1) Under the '*Keyword List*' tab, click on the keyword acceptable to be the node, for example, the word '*Nurses*'. The concordance program displays under the '*Concordance*' tab to show the concordance lines the word '*Nurses*' appears in the entire corpus with the frequency of occurrence.

2) To find the collocates of a node, click on the '*Collocates*' tab.

Before clicking on '*Start*' button, there are three preferences to be set as follows:

Under '*Window Span*', set as '*From 0* to *3R*' as the study intended to investigate the collocates with three-word span on the right side of the nodes.

Under '*Sort by*', select '*Sort by Stat*' as the study gave more important to the MI score than the frequency of occurrence.

Under '*Min. Collocate Frequency*', set as '*10*' as the study set the number of the co-occurrence of the pairs at $\geq$ 10. However, the number of occurrence could be reduced in case of the number of co-occurrence did not meet the criteria that is when the frequency of occurrence is less than 10. In that case, the intention would be paid only on the MI value of the pair which the present study is set at $\geq$ 5. Also, in this case, the first collocate with the highest MI value and the highest frequency of occurrence would be selected.

Once the preferences were set accordingly, click on '*Start*' button.

3) After clicking on '*Start*' button, the collocates of the node would be displayed if they were any. The strength of each pair of lexical collocation identified was measured on the basis of Mutual Information (MI) as provided by the AntConc. In order to measure the association strength of each pair, there is the comparison between the study corpus, the SCNRA, and the reference corpus. The reference

corpus used in this study is the British National Corpus (BNC) as it is readily available and comprises over 100 million words. It is also one of the largest corpora of general English and commonly used as a reference corpus.

The acceptable association strength of co-occurring pair is the MI value of ≥ 3 (Cheng, 2012). The frequency of the co-occurrence is one important factor to be considered. For the present study, the frequency of co-occurrence of each pair has been set at ≥ 10 as the higher rate of co-occurrence means higher the need for students of nursing to be familiar with. In case of a pair co-occurs less than 10 times but has MI value ≥ 5, the collocate with the highest occurrence and MI value would be chosen. As well, in case of a node has more than one collocates that meet the criteria, all of them would be selected.

As the criteria set for the study that the collocates have to have the MI value at least 5 with the frequency of occurrence at least 10, not all of the words on the list may be acceptable collocates. Even those words that meet the criteria may not always eligible to be acceptable pairs. To check whether a collocate is acceptable pair of a node or not, by clicking on the word, it will show how the pair co-occur in the sample corpus. In case of no collocates with the number of occurrence of 10 or more with the MI value at least 5, the reduction of the '*Min. Collocate Frequency*' could be applied. The acceptable pairs of collocations were recorded for each keywords or nodes as the results and for further analysis.

4) The process was repeated for other keywords/nodes to find their collocates. The results of each keyword with their collocates were recorded using Microsoft Excel spreadsheet.

5) The collocations generated from the keywords or nodes found were then classified according to the types of combination. The combination patterns of the lexical collocations in the present study have been adapted from that of Benson et al. (2010). That is from the original seven combinations, the present study has adapted to six combinations. This is to create lexical collocations of two word pair on the right side of the nodes within the range of four word span. This combination is meant to help add up to single word vocabulary which should be a proper stepping stones for students to learn as the extension of the normal single unit vocabulary. The comparison between the combinations given by Benson et al. (2010) and the adapted combinations for the present study is shown in Table 3.3 below.

Table 3.3 Lexical Collocations: Benson et al. (2010) and Adapted in Comparison

Types	Combinations by Benson et al. (2010)	Types	Combinations adapted
L1	Verb + Noun		
L2	Verb + Noun	L1	Verb + Noun
L3	Adjective + Noun	L2	Adjective + Noun
L4	Noun + Verb	L3	Noun + Verb
L5	Noun + *of* Noun	L4	Noun + Noun
L6	Adverb + Adjective	L5	Adverb + Adjective
L7	Verb + Adverb	L6	Verb + Adverb

From Table 3.3 above, among the seven lexical collocations with the combinations developed by Benson et al. (2010), L5 combination is not in the two word combinations having '*of*' between the nouns. Also for the L1 and L2, both are the combinations of Verb + Noun. In order to create lexical collocations within the range of four word span, the L5 is adapted by looking only for the combinations of Noun + Noun without having '*of*' between them. In this case, however, if there is a

combination between nouns occurs within the range of four word span with '*of*' in it as in Noun + *of* + Noun, such combination is the legitimate combination of the present study. The Verb + Noun combinations in L1 and L2 are reduced to have only one Verb + Noun combinations. This is to reduce the complexity of the combinations. Figure 3.3 below shows the process of identifying keywords and their collocates.
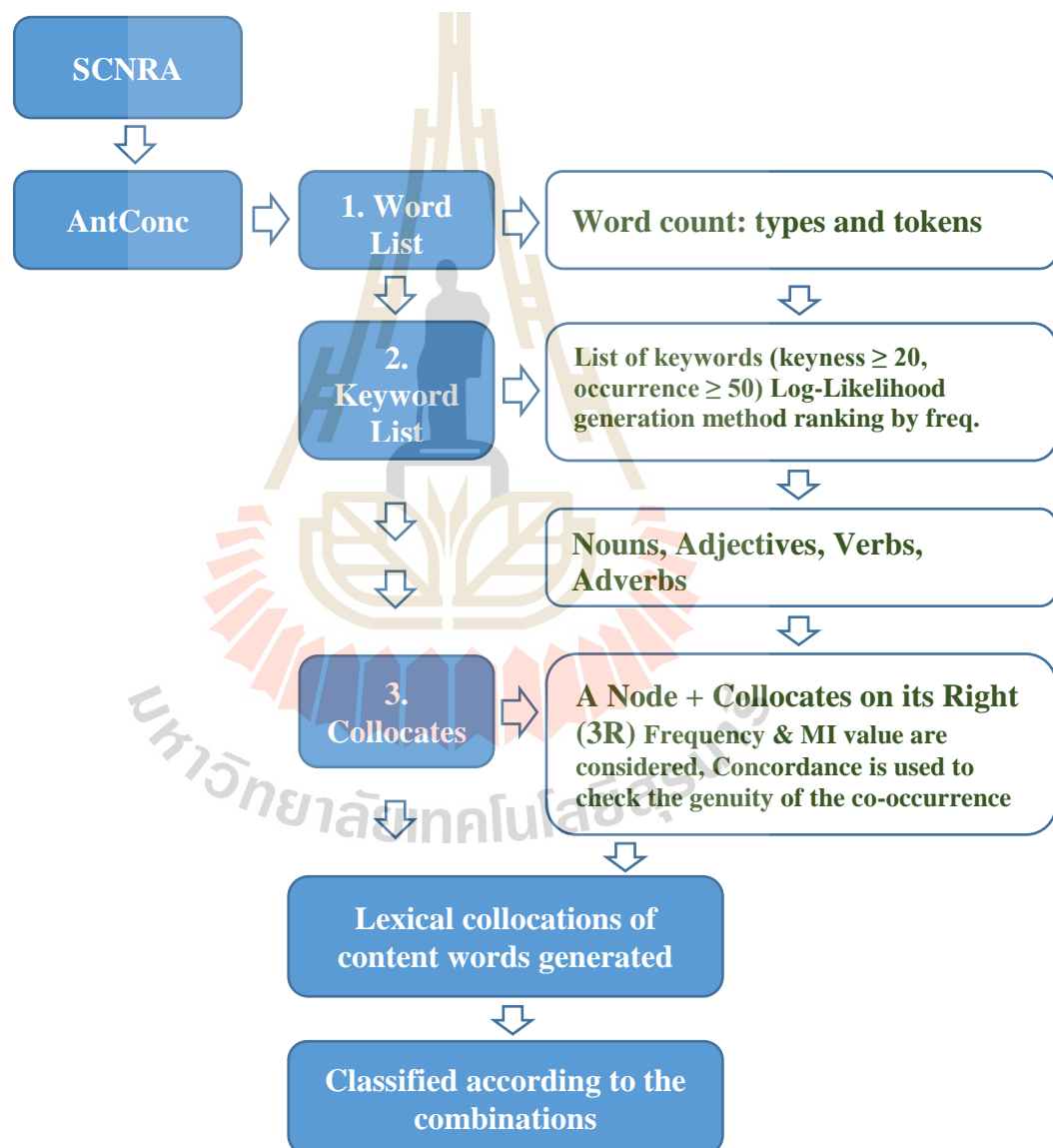
Figure 3.3 The process of identifying keywords and collocations from SCNRA

The collocations gained from the SCNRA, particularly lexical collocations, were then used for the second part of the study, the test of collocation knowledge of the fourth year nursing students at SUT.

3.2.4 Pilot Study

The objectives of the pilot study were: (1) to examine the plausibility of the research project which intends to explore lexical collocations found in a Sample Corpus of Nursing Research Articles (SCNRA), and (2) to explore whether any adjustments were necessary in order to conduct the main study successfully.

The pilot study was conducted with a sample corpus of nursing research articles compiled from 10 journals in the field of nursing intended for the main study. The pilot study was conducted by following the procedures designed for the main study, to investigate two immediate word lexical collocations. The differences are that the number of the articles used where five research articles were taken from each of the listed journals. These selected research articles were also taken from the earlier issues than research articles to be used for the main study to avoid the repetition of articles in the main study.

The sample corpus of 50 research articles from the 10 journals in the field of nursing comprises 11,517 types and 225,248 tokens. 500 keywords were selected according to the ranking of their keyness value for further analysis. When classified into categories according to the part of speech, the 500 selected keywords consist of 348 nouns (65.5%), 118 adjectives (22.2%), 58 verbs (10.9%), and 7 adverbs (1.32%)

The identified keywords were then used as 'nodes' for the next step of the study which was to find out their collocates. In the pilot study, as the lexical approach was emphasized, lexical collocations of the nodes had been explored. From the

keywords according to their parts of speech, the collocations found from the sample corpus were divided into three groups as follows: 1) lexical collocations according to the set framework; 2) collocations not according to the set framework; and 3) keywords/nodes with no collocates.

For the first group, which presents collocations the present study intends to investigate, it reveals the lexical collocations of 281 keywords or nodes that account for 52.92 percent. The second group, collocations which are not according to the set framework, comprises 174 keywords or nodes with their collocates other than lexical collocations. The majority of these collocations can be categorized as grammatical collocations. The last group, keywords or nodes with no collocates, comprises 76 keywords that account for 14.31 percent with the majority being the nouns.

From the findings gained from the pilot study, it is found that from the 500 keywords generated, the majority of them, 65.5 percent, being the nouns. The adjectives come the second place at over 22 percent. The verbs come third at over 10 percent and the adverbs are the least in number, just over 1 percent. Among the three groups categorized according to the combinations, the majority being lexical collocations with the combinations according the set framework of over 52 percent, generated from 281 keywords. The rest of the keywords give the results not according to the set framework that can be divided into two groups: grammatical collocations and words with no collocations.

As the main focus of the study is on the group of lexical collocations to further apply in the assessment of SUT nursing students' knowledge of lexical collocations, the proportion of the group with over 52 percent and 281 in number is considered large enough to meet the purpose. Although the results reveal that there are

combinations of words both within the framework and outside of the framework, all of the combinations are worth investigation as they actually occur in the real use of the language. Thus, despite the focus of the study is on the lexical collocations of the set framework, knowing other possible collocates or no collocates is still useful in effective communication and use of the language as well as the EFL/ESL learning and teaching. With the larger sample size in the main study, it is believed that the results could be slightly different and more statistically significant.

As the objectives of this pilot study were to examine the plausibility of the research project as well as to explore whether any adjustments were necessary in order to successfully conduct the main study, the results of the pilot study assured that this research project was plausible. However, there was a concern in terms of the meaningfulness of the combination of two immediate word pairs. Therefore, in the main study, the range of word span of collocations has been expanded from one position on the right side of a node (1R) to three words (3R) in the hope of finding collocations with more meaningful combinations. Besides, the pilot study in which the AntConc has been used as the concordance tool has proved that the AntConc is the efficient tool for the study. The complete report of the pilot study is attached in Appendix B.

3.3 Research Part II: Tests of Lexical Collocation Knowledge

This second part of the present study is to examine the fourth year nursing students at SUT on the knowledge of lexical collocations in their field which have gained from the SCNRA in the first part of the study. The results gained from this part

of the study were used to answer Research Question number 3 and 4 of the present study.

RQ 3: How much collocational knowledge do the fourth year nursing students at Suranaree University of Technology (SUT) have based on a test of lexical collocations extracted from the SCNRA?

RQ 4: How much does corpus-based instruction help improve the knowledge of collocations for the fourth year nursing students at SUT?

3.3.1 Participants

Fifty-one fourth year undergraduate nursing students at SUT in trimester 1 of the academic year 2017 were the participants in the test of collocation knowledge. As this group of students was in their final year of study and preparing themselves for their career, it is essential to determine their collocational knowledge in their professional field. This can be a useful indicator for measuring whether the students are sufficiently equipped with knowledge of collocations in their field, the knowledge which should be very useful for both their professional lives and academic advancement.

From the questionnaire asking for their grades received from English courses taken earlier in their previous years of study namely English I to English V, most of them received Bs and higher grades, with some students reported to have Cs. This indicates that this group of students had satisfactory background in English language. The highly competitive in entering the program also is believed to be another factor that make these students to have higher academic performance before entering and during in the program.

3.3.2 Lexical Collocation Test

The test of collocation knowledge of nursing students mainly involves the construction of the test as well as the administration of the test and scoring. There were a pre-test and a post-test. This means that the same test was administered before and after lessons on lexical collocations provided. This is to compare how corpus-based instruction affects the performance of the students. The lesson plan for the lessons provided to the nursing students in details is shown in Appendix C.

3.3.2.1 Procedures in the Construction of Lexical Collocation Test

The procedures involved in this part of the study concern with the test format to be used, the selection of lexical collocations to be tested, the test construction, as well as the pilot of the test.

3.3.2.1.1 Test format

According to Jaén (2007), tests of collocation knowledge can be divided into two types: a test to measure receptive knowledge and to measure productive knowledge. A test of receptive knowledge intends to explore the ability of test takers whether they can select the most appropriate collocates of the nodes given in certain situations. The test formats are those in which alternatives are provided such as in a multiple choice format. For a test of productive knowledge, this type of test is designed to measure whether the test takers are able to use the collocations appropriately in given situations. The test formats are gap filling tasks where test takers are required to fill the missing pairs of the given collocations or translation tasks where test takers have to translate the given collocations from L1 to L2 or vice versa.

The test format to be applied in the present study in order to test collocational knowledge of nursing students at SUT consists of two types of test: a test to measure receptive ability and a test to measure productive ability. For a receptive ability test, a multiple-choice test and a gap-filling with choices provided were used. For a productive ability test, a short sentence writing task was applied since it should very well be able to reflect the ability to produce the language of the test takers.

Since the test comprises the total of 60 test items, the receptive test consists of 30 items of multiple-choice format and 20 items of gap-filling task. The productive test consists 10 items of a short sentence writing task. The number of tasks and test items set are considered adequate amount to allow the test takers in revealing their knowledge to be measured. The time allow for test takers was two hours. The number of test items and the duration allowed were appropriate to keep the test takers to stay focused on the tasks which could give the optimal reliability to the results.

3.3.2.1.2 Item selection

As the lexical collocations gained from the SCNRA comprise over 2,000 pairs, it is not possible to put all of them into the test. Therefore, the collocations to be tested are those pairs with the highest frequency of occurrence. Therefore, the most 200 frequent collocation pairs were randomly put to the test. Also to prevent confusion, only the most 200 frequent pairs of each node (Appendix M) were used as the target collocation pairs regardless of the types of combination.

3.3.2.1.3 Test construction

In this study, after lexical collocations in the SCNRA had been identified and classified, the test of nursing collocations was constructed. The randomly selected 60 collocation pairs were put into three groups according to the test tasks: 30 collocations

for a multiple-choice test, 20 collocations for a gap-filling task, and 10 collocations for a short sentence writing task. This means that for the receptive test, there are a multiple-choice test and a gap-filling task, while the productive test comprises a short sentence writing task.

A multiple choice test consists of 30 discrete sentences with one blank space for the correct pair of the given collocation to be filled in on its right side. Four choices are given for each item.

Example: *health consequences*
The health _____ of tobacco products are well established.
(a) aftermaths   (b) properties   (c) consequences   (d) reactions

A gap-filling task consists of 20 discrete items with one blank in each sentence similar to those of a multiple choice test. However, to narrow down the possible collocates to fill the gaps as well as to make it clear for the test to have only one correct answer for each item, the test items are divided into four parts of five test items in each. In each part, the five correct collocates for each item are provided in a box with a choice from "a" to "e".

Example: a. stage   b. exposure   c.index
1. The implementation _____ involves putting resources in place.
2. This study was to explore multipollutant _____ assessments.

A short sentence writing task consists of 10 collocation pairs for the test takers to write a meaningful sentence containing the each given pair.

For example: *public sector*, *health problems*
public sector: _____
health problems: _____

To make the full use of the SCNRA and to make the test be more specific to nursing context and well as to enrich the test with the authentic language in the real use environment, the sentences to be used to construct the test in order to test each collocation have been taken from the SCNRA. However, simplification may be made to reduce the complexity of the sentences. The test comes with the answer sheet which is separately provided.

3.3.2.1.4 Piloting of the Test

Once the Nursing Collocation Test had been constructed, it was piloted with the fourth year students of nursing during trimester 3 of the academic year 2016 at SUT as these group of students were closest in terms of the target population. This is to assure the validity and reliability as well as practicality of the test. Item analysis was also applied in this process by examining three types of information: difficulty index, discrimination index, and analysis of response options. According to Haladyna, Downing, and Rodriguez (2002), item analysis is "a process of examining class-wide performance on individual test items" (p. 309). The difficulty index indicates the item's difficulty by the number of test takers who answer a particular test item correctly. The discrimination index, meanwhile, is a basic measure of the validity of a test item. It is a measure of an item's ability to discriminate between those who scored high on the total test and those who scored low. The analysis of response options, however, examines the performance of individual distractors on multiple-choice items. By calculating the proportion of test takers who chose each answer option, it would reveal that which distractors are 'working' and appear attractive to test takers who do not know the correct answer, and which distractors are not being chosen. The analysis of response options, therefore, is a great way to fine tune and improve

validity and reliability of the test items. The details analysis of the pilot test is shown in Appendix D. The finalized version of the Nursing Collocation Test had been improved and adjusted according to the results found from the pilot test. The complete version of the Nursing Collocation Test which was later used in the pre-test and the post-test is shown in Appendix E.

3.3.2.2 Test Administration and Scoring

The test had been administered with 51 target group of the fourth year nursing students at SUT during the first semester of academic year 2017 in a traditional paper-based manner. The duration allowed for the test was two hours. The test takers were seated in the same room and separated as far as possible to avoid distracting and cheating. The test takers who finished the test before the time given were allowed to leave the room.

In terms of scoring, as each test item has equal score of one point, the total raw score of the test was 60. However, the scores can be viewed as the whole test score and separately according to each of the three parts of the task types. The total score were converted into percentage for the convenience of analysis and clear measurement. The five scale score range was applied by dividing the range equally to form five scale level of performance as shown in Table 3.4 below.

Table 3.4 Score range and performance

No.	Score range	Performance
1	80 - 100	Excellent
2	60 - 79	Good
3	40 - 59	Fair
4	20 - 39	Poor
5	0 - 19	Very poor

For the multiple-choice and gap-filling task, as all of the test items contain only one best answer, the inter-raters were not needed. For a sentence writing task, although there are only sentences and with the quite clear holistic marking criteria as shown in Table 3.5 below, the inter-rater was still applied in order to assure the reliability of the results.

Table 3.5 Criteria for marking sentence writing task

Score	Criteria
0	No answer or answer with no meaningful expressions of the given pair.
0.5	Answer with some meaningful expressions of the pair, but not completely clear or with grammatical errors.
1	Answer with meaningful expressions of the pair with a complete and correct sentence structure.

After the pre-test had been administered, lessons on nursing collocation were provided. The post-test of nursing collocations were administered aftermath of the lessons being delivered. The scores of the pre-test and the post-test were compared to examine the test takers' performance in the two tests.

3.3.3 Data Analysis

For the first part of the study, the data analysis was both quantitative and qualitative. This involved classification of the analyzed data in order for answering the first two research questions. For the second part, the results of the test were examined quantitatively using Microsoft Excel and the Descriptive Statistics in the Statistical Package for Social Sciences (SPSS). The scores were analyzed as a whole as well as separately and compared between three test formats. The results of the test are meant to provide a clear answer to the last two research questions of the study.

3.4 Summary of the Chapter

This chapter presented research methodology in the present study. The research design showed that there are two main parts: collocation identification and classification part and the test of collocation knowledge part. The main instruments for the first part of the study are the selected nursing research articles for building a Sample Corpus of Nursing Research Articles (SCNRA), and a corpus analysis tool. The One-Group-Pretest-Posttest was employed in order to investigate nursing lexical collocation knowledge of the nursing students. The data analysis presented how the collected data were analyzed in order for answering the research questions of the study. The following chapter provides the results as well as the discussion of the present study.
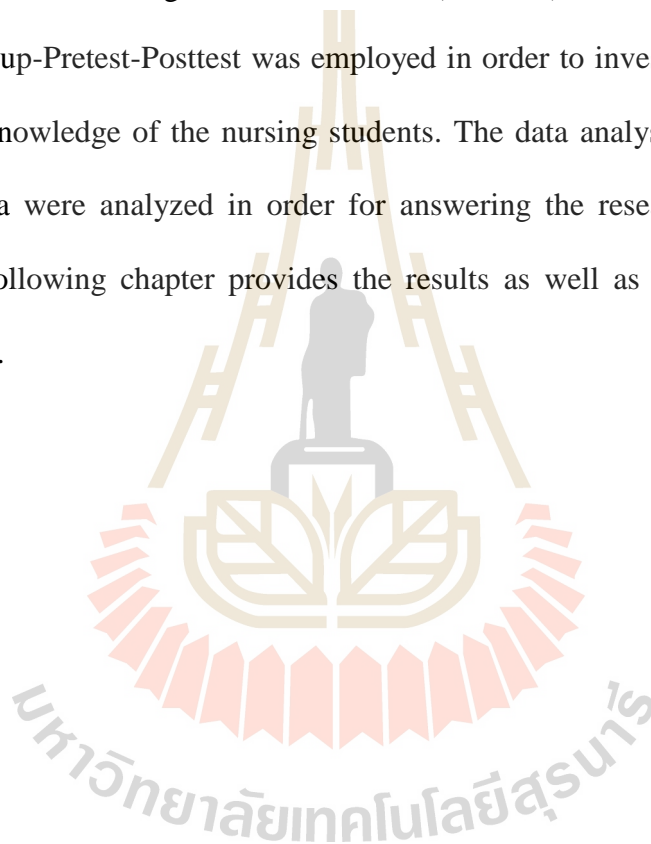
CHAPTER 4

RESULTS AND DISCUSSION

The findings of the present study consist of two parts as there have been two phases of study. Findings from part I provide the results of the lexical collocations extracted from the SCNRA. It starts from identifying the keywords from the SCNRA, and then identifying collocates of each keyword to find their collocation pairs. The findings of this part are the answers for research questions 1 and 2. The findings of the research from part II are the pre-test and the post-test results which are the product of the first part's study. This pre-test's results are used as the means for determining collocational knowledge of the nursing students. The post-test's results are analyzed in relation to the pre-test's results to evaluate how much students' performance have improved as the consequence of a corpus-based instruction of the collocations. The discussions concerning the findings are provided on the basis of each research question.

4.1 Research Part I: Identifying Lexical Collocations

This part of the study aims to provide answers to research questions 1 and 2 of the present study.

RQ 1: What are the keywords in the SCNRA based on the frequency of occurrence at ≥ 50 and the keyness value at ≥ 20? What is the proportion according to their parts of speech?

The SCNRA – a compilation of 300 research articles from 10 academic journals in the field of nursing comprises 28,054 types and 1,253,992 tokens. Under the set criteria where the frequency of occurrence is at ≥ 50 and the keyness value is at ≥ 20, there were 855 keywords generated. However, there were some words other than the content words appeared on the keyword list. These words are mainly the function words and pronouns, proper nouns which are not in the focus of the present study, thus they were deleted. Therefore, the number of the keywords found from the SCNRA to be used as "nodes" was 717. The full list of the keywords is shown in Appendix F. Most of the keywords generated have surprisingly high keyness value. The keyword with the highest keyness value is '*nurses*' with the keyness value at 34,638.35, while the lowest keyness value is '*understandings*' with the keyness value at 197.98. The ten keywords with the highest and lowest keyness values are shown in Table 4.1 below.

Table 4.1 Ten highest and lowest keyness value keywords

Ten highest keyness value keywords			Ten lowest keyness value keywords		
No.	Keywords	Keyness	Freq.	No.	Keywords	Keyness	Freq.

No.	Keywords	Keyness	Freq.	No.	Keywords	Keyness	Freq.
1	nurses	34638.34	5115	1	understandings	197.978	50
2	care	25519.82	6600	2	recommendations	198.375	146
3	health	25131.96	6540	3	category	198.471	171
4	study	21880.57	5739	4	suicidal	198.771	55
5	participants	19288.67	3058	5	onset	199.165	87
6	patients	19121.00	4895	6	clinically	199.85	62
7	nursing	17631.76	2991	7	perceive	200.212	90
8	patient	13012.95	3019	8	cluster	200.501	91
9	mental	9997.944	2245	9	indicates	200.542	141
10	nurse	9905.078	1930	10	problematic	200.658	84

At this point, the first part of research question 1 has been answered with the finalized number of 717 keywords. However, the answer for the second part of the

research question 1 is to be worked out alongside with the findings of the answers for research question 2, identifying collocations from those keywords.

Following the set criteria, the collocates of each keyword or the node have been identified and recorded according to their combination types on Microsoft Excel Spreadsheet. After the process of identifying collocates of the keywords has been completed, it is possible to categorize the keywords according to their parts of speech. The results showed that the majority of the keywords are the nouns, following with the adjectives, the verbs, and the adverbs respectively. The details of the numbers of each part of speech of the keywords and their percentage are indicated in Table 4.2 below.

Table 4.2 Keywords generated from SCNRA according to their parts of speech

No.	Parts of speech	Numbers	Percentage
1	Noun	463	63.51
2	Verb	98	13.44
3	Adjective	157	21.54
4	Adverb	11	1.51
	Total	**729**	**100**

Table 4.2 above shows that the keywords generated from the SCNRA are the content words found as keywords of 717 words. It is noticeable that the total number of the keywords has increased from 717 to 729. This is because along the process of identifying collocation pairs, it has been found that some of the keywords functioned more than one part of speech. Keywords such as '*use*' and '*need*' were found functioning as nouns and verbs as well. The examples are shown in Table 4.3 below.

Table 4.3 Keywords found function more than one part of speech

Keywords	Function as nouns	Function as verbs
need	(the) need (to) consider, (the) need (to) understand	need assistance
use	(the) use (of) technology	use lubricants

Another reason for the increase number of the keywords is that some compound words were found emerged from the original single keywords. The examples are presented in Table 4.4 below.

Table 4.4 Compound words emerged from single keywords

Keywords	Become compound words	Example of collocations
customer	customer-oriented	customer-oriented behaviors
evidence	evidence-based	evidence-based practice
family	family-centered	family-centered care
high	high-risk	high-risk behaviors
hospital	hospital-based	hospital-based oncology
nurse	nurse-led	nurse-led clinic
parent	patent-child	parent-child relationships
risk	risk-assessment	risk-assessment (*and*) management

Table 4.4 above provides the evidence to clarify the increased number of the keywords along the process of identifying collocates of the keywords. Due to some of the keywords which are single words combining with other words, new words which are compound words were formed. As a result, the number of the keywords increased.

Among these keywords, the majority of them are the nouns ($N = 463$), accounted for 63.51 percent. The adjectives ($N = 157$) come the second accounted for 21.54 percent. The verbs ($N = 98$) are accounted for 13.44 percent. The smallest number among them is the adverbs ($N = 11$) accounted for 1.51 percent. The full list of these keywords categorized according to their parts of speech is shown in Appendix I.

The prevalent number of content words such as nouns, adjectives, and verbs in the keywords seems to be common in all corpora. This findings can also be found in the studies conducted by scholars in the field such as Coxhead's (2000) Academic Word List; Mudraya's (2006) one hundred most frequent word families in the Student Engineering Word List; Wang, Liang, and Ge's (2008) Medical Word List; Ward's

(2009) Basic Engineering List generated from his Engineering Corpus; and Yang's (2015) Nursing Academic Word List.

The commonness of the nouns, adjectives and verbs in the corpora also reflects in a number of studies exploring the use of collocation pairs of these keyword types such as the verb-noun collocations in relation to the language proficiency and the knowledge of the verb-noun collocations in EFL learners (Ebrahimi-Bazzaz et al., 2012); the nouns and their collocates (Miščin, 2013); the use of adjective-noun collocations in comparison between learners in the regular and English programs (Suwitchanphan & Phoocharoensil, 2014); and the use of noun-noun collocations in learners' academic writing (Parkinson, 2015).

When comparing the 120 most frequent keywords from the SCNRA with the 120 most frequent academic word families in the Nursing Research Articles Corpus (NRAC) provided by Yang (2015), it is found that 36 words (29.17%) from the SCNRA match with that of the NRAC. These matched keywords are shown in Table 4.5 below.

Table 4.5 Matched keywords in the 120 most frequent of the SCNRA and the NRAC

approach	assessment	cancer	caregivers	clinical
conducted	data	diagnosis	factors	focus
individual	intervention/s	interview/s	items	medical
medication	mental	outcomes	participants	perceived
physical	positive	previous	process	professional/s
research	role	scores	significant	specific
status	symptoms			

Illustrated above in Table 4.5 are the common keywords found in the two corpora of nursing research articles. It may not seem very high in terms of number and percentage despite the corpus size of the two corpora is relatively similar. The corpus size of the SCNRA is 1,253,992 words, while the NRAC's is 1,006,934 words.

However, the possible cause for not having remarkably high common keywords in the two corpora may stem from the range of nursing journals selected for each corpus. That is the SCNRA comprised 300 research articles from journals in six subject areas, while the NRAC was built from 252 research articles from journals in 21 subject areas which could lead to the difference of words found in the two corpora.

Another factor that may contribute to the difference in the word list found from these two corpora is the way the keywords were listed. The Nursing Academic Word List (NAWL) was generated by means of word families, while the keyword list extracted from the SCNRA in the present study was produced on the basis of individual words. By producing the word list on the basis of word families, it is not possible to determine the frequency of the words when classified according to the part of speech. The different corpus tools used and the criteria applied in extracting the word list may also be the cause of the different results of these two corpora.

With a relatively high match of keywords in these two sample corpora, it implies that these keywords are commonly used in research articles in the field of nursing. Therefore, it should be useful for nursing students as well as nursing practitioners to know these words. This also indicates the pedagogical importance in that the EAP and ESP instructors in the field of nursing could emphasize these words and raise students' awareness of these commonly found words in research articles of their field.

RQ 2: What are the lexical collocations of the keywords in the SCNRA? What is the proportion according to each type of combinations?

Once the keywords from the SCNRA had been identified, the process of identifying their collocates began. The identifying collocation pairs of the keywords

followed the set criteria, namely the word-span of 3 on the right side of the node (3R)

and the set frequency of co-occurrence as well as the mutual information value. Under

the set criteria, each keyword of four parts of speech can co-occur with different types

of words. The summary of the findings is shown in Table 4.6 below.

Table 4.6 Number of Lexical Collocations in SCNRA according to Types of combination

No.	Combination Types	No. of collocation pairs	(%)
1	Noun + Noun	889	41.39
2	Adjective + Noun	610	28.40
3	Noun + Verb	240	11.17
4	Verb + Noun	128	5.96
5	Noun + Adjective	84	3.91
6	Adjective + Adjective	82	3.82
7	Verb + Adjective	34	1.58
8	Verb + Verb	24	1.12
9	Adjective + Verb	13	0.61
10	Adverb + Verb	12	0.56
11	Adverb + Adjective	11	0.51
12	Verb + Adverb	10	0.47
13	Adverb + Noun	7	0.33
14	Noun + Adverb	4	0.19
	Total	**2148**	**100**

Table 4.6 above shows that there are 2,148 collocation pairs generated from

the keywords earlier generated. The majority of the collocation pairs is the 'Noun +

Noun' combinations with 889 pairs. The number of collocations generated under this

combination types are accounted for over 40 percent of all combination types. The

'Adjective + Noun' combination comes second with 610 pairs (28.4%). The 'Noun +

Verb' combination comes third with 240 collocation pairs (11.17%). The 'Noun +

Adverb' generates the least number of combinations at four pairs (0.19%). Examples

of collocation pairs of each combination type are shown in Table 4.7 below.

Table 4.7 Examples of Lexical Collocations extracted from SCNRA

Nodes	Collocates	Examples
Noun	Noun	care provider/s, health care, nurses (*and*) physicians, patients (*and, and their*) families, children (*with, diagnosed with*) cancer
	Verb	study (*was*) conducted, patients (*were*) admitted, children (*were*, *had been*) diagnosed, information (*was, could be*) provided
	Adjective	patients (*with*) stable, children (*with, diagnosed with*) chronic, use (*of*) antipsychotic, risk (*of, of developing*) adverse
	Adverb	parents (*of, of children*) newly, responses (*including, ranging from*) strongly
Adjective	Noun	mental (*ill, and physical*) health, social support, important role, clinical (*nursing*) practice, different types
	Adjective	social cognitive, physical (*and*) psychological, high (*and*) low, medical (*and, to a*) surgical, positive (*and*) negative
	Verb	important (*to*) note, significant (*difference was*) noted, ethical (*approval was*) obtained
Verb	Noun	reported feeling, associated (*with, with higher*) suicide, provided (*with, with adequate*) information, compared (*with, with other*) women
	Adjective	reported (*a, to be*) moderate, need (*to be*) aware, associated (*with, with a*) higher, provide (*a*) safe,
	Verb	use (*to*) measure, need (*to*) develop, stated (*they, that they*) know
	Adverb	described above, showed (*no*) statistically, viewed positively
Adverb	Verb	significantly associated, specifically designed, positively related
	Adjective	significantly (*associated with*) higher, statistically significant
	Noun	significantly (*higher, higher pain*) scores, approximately (-) minutes

Altogether, there are 14 combination types of lexical collocations found in the SCNRA under this investigation. This means that there are more combination types than the framework set by the present study adapted from that of Benson et al.'s (2010). Therefore, the findings shown in Table 4.6 above can be categorized into two groups: lexical collocations with combination types according to the set framework

and lexical collocations with combination types not according to the set framework. The former group consists of combination type numbers 1, 2, 3, 4, 11, and 12, while the latter group consists of combination type numbers 5, 6, 7, 8. 9, 10, 13, and 14. The former group comprises 1,888 collocation pairs accounted for 87.9%, the counterpart latter group consists of 260 collocation pairs accounted for 12.1%. These two groups of lexical collocations are shown in Table 4.8 and Table 4.9 below. The lists of these two groups of lexical collocations are shown in Appendix G and Appendix H respectively.

Table 4.8 Lexical Collocations according to the Set Framework

Combination type No.	Combinations	No. of collocation	(%)
1	Noun + Noun	889	47.09
2	Adjective + Noun	610	32.31
3	Noun + Verb	240	12.71
4	Verb + Noun	128	6.78
11	Adverb + Adjective	11	0.58
12	Verb + Adverb	10	0.53
	Total	**1888**	**100**

Table 4.9 Lexical Collocations Not according to the Set Framework

Combination type No.	Combinations	No. of collocation	(%)
5	Noun + Adjective	84	32.31
6	Adjective + Adjective	82	31.54
7	Verb + Adjective	34	13.08
8	Verb + Verb	24	9.23
9	Adjective + Verb	13	5.00
10	Adverb + Verb	12	4.62
13	Adverb + Noun	7	2.69
14	Noun + Adverb	4	1.54
	Total	**260**	**100**

When considering the collocations according to the framework of the study which has been adapted from that of Benson et al.'s (2010), it is found that the majority of the collocations gained from the present study are under the set framework. This high in number and percentage of the collocations indicates that the

collocations under the set framework are common combination types. For the collocations which are not according to the set framework, as shown in Table 4.9, there are eight combination types that belong to this group. There are altogether 260 collocation pairs and accounted for 12.1 percent of the total collocations generated from the SCNRA. These collocations constitute combination types which are less commonly found than those in the set framework. These uncommon combination types of collocations have been generated in the present study because of the set criteria for identifying collocations by which the collocates are identified within the range of 3 word-span on the right side (3R) of each node. Since wider word-span leads to more combination types to occur, therefore, with the set word-span in the present study, it allows more combination types to occur. Among these combination types, there are also the less common collocation pairs. These uncommon combination types are also not found and not recommended by scholars and researchers from previous studies in the field. This may be the reason they were excluded in the combination types suggested by Hausmann (1990), Hill (2000), and Benson et al. (2010), except the Adverb + Verb combination that is found suggested by Hill (2000).

When considering the keywords or nodes with their collocates, it is interesting to find out the nodes with the most collocates. The five keywords/nodes of each part of speech thus have been identified with the number of the collocates and parts of speech of the collocates that co-occur with them. The five noun keywords with the most collocates are: *health*, *nursing*, *risk*, *studies*, and *family* respectively. The five adjective keywords with the most collocates are: *physical*, *mental*, *social*, *each*, and *high*. The five verb keywords with the most collocates are: *used*, *diagnosed*,

associated, *received*, and *manage*. The five adverb keywords with the collocates are: *significantly*, *positively*, *negatively*, *approximately*, and *strongly*. The details of each type of keywords/nodes with the most collocates are presented in Tables 4.10 to 4.13 below.

Table 4.10 Five noun keywords with highest number of collocates

Keywords (No.)	Nouns	Verbs
health (25)	literacy, professionals, (*x*) provider/s, promotion, centres, check/s, professions, crisis/es, commission, service/s, insurance, facility/ies, complaints, (*x,xx*) recommendation, sciences, (*x,xx*) excellence, workforce, behaviours, (*x,xx*) care, status, problems, system/s, outcomes, practitioners, issues	-
nursing (23)	curricula, home/s, shortage, (*x*) placement, assistants, profession, pathway, student/s, discipline, educators, (*x*) environments, practice, workforce, notes, science, teamwork, staff, rounds, leaders, (*x*) adjustment, documentation, competence, interventions	-
risk (16)	(*x,xx*) ulceration, (*x*) infertility, (*x*) stunting, assessment, taking, (*xx*) defects, (*x,xx*) reduction, (x) tools, (*x,xx*) safety, factor/s, (*x,xx*) harm, (*x,xx*) developing, (*x,xx*) bias, management, behaviors, (*x,xx*) suicide	-
studies (15)	-	investigating, (*x,xx*) examined, examining, (*x*) shown, (x) investigated, (*x,xx*) explored, exploring, focusing, involving, (*x,xx*) conducted, show, (*x*) published, (*x,xx*) evaluated, reporting, (*x*) focused
family (13)	member/s, normalcy, harmony, advocate, (*x,xx*) friends, caregiver/s, functioning, empowerment, involvement, systems, strengths, dynamics, conversations	-

Table 4.11 Five adjective keywords with highest number of collocates

Keywords (No.)	Nouns	Verbs	Adjectives
physical (22)	activity, fitness, assault, disorder, performance, restraint, illnesses, functioning, (*x*) tests, abuse, function, (*x,xx*) health, (*x,xx*) problems, aggression, difficulties, condition/s, (*x,xx*) violence	-	(*x,xx*) psychosocial, (*x,xx*) psychological, (*x*) emotional (*x,xx*) sexual, (*x,xx*) cognitive

Keywords (No.)	Nouns	Verbs	Adjectives
mental (21)	(*x*) triage, (*x*) crisis/es, (*x*) commission, (*x,xx*) health, illness/es, (*x*) facilities, (*x*) service/s, (*x,xx*) workforce, (*x*) disorder/s, (*x*) practitioners, (*x*) practitioners, (*x*) teams, (*x*) simulation, (*x*) consumers, (*x*) problems, (*x*) issues, state, (*x*) settings, (*x*) professionals, (*x*) clinicians	-	-
social (21)	cohesion, gradient, science/s, capital, isolation, network/s, support, norms, worker/s, class, relations, (*xx*) contexts, media, (*x*) theory, activities, interaction/s, functioning	-	(*x,xx*) political, (*x,xx*) economic, cognitive, (*x*) cultural
each (20)	(*x*) transcript, (*x*) item, (*x,xx*) dimension, (*x,xx*) site, (*x*) country, (*x,xx*) indicator, subscale, (*x*) component, participant, (*x*) category, session, domain, (*x*) theme, (*x*) member, (*x*) source, (*x*) variable, (*x*) interview, (*x,xx*) question, year	(*x,xx*) rated	-
high (17)	secure, (*x,xx*) neuroticism, (*x*) turnover, (*x,xx*) extraversion, school, profile, level/s, workload, priority, (*x,xx*) rate/s, prevalence, (*x*) demands, degree, (*x,xx*) burnout, (*x*) score/s, reliability	-	(*x,xx*) low

Table 4.12 Five verb keywords with highest number of collocates

Keywords (No.)	Nouns	Verbs	Adjectives
used (9)	-	(*x*) analyse/ze, (*x*) compare, (*x*) collect, (*x*) assess, (*x*) measure, (*x*) examine, (*x*) describe, (*x*) evaluate	(*x*) cross-sectional
diagnosed (7)	(*x*) schizophrenia, (*x,xx*) cancer, (*x,xx*) diabetes, (*xx*) patients, families	-	(*x*) oncology, pediatric
associated (6)	-	-	(*x,xx*) decreased, (*x,xx*) lower, (*x,xx*) increased, (*x,xx*) higher, (*x,xx*) severe, (*x,xx*) greater
received (5)	(*x*) approval, (*x,xx*) attention, (*x,xx*) training, (*x,xx*) education, (*x,xx*) treatment	-	-
manage (5)	(*x,xx*) condition, (*x*) body, (*x*) child, (*x,xx*) health, (*x,xx*) care	-	-

161

Table 4.13 Five adverb keywords with highest number of collocates

Keywords (No.)	Nouns	Verbs	Adjectives
significantly (9)	(*x,xx*) scores	correlated, associated, (*x,xx*) related, increased	(*xx*) higher, (*xx*) lower, (*x*) likely, different
positively (3)	(*x,xx*) job	associated, related	-
negatively (3)	-	correlated, affect, associated	-
approximately (3)	(*x*) min(utes), half, (*xx*) people	-	-
strongly (2)	-	-	disagree/d, agree

As seen in Tables 4.10 to 4.13 above, it is noticeable that the noun nodes are more likely to co-occur with noun collocates and some possibility to co-occur with verb collocates. The adjective nodes are more common to take noun collocates with a high chance to co-occur with adjective collocates and some chance to co-occur with verb collocates. The verb nodes have higher possibility of co-occurrence with noun and adjective collocates and may co-occur with other verbs. The adverb nodes, although found in small number of collocation pairs, have slightly equal possibility to co-occur with noun, adjective, and verb collocates.

When ranking the lexical collocations according to the frequency of occurrence, it is noticeable that the majority of the 200 most frequent collocations (Appendix J) are in Adjective + Noun and Noun + Noun combinations. Therefore, it should be useful to compile the list of most frequent collocations of these two combination types. The list of 100 most frequent collocations of these two combination types are shown in Appendix K and Appendix L respectively. The prominent number of Adjective + Noun and Noun + Noun combinations is consistent with the study conducted by Biber et al. (2011) where they discover that complex noun phrases are prevalent features found in research articles.

When examining the examples of collocation pairs generated as shown in Tables above particularly in Table 4.7 and Table 4.10, it is noticeable that the majority of the verbs collocates of the noun nodes are those in the passive forms. For example, study (*was*) conducted, patients (*were*) admitted, children (*were*, *had been*) diagnosed, information (*was*, *could be*) provided. This could have pedagogical implication in the way that this linguistic feature should be aware of and pointed out to the learners.

Apart from classification based on types of combination, it is noticeable that the lexical collocations found from the SCNRA can also be categorized into two main groups: nursing specific collocations and general academic collocations. The categorization of these two groups has been arranged in reference to a rating scale for finding technical words designed by Chung and Nation (2003). The examples of these two groups of collocations are shown below in Tables 14.14 and 14.15 respectively.

Table 4.14 Nursing specific collocations

No.	Nodes	Collocates	No.	Nodes	Collocates
1	mental	(*ill*, *and physical*) health	26	risk	factor/s
2	health	care	27	pediatric	oncology
3	family	member/s	28	nursing	staff
4	health	service/s	29	patients	(*and*, *and their*) families
5	physical	(*ill*, *and mental*) health	30	parents	(*of*, *and their*) children
6	nursing	home/s	31	critical	care
7	mental	illness/es	32	health	problems
8	care	provider/s	33	health	status
9	service	user/s	34	patient	safety
10	palliative	(*and supportive*) care	35	health	outcomes
11	family	caregiver/s	36	health	literacy
12	social	support	37	acute	(*psychiatric*, *and primary*) care
13	physical	activity	38	intensive	(*follow-up*, *support and*) care
14	health	(*care*) provider/s	39	illness	belief/s
15	quality	(*of*, *of nursing*) care	40	childhood	cancer
16	mental	(*health*) service/s	41	chronic	(*disease*) condition/s
17	nursing	practice	42	psychological	distress
18	quality	(*of*) life	43	primary	(*family*) caregiver/s
19	patient	education	44	medication	administration
20	chronic	(*physical*, *conditions for*)	45	care	settings

21	risk	assessment	46	health	system/s
22	registered	nurse/s	47	health	issues
23	nurse	manager/s	48	depressive	symptoms
24	nursing	student/s	49	patient	satisfaction
25	clinical	(*nursing*) practice	50	anxiety	(*and*) depression

Table 4.15 General academic collocations

No.	Nodes	Collocates	No.	Nodes	Collocates
1	present	study	26	assessment	tool/s
2	data	collection	27	mean	age
3	focus	group/s	28	qualitative	(*research*) study/ies
4	previous	studies	29	control	group/s
5	significant	difference/s	30	research	question/s
6	higher	(*mean, and lower*) score/s	31	review	board/s
7	high	level/s	32	internal	consistency
8	mean	score/s	33	strongly	agree
9	sample	size/s	34	age	group/s
10	data	(*were*) collected	35	response	rate/s
11	age	(*of -, ranged between -*) years	36	age	(*and*) gender (*and socioeconomic*)
12	higher	level/s	37	demographic	characteristics
13	aged	(*under -*) years	38	everyday	life/ves
14	statistically	significant	39	each	(*questionnaire*) item
15	previous	research	40	work	environment/s
16	total	(*mean, health literacy*) score/s	41	aim	(*of this*) study
17	current	study	42	descriptive	statistics
18	study	(*was*) conducted	43	knowledge	(*and*) skill/s
19	research	team	44	marital	status
20	data	analysis	45	team	members
21	informed	consent	46	institutional	review
22	score/s	indicate/ed/ing	47	institutional	(*review*) board/s
23	systematic	review/s	48	content	analysis
24	inclusion	(*and exclusion*) criteria	49	limitation/s	(*of the*) study
25	participate	(*in the*) study	50	participants	(*were*) asked

Once the two groups of collocations have been clearly classified, in terms of pedagogical implications, this can greatly benefit both the learners and the teachers. With the clearer target collocations, the teachers should be able to design lessons with a clearer purpose by focusing on each group of collocations separately. Therefore, this should be easier for the learners to notice and recognize the collocations of the two groups.

4.2 Research Part II: Nursing Students' Lexical Collocation Knowledge

After the lexical collocations in the SCNRA had been identified, the Nursing Collocation Test have been constructed and administered with nursing students at SUT. The pre-test results are used as the indicator of the students' knowledge of lexical collocations being tested. The post-test, which was administered following the lexical collocation lessons are used in comparison with the pre-test's results to evaluate how corpus-based instruction affects their performance. The results gained from this part of the study are meant for answering Research Questions 3 and 4 of the present study.

4.2.1 Nursing Students' Collocation Knowledge: Pre-test results

With the pre-test results, the research question number three can be answered.

RQ 3: How much collocational knowledge do the fourth year nursing students at Suranaree University of Technology (SUT) have based on a test of lexical collocations extracted from the SCNRA?

The knowledge of nursing students on collocations found from the SCNRA is determined from the students' performance on the pre-test of the Nursing Collocations Test. The test was administered with the fourth year nursing students ($N = 51$) at SUT during trimester 1 of the academic year 2017. The summary of the test results in descriptive statistics is shown in Table 4.16 below.

Table 4.16 Pre-test's Results in Descriptive Statistics

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Prepart1 (30)	51	9.00	25.00	17.0392	3.96465
Prepart2 (20)	51	4.00	15.00	9.9804	3.01656
Prepart3 (10)	51	.00	8.50	3.6373	2.18879
PreTotal (60)	51	16.00	46.50	30.6569	7.41383

As the results shown in Table 4.16 above, in terms of the total score of 60, the mean score is just over 50 percent of the total score ($\bar{x} = 30.66$). The maximum score is 46.5, while the minimum score is 16. This results in a wide diversity of scores from the mean among the test takers ($SD = 7.41$). It also shows that the students in this group have different level of knowledge on the nursing collections being tested.

In terms of individual parts, part 1 includes the multiple-choice test of 30 items. The mean score is slightly over 50 percent of the total score of 30 ($\bar{x} = 17.04$). The maximum score is 25, while the minimum is 9. This indicates that there is less diverse of scores from the mean among the test takers ($SD = 3.96$). It shows the majority of the students have slightly different level of knowledge on the test, but less than that of the total score's.

For part 2 of the test, the gap-filling of 20 items, the mean score is just under 50 percent of the total score of 20 ($\bar{x} = 9.98$). The maximum score in this part is 15, while the minimum is 4. This leads to a slightly less diversity of scores from the mean of the group ($SD = 3.02$). It indicates that this group of students still have slightly different level of knowledge on the test, but less than that of the total score's and part 1 score's.

For part 3 of the test, 10 items of a sentence writing task, the mean score is a great deal under 50 percent of the total score of 10 ($\bar{x} = 3.64$). The maximum score is 8.5, while the minimum score is 0. However, the score of this part indicates less diversity of scores from the mean of this part ($SD = 2.19$). This also shows that the majority of the students have similarly low level of knowledge of this part.

In order to clearly demonstrate the students' performance on the pre-test, their scores both the total and each part were converted into percentage. Then the students

were grouped according to their scores in the five score range to give a clear picture on how they had performed on the pre-test as shown in Tables 4.17 to 4.20 below.

Table 4.17 Students' performance on the pre-test for the total score in percentage

Performance levels	Score ranges	No. of students	Percentage
Excellent	80 - 100	0	0
Good	60 - 79	13	25.49
Fair	40 - 59	**31**	**60.78**
Poor	20 - 39	7	13.73
Very poor	0 - 19	0	0
	Total	**51**	**100**

From the overall performance of the pre-test on the total score shown on Table 4.17, the performance of the majority of the students is ranked "*Fair*" (N = 31; 60.78 %). The scores of these students ranged from 40 to 59 on the scale of 100. The performance of the second group is "*Good*" (N = 13; 25.49 %). Their scores ranged from 60 to 79 on the scale of 100. The performance of another group is "*Poor*" (N = 7; 13.73 %). Their scores ranged from 20 to 39 on the scale of 100. None of the students' score could reach 80 and higher which is ranked "*Excellent*", the highest level of performance. Likewise, none of them scored under 20 on the scale of 100 which is "*Very poor*" and the worst level of performance. This indicates that most of these nursing students have a fair knowledge of nursing collocations. Some of them have good knowledge of nursing collocations, while some of them have poor knowledge of nursing collocations.

Table 4.18 Students' performance on Part 1 of the pre-test in percentage

Performance levels	Score ranges	No. of students	Percentage
Excellent	80 - 100	2	3.92
Good	60 - 79	**22**	**43.14**
Fair	40 - 59	**22**	**43.14**
Poor	20 - 39	5	9.80
Very poor	0 - 19	0	0.00
	Total	**51**	**100.00**

As seen in Table 4.18, the majority of students' performance on the pre-test for Part 1 in the five-level scale were divided into two groups ($N = 22$; 43.14% each). The results shows that students' performance equally fell under "*Good*" and "*Fair*" levels. The second group ($N = 5$; 9.8%) fell under "*Poor*", while the third group ($N = 2$; 3.92%) fell under "*Excellent*". There was no student performed under "*Very poor*" in this part.

Table 4.19 Students' performance on Part 2 of the pre-test in percentage

Performance levels	Score ranges	No. of students	Percentage
Excellent	80 - 100	0	0
Good	60 - 79	17	33.33
Fair	40 - 59	**24**	**47.06**
Poor	20 - 39	10	19.61
Very poor	0 - 19	0	0
	Total	51	100.00

Table 4.19 above shows students' performance in the five levels of performance scale for Part 2 of the pre-test, the majority of them ($N = 24$; 47.06%) fell under "*Fair*" level. The second group ($N = 17$; 33.33%) fell under "*Good*", while the third group ($N = 10$; 19.61%) fell under "*Poor*". There was no student performance ranked under either "*Excellent*" or "*Very poor*" level in this part.

Table 4.20 Students' performance on Part 3 of the pre-test in percentage

Performance levels	Score ranges	No. of students	Percentage
Excellent	80 - 100	3	5.88
Good	60 - 79	8	15.69
Fair	40 - 59	11	21.57
Poor	20 - 39	**18**	**35.29**
Very poor	0 - 19	11	21.57
	Total	51	100.00

As seen in Table 4.20 above, students' performance in the five levels of performance scale for Part 3 of the pre-test reveals that the majority of them ($N = 18$; 35.29%) fell under "*Poor*" level. The second group consisted of two different levels

(N = 11; 21.57% each) fell under "*Fair*" and "*Very poor*". The third group (N = 8; 15.69%) fell under "*Good*". There were a few students (N = 3; 5.88%) performed under "*Excellent*" in this part.

When examining students' performance on each part of the pre-test (as seen in Tables 4.18, 4.19, 4.20), it is noticeable that while there was no student taking part 1 and part 2 of the pre-test performed under the "*Very poor*" level. However, there were 11 students (21.57 %) in part 3 performed under this level, which is the worst in the five performance scale. This is understandable since writing, even though at a sentence level, which is a productive knowledge as opposed to a receptive knowledge, requires higher knowledge and skill to enable the learners to perform or produce the language (Jaén, 2007; Nation, 2001). Studies conducted by a number of scholars also confirm that EFL/ESL learners performed better in receptive tests than productive ones (Bueraheng & Laohawiriyanon, 2014; Khittikote, 2011; Torabian et al., 2014).

With the performance of the majority of the fourth year nursing students at the "*Fair*" level indicated by the overall pre-test scores and slightly lower on a sentence writing task, this shows that there is still a great deal of room for improvement. This is consistent with the study conducted by El-Dakhs (2015) that reveals the collocational competence of learners was notably unsatisfactory. A study conducted by Keshavarz and Salimi (2007) also suggests the importance of improving EFL/ESL learners' collocational knowledge to enhance their proficiency level. Ebrahimi-Bazzaz et al. (2012) also endorses that there is a high positive relationship between collocational competence and general language proficiency of learners. With his awareness that Croatian medical students still need to improve their collocational knowledge, Miščin (2013) studied most frequent mistakes students made and compiled a glossary of most

frequent verb collocations. Suwitchanphan and Phoocharoensil (2014) also suggest enhancing Thai learners' collocational competence, especially adjective-noun collocations.

4.2.2 Nursing Students' Collocation Knowledge: Post-test results

Following the administration of the pre-test and a workshop on nursing collocations extracted from the SCNRA, the post-test was administered to measure the effect of corpus-based instruction on students' performance. This also should answer the research questions number four of the present study.

RQ 4: How much does corpus-based instruction help improve the knowledge of collocations for the fourth year nursing students at SUT?

In order to find out how much corpus-based instruction of lexical collocations extracted from the SCNRA could help nursing students improve their performance, the post-test was administered aftermath of the workshop had been provided. The summary of the post-test results in descriptive statistics is shown in Table 4.21 below.

Table 4.21 Post-test's Results in Descriptive Statistics

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
Postpart1 (30)	51	12.00	29.00	20.4706	4.36968
Postpart2 (20)	51	8.00	20.00	13.3529	3.06479
Postpart3 (10)	51	1.00	9.00	5.6176	1.94059
PostTotal (60)	51	25.50	53.50	39.4412	7.65483

As shown in Table 4.21 above, in terms of the post-test total score of 60, the mean score is nearly two third of the total score ($\bar{x}$ = 39.44). The maximum score is 53.5, while the minimum score is 25.5. This, however, still results in a wide diversity of scores from the mean among the test takers (SD = 7.65). It also indicates that the students still have different level of knowledge on the nursing collections being tested.

In terms of individual parts, part 1 includes the multiple-choice test of 30 items. The mean score is slightly over two third of the total score of 30 ($\bar{x} = 20.47$). The maximum score is 29, while the minimum is 12. This indicates that there is less diverse of scores from the mean among the test takers ($SD = 4.37$). It shows the students have slightly different level of knowledge on the test, but less than that of the total score's.

For part 2 of the test, the gap-filling of 20 items, the mean score is well over 50 percent of the total score of 20 ($\bar{x} = 13.35$). The maximum score in this part is the full score of 20, while the minimum is 8. This leads to a slightly less diversity of scores from the mean of the group ($SD = 3.06$). It indicates that the students still have slightly different level of knowledge on the test, but less than that of the total score's and part 1 score's.

For part 3 of the test, 10 items of a sentence writing task, the mean score is just over 50 percent of the total score of 10 ($\bar{x} = 5.62$). The maximum score is 9, while the minimum score is 1. However, the score of this part indicates the least diversity of scores from the mean ($SD = 1.94$). This also indicates that the students have fairly similar level of knowledge of this part.

In order to find out how much the students' scores have improved as a result of the corpus-based instruction, the results of the post-test were then compared with those of the pre-test. The list of raw scores of the pre-test and the post-test is shown in Appendix N. In terms of statistical differences between the pre-test and the post-test, a paired samples t-test's results are shown in Table 4.22 below.

Table 4.22 The statistical differences between the Pre-test and the Post-test

Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	
Pair 1	Prepart1 - Postpart1	-3.43137	3.28180	.45954	-4.35439	-2.50835	-7.467	50	.000
Pair 2	Prepart2 - Postpart2	-3.37255	2.97295	.41630	-4.20870	-2.53639	-8.101	50	.000
Pair 3	Prepart3 - Postpart3	-1.98039	1.89990	.26604	-2.51475	-1.44604	-7.444	50	.000
Pair 4	PreTotal - PostTotal	-8.78431	5.34065	.74784	-10.28640	-7.28223	-11.746	50	.000

In Table 4.22 above, the results of Pair 1 show that there is a significant difference in the mean scores of part 1 of the test between the pre-test and the post-test; $t(50) = (-7.47)$, $p = <0.001$. Similarly in Pair 2, there is a significant difference in the mean scores of part 2 of the test comparing the pre-test and the post-test; $t(50) = (-8.10)$, $p = <0.001$. Pair 3 also shows a significant difference in the mean scores of part 3 of the test in comparison between the pre-test and the post-test; $t(50) = (-7.44)$, $p = <0.001$. There is also a significant difference in the mean scores of the pre-test and the post-test in terms of the total score; $t(50) = (-11.75)$, $p = <0.001$. This shows that there has been a significant improvement of students' performance on the test.

In order to examine students' performance based on the results of both the pre-test and the post-test, the raw scores were converted into the total score of 100. Then the students were grouped according to their scores in the five score range to give a clear picture on how they had performed on both the pre-test and the post-test. For the pre-test, this can indicate the level of their knowledge of nursing collocations prior to the teaching or the providing of the workshop on nursing collocations. For the post-

test, this can compare to that of the pre-test in examining whether teaching is necessary or not for collocations.

For the overall test or the total score of 60, the distribution of the students' performance in the five score range from 0 to 100 is shown in Table 4.23 below.

Table 4.23 Students' performance in comparisons between the Pre-test and the Post-test for the total score in percentage

Total score		Pre-test		Post-test		Changed
Performance	Score ranges	N	%	N	%	%
Excellent	80 - 100	0	0	10	19.61	19.61
Good	60 - 79	13	25.49	**21**	**41.18**	15.69
Fair	40 - 59	**31**	**60.78**	20	39.22	-21.57
Poor	20 - 39	7	13.73	0	0.00	-13.73
Very poor	0 - 19	0	0	0	0	0.00
	Total	**51**	**100**	**51**	**100**	

Table 4.23 above shows that there has been an improvement of students' performance. The majority of students (N = 21, 41.18 %) gained satisfactory results under the score range of 60 to 79 at the "*Good*" level with over 15 percent increase in number. More importantly, there were 10 students (19.61 %) whose results were under the score range of 80 to 100 at the "*Excellent*" level from which no students attained this level of performance in the pre-test. There were 20 students (39.22 %) whose results were under the score range of 40 to 59 indicating "*Fair*" level of performance which has decreased over 21 percent comparing to that of the pre-test's. None of the students scored under 40 out of 100 in the post-test.

For part 1 of the test, the multiple-choice test, the distribution of students' performance of both the pre-test and the post-test in the five score range is shown in Table 4.24 below.

Table 4.24 Students' performance in comparisons between the Pre-test and the Post-test for the Part 1 score in percentage

Part 1		Pre-test		Post-test		Changed
Performance	Score ranges	N	%	N	%	%
Excellent	80 - 100	2	3.92	13	25.49	21.57
Good	60 - 79	**22**	**43.14**	**23**	**45.10**	1.96
Fair	40 - 59	**22**	**43.14**	15	29.41	-13.73
Poor	20 - 39	5	9.80	0	0.00	-9.80
Very poor	0 - 19	0	0.00	0	0.00	0.00
	Total	**51**	**100**	**51**	**100**	

In Part 1 (multiple-choice) of the test, there has been an improvement of students' performance. Firstly, there has been over 21 percent increase of students who scored at "*Excellent*" level from only two students in the pre-test to 13 students in the post-test. Secondly, there has also been an increase in number of students who scored at "Good" level, from 22 to 23 students. Thirdly, the number of students who scored under "Fair" level has decreased over 13 percent from 22 students in the pre-test to 15 in the post-test. Finally, in the post-test, there was no student who scored under "Poor" and "Very poor" levels, while there were five students in the pre-test.

For part 2 of the test, the gap-filling format, the distribution of students' performance of both the pre-test and the post-test in the five score range is shown in Table 4.25 below.

Table 4.25 Students' performance in comparisons between the Pre-test and the Post-test for the Part 2 score in percentage

Part 2		Pre-test		Post-test		Changed
Performance	Score ranges	N	%	N	%	%
Excellent	80 - 100	0	0	13	25.49	25.49
Good	60 - 79	17	33.33	**21**	**41.18**	7.84
Fair	40 - 59	**24**	**47.06**	17	33.33	-13.73
Poor	20 - 39	10	19.61	0	0.00	-19.61
Very poor	0 - 19	0	0	0	0	0.00
	Total	**51**	**100**	**51**	**100**	

In this part, there has also been a significant improvement in students' performance. Firstly, there has been over 25 percent increase in the number of students whose score could reach the "*Excellent*" level, while there was no student achieving such as high level in the pre-test. Secondly, there has been over 7 percent increase in number of students who scored under "*Good*" level, from 17 students did in the pre-test to 21 in the post-test. Thirdly, the number of students who scored under "*Fair*" level has decreased over 13 percent, from previously in the pre-test of 24 students to 17 in the post-test. Finally, there was no student scored under either "*Poor*" or "*Very poor*" in the post-test, while there were 10 students accounted for nearly 10 percent who did it in the pre-test.

For part 3 of the test, the sentence writing part, the distribution of students' performance of both the pre-test and the post-test in the five score range is shown in Table 4.26 below.

Table 4.26 Students' performance in comparisons between the Pre-test and the Post-test for the Part 3 score in percentage

Part 3		Pre-test		Post-test		Changed
Performance	Score ranges	N	%	N	%	%
Excellent	80 - 100	3	5.88	8	15.69	9.80
Good	60 - 79	8	15.69	15	29.41	13.73
Fair	40 - 59	11	21.57	**20**	**39.22**	17.65
Poor	20 - 39	**18**	**35.29**	6	11.76	-23.53
Very poor	0 - 19	11	21.57	2	3.92	-17.65
	Total	**51**	**100**	**51**	**100**	

In this part, the post-test results show a positive change in the students' performance as more students achieved higher scores. Firstly, at "*Excellent*" level, there has been nearly 10 percent increase of the number of students whose score could achieve this level, from only three students in the pre-test to eight in the post-test. Secondly, there has also been a positive change at the "*Good*" level with over 13

percent increase in number of students who could attain scores in the post-test to reach this level, from eight to 15 students. Thirdly, there has been over 17 percent increase in number of students who scored at the "*Fair*" level, from 11 students to 20 students. Fourthly, there has been a significant decrease in number of students who scored under the "*Poor*" level, over 23 percent from 18 students in the pre-test to only 6 in the post-test. Finally, there has also a considerable decrease in number of students who scored under "*Very poor*" level, over 17 percent from previously 11 students in the pre-test to only two in the post-test.

Since the nursing students' performance on the pre-test revealed that improvement is needed, it is interesting to find out how a workshop on lexical collocations found from the SCNRA could affect students' performance. With the time restraint coupled with the tight schedule of the students, the workshop introduced students with the 100 most frequent Noun + Noun, Adjective + Noun, 50 most frequent Noun + Verb, Verb + Noun, and 200 most frequent collocations both with repeated and not repeated nodes. The pre-test and post-test were also administered before and after the workshop. After the workshop was provided and the post-test was administered, the test's results revealed the statistically significant improvement on the students' scores on both the total score and on each part of the test as shown on the paired-samples t-test in Table 4.22.

With statistically significant improvement of the students revealed in the comparison of the pre-test's and post-test's results as well as the level of performance compared on the five levels on the scale of 100, teaching of collocations has positive effects on EFL/ESL learners and should be put into wider practice. This is consistent with a number of studies such as studies conducted by Eidian et al. (2013); Khonamri

and Roostaee (2014); Ördem (2013); Pirmoradian and Tabatabaei (2012); Shooshtari and Karami (2013); Szudarski (2012); and Usen and Musigrungsi (2015). It is evident that corpus-based instruction of collocations positively affects students' performance. As collocations are one useful element in achieving natural and native-like level of language performance, corpus-based instruction could be encouraged and put into practice.

Awareness raising, as suggested by Vasiljevic (2014), is an important tool that can facilitate the knowledge of collocations. The notion of noticing recommended by Skehan (1998) is also another productive strategy for language learning in general not particularly for collocations. Corpus-based instruction is one effective instrument to enable learners to notice the features of the language being taught. With guidance through corpus-based instruction with the application of three Is suggested by McEnery et al. (2006) and providing explanation, showing examples, or other means, coupled with learners' cognitive processes, the learners are more likely to be able to notice the features of the language being taught. Once noticed and perceived the targeted language features, those consciously learned features are then kept in memory system to be retrieved for use later when the occasion arises (McCarthy, 1990). This memory system is also called "mental lexicon" (Krashen, 1987; Hulstijn, 2000; Lewis, 2002b; Takač, 2008).

4.3 Summary of the Chapter

This chapter presented the findings of the present study. The findings of the research part I revealed the lexical collocations generated from the Sample Corpus of Nursing Research Articles (SCNRA) as well as the classifications of those

collocations found. It also discussed regarding the keywords extracted which were prevalently the nouns, adjectives, and verbs. The collocates of those keywords, which were used as nodes, were, hence, predominantly with Noun + Noun, Adjective + Noun, and Noun + Verb. These combination types of the collocations were also within the set framework adapted from that of Benson et al.'s (2010). There were some uncommon combination types which were not in accordance with the set framework. This phenomenon occurred as a result of the word-span for the collocates, which was three words from the right side of the node (3R). The wider the word-span allowed for the collocates, the more chances for uncommon combination types were to occur.

The findings of the research part II showed the results of the Nursing Collocation Test in comparison between the pre-test and the post-test. The students' performance had also been examined as to assess the knowledge of the students on nursing collocations found from the study. The discussions of the findings were provided concerning the collocational knowledge of the fourth year nursing students at SUT as revealed by the pre-test. With the level of performance gained, which is in the middle of the scale of five, there was a room for improvement. The students' performance on the post-test administered aftermath of the workshop revealed that their performance had significantly improved. This suggests that corpus-based instruction is beneficial for EFL/ESL learners to improve students' performance and awareness of collocations.

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

This final chapter provides the conclusion of the present study, the pedagogical implications gained from the present study, and the recommendations for further research.

5.1 Conclusion of the Present Study

As it is crucial for professional nurses to be efficient in English for both their professional advancement and further study, the present study has attempted to facilitate the learning and teaching of ESP in the field of nursing by filling the gap that prevents effective and natural use of English, collocations. Therefore, the present study was conducted (1) to identify and classify keywords found in the Sample Corpus of Nursing Research Articles (SCNRA) published in international journals in the field of nursing; (2) to explore lexical collocations extracted from the SCNRA using keywords found as 'nodes' to find their 'collocates', and to classify collocations produced according to their combinations; (3) to assess collocation knowledge of the fourth year nursing students at SUT based on the collocations identified from the SCNRA; and (4) to provide lessons on nursing collocations and assess the effects of corpus-based instruction. In order to achieve the mentioned objectives, the present study comprised the following research questions.

1) What are the keywords in the SCNRA based on the frequency of occurrence at ≥ 50 and the keyness value at ≥ 20? What is the proportion according to their parts of speech?

2) What are the lexical collocations of those keywords in the SCNRA? What is the proportion according to each type of combinations?

3) How much collocational knowledge do the fourth year nursing students at Suranaree University of Technology (SUT) have based on a test of collocations extracted from the SCNRA?

4) How much does corpus-based instruction help improve the knowledge of collocations for the fourth year nursing students at SUT?

The study consisted of two main parts: 1) identification and classification of lexical collocations from the SCNRA, and 2) collocational knowledge of fourth year nursing students at SUT. The present study started with the research part 1 by compiling 300 research articles from 10 selected academic journals in the field of nursing accessible online via SUT's Library Resources. The SCNRA which comprised over 1.25 million words generated 717 keywords under the set criteria. Among these keywords found, the majority were the nouns ($N = 463$; 63.51%), followed by the adjectives (N = 157; 21.54%), the verbs ($N = 98$; 13.44%) and the adverbs ($N = 11$; 1.51%). These keywords then were used as the "nodes" to find their collocates with 3 word-span on the right side (3R). 2,148 pairs of lexical collocations were identified with 14 combination types. The majority of the collocation pairs found were Noun + Noun ($N = 889$; 41.39%), followed by Adjective + Noun ($N = 610$; 28.4%), and Noun + Verb ($N = 240$; 11.17%). Among these 14 combination

types, the majority of them were in accordance with the set framework of six combination types adapted from Benson et al. (2010).

Once the lexical collocations had been identified and classified, the research part 2 was able to be commenced. In order to assess collocational knowledge of the fourth year nursing students at SUT, a Nursing Collocation Test was constructed. A list of 200 most frequent collocations identified from the SCNRA was used to construct the test. To avoid confusion, these 200 pairs of collocations were chosen from those of the highest frequency of each node. These 200 pairs of lexical collocations were then randomly selected for 60 test items. These 60 test items consisted of three parts: 30 items of a multiple-choice test; 20 items of a gap-filling test; and 10 items of a sentence writing task. The tryout of the test was carried out with 38 fourth year nursing students during trimester 3 of the academic year 2016. The tryout test was examined for the Difficulty Index, Discrimination Index, and Reliability Index. The improvement of the test had been made accordingly to the analysis.

The Nursing Collocation Test was administered with 51 fourth year nursing students at SUT during trimester 1 of the academic year 2017 before and after a 12 hour-workshop on nursing collocations organized for them. The pre-test's mean of the total score was 30.66 with the standard deviation of 7.41 ($\bar{x} = 30.66$; $SD = 7.41$). This revealed that the overall knowledge of collocations of the majority of the students was at a "*Fair*" level. The post-test's mean of the total score was 39.44 with the standard deviation of 7.65 ($\bar{x} = 39.44$; $SD = 7.65$). To evaluate the effects of corpus-based instruction on the students' performance, the results of the pre-test and the post-test were compared using the paired samples t-test. The results showed that there were

statistically significant improvements of the students' performance in the total score ($t(50) = (-11.75)$, $p = {<}0.001$) as well as in each of the three parts of the test: part 1 ($t(50) = (-7.47)$, $p = {<}0.001$); part 2 ($t(50) = (-8.10)$, $p = {<}0.001$); and part 3 ($t(50) = (-7.44)$, $p = {<}0.001$). In terms of the students' performance, the majority of the students' scores increased and put them one level higher in all parts from "*Fair*" to "*Good*", except for part 3, a sentence writing task, the level of performance was improved from "*Poor*" to "*Fair*".

The present study has added more resources to the ESP as well as EAP and EFL/ESL teaching and learning by means of producing both keyword list and lexical collocations. It has opened up and widened the road to success for nursing students and nursing professionals who seek means for effective and natural communication in their profession. The present study also has specifically filled the gap or at least narrowed down the gap in terms of collocations in the field of nursing both locally in Thailand and beyond in raising the awareness of collocations through the collocation test and corpus-based instruction.

5.2 Pedagogical Implications from the Present Study

As the present study has examined lexical collocations in the field of nursing through the building of a sample corpus of nursing research articles and investigated the knowledge of the fourth year nursing students at SUT, there are a number of pedagogical implications emerged from the present study. These implications are as follows:

Firstly, applying a corpus tool or concordance tool in extracting collocations from teaching and learning materials such as textbooks and related documents can be

beneficial in ESP such as the field of nursing as presented in the present study, as well as in EAP and EFL/ESL teaching and learning in general. By extracting collocations from materials used in each course of particular field of study and classifying as well as making the list of the collocations, this should make it clearer and easier for both the instructors and the learners to notice those collocations. This also is a great way to raise awareness of collocations among the EFL/ESL learners.

Secondly, corpus-based instruction of collocations is beneficial in helping EFL/ESL learners to be aware of the target collocations. With guidance from teachers, coupled with proper teaching and learning materials and exercises, the learners can gain a clear idea about collocations, their usefulness, and how the pairs occur in real contexts.

Thirdly, incorporating a corpus-based teaching and learning of collocations can also be beneficial in the teaching and learning of English from general EFL/ESL to EAP and ESP. By showing the concordance lines from a corpus built from materials used in the course such as textbooks and related documents, the learners can see the target collocations in the real use and authentic contexts (Gaskell & Cobb, 2004; Yoon & Hirvela, 2004; Kırkgöz, 2006; Lee & Swales, 2006; Wang et al., 2015). This should encourage the correct and natural use of the collocations.

Finally, testing students' collocational knowledge can be a helpful means in examining how much students know the target collocations. This is useful information for teachers to introduce students to those collocations, as well as how much attention is needed for particular pairs of collocations. The post-test or the test after collocations being taught also helps the teachers in assessing how much the students'

performance have improved after being taught, as well as how effective the lessons and the teaching methods are when being delivered by the teachers.

5.3 Recommendations for Further Research

Firstly, as the present study has focused solely on lexical collocations, it is recommended that a further research could be conducted by extending the focus to grammatical collocations. This could be done either by only focusing on the grammatical collocations or together with the lexical collocations so that the comparisons between the two could be made.

Secondly, as the present study has focused on the field of nursing, further research could be conducted by focusing on a variety of other fields which have not been conducted. This is to make collocations in other academic fields be more accessible and to facilitate the teaching and learning of collocations in those fields.

Thirdly, as research articles from academic journals have been used in the present study, further research could be conducted using other types of materials such as textbooks, magazines, news' columns, and the like. This is to make collocations available from various genres.

Finally, as the present study applied the whole IMRD parts of the research articles, further research could be conducted by separating the four different parts and comparing the results gained from those different parts of the research articles. This is to examine the similarities and differences of collocations of the same genre for their different parts.

5.4 Summary of the Chapter

This chapter has presented the conclusion of the present study. In the first part, it summarized the whole research study from the beginning to the end. In the second part, it suggested the pedagogical implications gained from the present study. Finally, it recommended the possibilities for further research which could be extended from the present study.
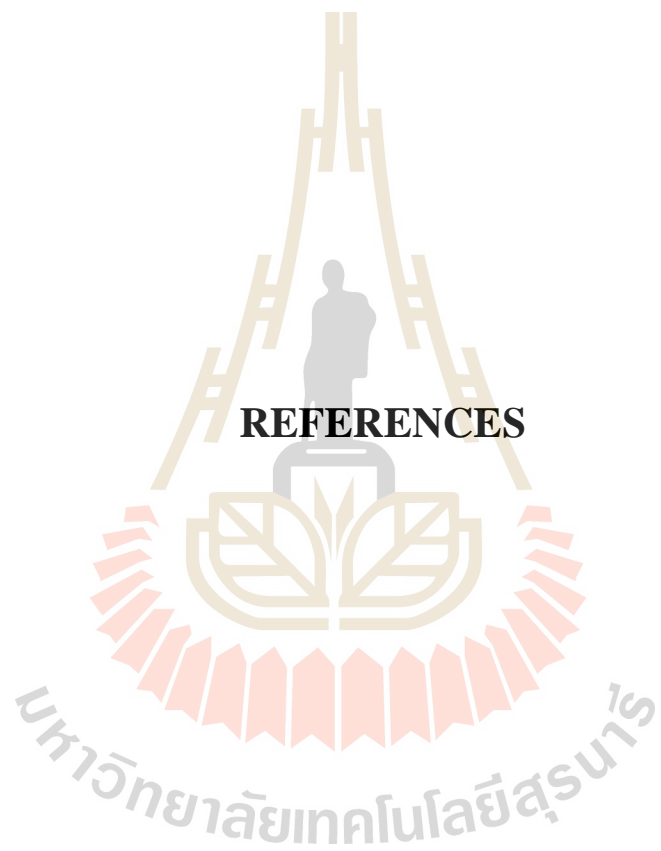
REFERENCES

REFERENCES

Abedi, Z., & Mobaraki, M. (2014). The effect of grammatical collocation instruction on understanding ESP texts for undergraduate computer engineering students. *Journal of Language Teaching and Research, 5*(3), 631-641. doi:10.4304/jltr.5.3.631-641

Ackermann, K., & Chen, Y.-H. (2013). Developing the academic collocation list (ACL) - A corpus driven and expert-judged approach. *Journal of English for Academic Purposes, 12*, 234-247.

Ahour, T., Rasoulizadeh, M., & Behnam, B. (2013). Analysis of the "Gone with the Wind" and its simplified version in terms of lexical structure. *International Journal of Applied Linguistics & English Literature, 2*(4), 39-49.

Aisenstadt, E. (1981). Restricted collocations in English lexicology and lexicography. *ITL: Review of Applied Linguistics, 53*, 53-61.

Akpınar, K. D., & Bardakçı, M. (2015). The effect of grouping and presenting collocations on retention. *TESL-EJ, 18*(4), 1-22.

Aldaba, R. M. (2013). *ASEAN economic community 2015: Labor mobility and mutual recognition arrangements on professional services*. The PIDS Discussion Paper Series No. 2013-04. Philippine Institute for Development Studies: Makati City, Philippines. Retrieved from http://www.pids.gov.ph

Alexander, M. (2011). Organizing the article. In C. Saver (Ed.), *Anatomy of writing for publication for nurses* (pp. 66-82). Indianapolis, IN: Sigma Theta Tau International.

Ali Zarei, A., & Baniesmaili, N. (2010). Learnability of various patterns of lexical collocations by Iranian upper-intermediate learners of English. *IJAL, 13*(2), 147-169.

Allerton, D. J. (1984). Three (or four) levels of word cooccurrence restriction. *Lingua, 63*, 17-40.

Anthony, L. (2014). *AntConc (Version 3.4.4)* [Computer Software]. Tokyo, Japan:Waseda University. Available from http://www.laurenceanthony.net/.

Arnaud, P. J. L., & Savignon, S. J. (1997). Rare words, complex lexical units and the advanced learner. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp.157-173). Cambridge, UK: Cambridge University Press.

Ary, D., Jacobs. L. C., & Sorensen, C. (2010). *Introduction to research in education* (8th ed.). Canada: Wadsworth.

Ashouri, S., Arjmandi, M., & Rahimi, R. (2014). The impact of corpus-based collocation instruction on Iranian EFL learners' collocation learning. *Universal Journal of Educational Research, 2*(6), 470-479.

Association of Southeast Asian Nations. (2008). *The ASEAN Charter*. Retrieved from http://www.asean.org/archive/publications/ASEAN-Charter.pdf

Bachman, L.F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford; UK: Oxford University Press.

Bahns, J. (1993). Lexical collocations: a contrastive view. *ELT Journal, 47*(1), 56-63.

Bahns, J. & Eldaw, M. (1993). *Should we teach EFL students collocations?* Paper given at 9th World Congress of Applied Linguistics Thessaloniki.

Bal, B. (2010). *Analysis of four-word lexical bundles in published research articles written by Turkish scholars*. Thesis, Georgia State University, 2010. Available at http://scholarworks.gsu.edu/alesl_theses/2

Balcı, Ö., & Çakır, A. (2012). Teaching vocabulary through collocations in EFL classes: The case of Turkey. *International Journal of Research Studies in Language Learning, 1*(1), 21-32.

Bardovi-Harlig, K., Mossman, S., & Su, Y. (2017). The effect of corpus-based instruction on pragmatic routines. *Language Learning & Technology, 21*(3), 76-103. Retrieved from http://llt.msu.edu/issues/october2017/bardovi-harligmossmansu.pdf

Barfield, A., & Gyllstad, H. (Eds.), (2009). *Researching collocations in another language*. Houndmills, UK: Palgrave Macmillan.

Barker, F. (2004). Using corpora in language testing. *Modern English Teacher, 13*(2), 63-67.

Baker, M. (1999). The role of corpora in investigating the linguistic behaviour of professional translators. *International Journal of Corpus Linguistics, 4*(2), 281-298.

Baker, P. (2006). *Using corpora in discourse analysis*. London, UK: Continuum.

Baker, P., & McEnery, T. (2005). A corpus-based approach to discourses of refugees and asylum seekers in UN and newspaper texts. *Journal of Language and Politics, 4*(2), 197-226.

Barbieri, F. (2005). Quotative use in American English: A corpus-based, cross-register comparison. *Journal of English Linguistics, 33*(3), 222-256.

Barbieri, F. (2007). Older men and younger women: A corpus-based study of quotative use in American English. *English World-Wide, 28*(1), 23-45.

Barbieri, F. (2009). Quotative be like in American English: Ephemeral or here to stay? *English World-Wide, 30*(1), 68-90.

Barlow, M. (2000). *MonoConc Pro (Version 2.2)* [Computer Software]. Available from http://www.athel.com/mono.html.

Bartsch, S. (2004). *Structural and functional properties of collocations in English: A corpus study of lexical and pragmatic constraints on lexical co-occurrence.* Germany: Gunter Narr Verlag Tübingen.

Belz, J., & Vyatkina, N. (2008). The pedagogical mediation of a developmental learner corpus for classroom-based language instruction. *Language Learning & Technology, 12*(3), 33-52.

Bennett, G. R. (2010). *Using corpora in the language learning classroom: Corpus linguistics for teachers.* Ann Arbor, MI: Michigan University Press.

Benson, M. (1989). The structure of the collocational dictionary. *International Journal of Lexicography, 2*, 1-14.

Benson, M., Benson, E. & Ilson, R. (2010). *The BBI Combinatory Dictionary of English: Your guide to collocations and grammar* (3rd ed.). Amsterdam, The Netherlands: John Benjamins.

Bettina, F. (2009). Keywords and frequent phrases of Jane Austen's Pride and Prejudice: A corpus-stylistic analysis. *International Journal of Corpus Linguistics, 14*(4), 492-523.

Biber, D. (2011). Corpus linguistics and the study of literature: Back to the future? *Scientific Study of Literature, 1*(1), 15-23.

Biber, D., & Conrad, S. (2010). Corpus Linguistics and Grammar Teaching. In *Pearson Longman English Language Teaching Newsletter*. Retrieved from http://longmanhomeusa.com/content/pl_biber_conrad_ monograph5_lo.pdf

Biber, D., Conrad, S., & Reppen, R. (2006). *Corpus linguistics: Investigating language structure and use*. Cambridge, UK: Cambridge University Press.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly, 45*(1), 5-35.

Biber, D., & Reppen, R. (2012). Editors' introduction: Corpus linguistics. In D. Biber and R. Reppen (Eds.), *Corpus linguistics, volume I, Lexical studies* (pp. xix-xiviii). London, UK; SAGE.

Binger., J. L. (1981). The nursing journal - learning resource, profession symbol, and commodity. *Image: Journal of Nursing Scholarship, 13*(10), 67-70.

Björk, B-C., Roos, A., & Lauri, M. (2009). Scientific journal publishing: Yearly volumeand open access availability. *Information Research, 14*(1) paper 391. Available from 12 January, 2009 at http://InformationR.net/ir/14-1/paper391.html.

Blackwell, N., & Fox Tree, J. E. (2012). Social factors affect quotative choice. *Journal of Pragmatics, 44*, 1150-1162.

Breyer, Y. (2009). Learning and teaching with corpora: Reflections by student teachers. *Computer Assisted Language Learning, 22*(2), 153-172.

Bueraheng, N., & Laohawiriyanon, C. (2014). Does learners' degree of exposure to English language influence their collocational knowledge? *International Journal of English and Literature, 4*(3), 1-10.

Cargill, M., & O'Connor, P. (2013). *Writing Scientific Research Articles: Strategy and Steps*. (2nd ed.). West Sussex, UK: Wiley-Blackwell.

Carter, R. (1998). *Vocabulary: Applied linguistic perspectives*. (2nd ed.). London, UK: Routledge.

Chambers, A. & O'Sullivan, Í. (2004). Corpus consultation and advanced learners' writing skills in French. *ReCALL, 16*(1), 158-172.

Charles, M. (2012). Proper vocabulary and juicy collocations': EAP students evaluate do-it-yourself corpus-building. *English for Specific Purposes, 31*(2), 93-102.

Charles, M. (2014). Getting the corpus habit: EAP students' long-term use of personal corpora. *English for Specific Purposes, 35*, 30-40.

Cheikhrouhou, H. (2014). Recognition of communication verbs with NooJ. In S. Koeva, S. Mesfar, and M. Silberztein (Eds.), *Formalising natural languages with NooJ 2013: Selected papers from the NooJ 2013 International Conference* (pp. 155- 170). Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

Cheng, W. (2010). What can a corpus tell us about language teaching? In A. O'Keeffe and M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 319-332). London, UK: Routledge.

Cheng, W. (2012). *Exploring corpus linguistics: Language in action*. London, UK: Routledge.

Chung, T. M., & Nation, P. (2003). Technical vocabulary in specialised texts. *Reading in a Foreign Language, 15*(2), 103-116.

Church, K. W., & Hanks, P. (1989). Word association norms, mutual information, and lexicography. In *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics* (pp. 76-83). Vancouver, Canada.

Conrad, S. (2002). Corpus linguistic approaches for discourse analysis. *Annual Review of Applied Linguistics, 22*, 75-95.

Cotterill, J. (2010). How to use corpus linguistics in forensic linguistics. In A. O'Keeffe and M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp.578-590). London, UK: Routledge.

Coulthard, M. (1994). On the use of corpora in the analysis of forensic texts. *International Journal of Speech Language and the Law, 1*(1), 27-34.

Cowie, A,, & Howarth, P. (1995). *Phraseological competence and written proficiency, paper presented at the British Association of Applied Linguistics Conference (BAAL).* Southampton, UK.

Coxhead, A. (2010). What can corpora tell us about English for Academic Purposes? In A. O'Keeffe and M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 458-470). London, UK: Routledge.

Crystal, D. (1995). *The Cambridge encyclopedia of the English language*. Cambridge, UK: Cambridge University Press.

Currie, J., & Chipps, J. (2015). Mapping the field of military nursing research 1990-2013: A bibliometric review. *International Journal of Nursing Studies, 52*, 1607-1616.

Dash, N. S. (2009). *Language corpora: Past, present and future*. New Delhi, India: Mittal Publications.

DePoy, E., & Gitlin, L. N. (2011). *Introduction to research: Understanding and applying multiple strategies* (4th ed.). St. Louis, MO: C. V. Mosby.

Deroey, K., L. B. (2011). What they highlight is…: the discourse functions of basic wh-clefts in lectures. *Journal of English for Academic Purposes, 11*(2), 112-124. doi:10.1016/j.jeap.2011.10.002

Díaz-Membrives, M., Farrero-Muñoz, S., & Lluch-Canut, M. T. (2012). Characteristicsof nursing publications in journals with impact factor. *Enfermeria Clinica,22*(5), 247-254. doi: 10.1016/j.enfcli.2012.06.006

Duan, M., & Quin, X. (2012). Collocation in English teaching and learning. *Theory and Practice in Language Studies, 2*(9), 1890-1894.

Dudley-Evans, T., & St John, M. J. (1998). *Developments in English for Specific Purposes: A multi-disciplinary approach*. Cambridge, UK: Cambridge University Press.

Durrant, P. (2009). Investigating the viability of a collocation list for students of English for academic purposes. *English for Specific Purposes, (28)*, 157-169.

Ebrahimi-Bazzaz, F., Samad, A. A., bin Ismail, I. A., & Noordin, N. (2012). Measuring collocational competence of Iranian learners by using C-test. *The Iranian EFL Journal, 2*, 227-240.

Eidian, F., Gorjian, B., & Aghvami, F. (2013). The impact of lexical collocation instruction on developing writing skill among Iranian EFL learners. *International Journal of Language Learning and Applied Linguistics World, 4*(3), 273-283.

El-Dakhs, D. A. S. (2015). The lexical collocational competence of Arab undergraduate EFL learners. *International Journal of English Linguistics, 5(*5), 60-74.

Englander, K. (2014). *Writing and publishing science research papers in English: A global perspective*. New York, NY: Springer.

Farr, F. (2008). Evaluating the use of corpus-based instruction in a language teacher education context: perspectives from the users. *Language Awareness, 17*(1), 25-43.

Farrokh, P. (2012). Raising awareness of collocation in ESL/EFL classrooms. *Journal of Studies in Education, 2*(3), 55-74.

Fernando, C. (1996). *Idioms and idiomaticity*. Oxford, UK: Oxford University Press.

Fillmore, C. J., & Atkins S. (1994). Starting where dictionaries stop: the challenge of corpus-lexicography. In B. T. S. Atkins and A. Zampolli (Eds.), *Computational approaches to the lexicon* (pp. 349-393). Oxford, UK: Oxford University Press.

Firth, J. R. (1957). *Papers in linguistics, 1934-1951*. Oxford, UK: Oxford University Press.

Flowerdew, L. (2002). Corpus-based analyses in EAP. In J. Flowerdew (Ed.). *Academic discourse* (pp. 95-114). London, UK: Longman.

Flowerdew, L. (2004). The argument for using English specialised corpora to understand academic and professional settings. In U. Connor & T. Upton (Eds.). *Discourse in the professions: Perspectives from corpus linguistics* (pp. 11-33). Amsterdam: John Benjamins.

Flowerdew, L (2010). Using corpora for writing instruction. In A. O'Keeffe and M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 444-457). London, UK: Routledge.

Forbes, A. (2009). Clinical intervention research in nursing. *International Journal of Nursing Studies, 46*, 557-568.

Fraser, B. (2009). Topic orientation markers. *Journal of Pragmatics 41*, 892-898.

Friginal, E. (2013). Developing research report writing skills using corpora. *English for Specific Purposes, 32*, 208-220.

Furui, S., Nakamura, M., Ichiba, T., & Iwano K. (2005). Analysis and recognition of spontaneous speech using corpus of spontaneous Japanese. *Speech Communication, 47*, 208-219.

Gaballa, H. E. M., & Al-Khayri, M. A. (2014). Testing collocational knowledge of Taif University English seniors. *IOSR Journal of Humanities and Social Science, 19*(11), 63-90.

Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An introduction* (8th ed.). Boston, MA: Pearson Education.

Gardner, D. (2008). Vocabulary recycling in children's authentic reading materials: A corpus-based investigation of narrow reading. *Reading in a Foreign Language, 20*(1). 92-122.

Garfield, E. (2006). The history and meaning of the journal impact factor. *Journal of American Medical Association, 295(*1), 90-93.

Gaskell, G., & Cobb, T. (2004). Can learners use concordance feedback for writing errors? *System, 32*, 301-319.

Gaskin, C. J., & Happell, B. (2014). Power, effects, confidence, and significance: An investigation of statistical practices in nursing research. *International Journal of Nursing Studies, 51*, 795-806.

Gavioli, L. (2005). *Exploring corpora for ESP learning*. Amsterdam: John Benjamins.

Gavioli, L., & Aston, G. (2001). Enriching reality: Language corpora in language pedagogy. *ELT Journal, 55*(3), 238-246.

Gerow, A., & Keane, M.T. (2011). Identifying metaphoric antonyms in a corpus analysis of finance articles. *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*. Boston, MA.

Gilquin, G., Granger, S., & Paqout, M. (2007). Learner corpora: The missing link in EAP pedagogy. *Journal of English for Academic Purposes, 6*, 319-335.

Gledhill, C. (2000). *Collocations in science writing: Volume 22 of language in performance*. Tübingen, Germany: Gunter Narr Verlag.

Godwin-Jones, R. (2010). Emerging technologies, from memory places to spacing algorithms: Approaches to second-language vocabulary learning. *Language Learning & Technology, 14*(2), 4-11.

Granger, S., & Bestgen, Y. (2014). The use of collocations by intermediate vs. advanced non-native writers: A bigram-based study. *IRAL, 52*(3), 229-252.

Granger, S., & Pacquot, M. (2008). Disentangling the phraseological web. In S.Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 27-49). Amsterdam, The Natherlands: Benjamins.

Grant, T. (2010). Text messaging forensics: Txt 4n6: Idiolect free authorship analysis? In J. Gibbons & M. T. Turell (Eds.), *Dimensions of forensic linguistics* (pp. 215-229). Philadelphia, PA: John Benjamins.

Gries, S. Th. (2009). *Quantitative corpus linguistics with R: A practical introduction*. New York, NY: Routledge.

Grundmann, R., & Krishnamurthy, R. (2010). The discourse of climate change: A corpus based approach. *Critical approaches to discourse analysis across disciplines, 4*(2), 125-146.

Gulec, N., & Arif Gulec, B. (2015). Lexical collocations (verb + noun) across written academic genres In English. *Procedia - Social and Behavioral Sciences, 182*, 433-440.

Gyllstad, H. (2013) Looking at L2 vocabulary knowledge dimensions from an assessment perspective – challenges and potential solutions. In C. Bardel, C. Lindqvist, & B. Laufer (Eds.), *Vocabulary acquisition, knowledge and use: New perspectives on assessment and corpus analysis* (pp. 11-28). Eurosla Monographs Series 2. Retrieved from http://eurosla.org/

Hajiveya, K. (2015). A corpus-based lexical analysis of subject-specific university textbooks for English majors. *Ampersand, 2*, 136-144.

Haladyna, T.M., Downing, S.M., & Rodriguez, M.C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education, 15*(3), 309-334.

Halliday, M.A.K. (1961). Categories of the theory of grammar. *Word, 17*(3), 241-292.

Hananfiyeh, M., & Keshi, A. K. (2013). The effect of corpus-based instruction and thesaurus-based teaching on Iranian EFL learners' grammatical knowledge. *Journal of Basic and Applied Scientific Research, 3*(2), 167-179.

Handl, S. (2008). Essential collocations for learners of English: The role of collocational direction and weight. In F. Meunier & S. Granger (Eds.) *Phrasology in foreign language learning and teaching* (pp. 43-66). Amsterdam, The Netherlands: John Benjamins Publishing.

Hanks, P. (2013). *Lexical analysis: Norms and exploitations*. Cambridge, MA: The MIT Press.

Hargreaves, P. (2000). Collocation and testing. In M. Lewis (Ed.) *Teaching collocation: Further developments in the lexical approach* (pp. 205-223). London, UK: Language Teaching Publications.

Hashemi, M., Azizinezhad, M., & Dravish, S. (2011). Collocation a neglected aspect in teaching and learning EFL. *Procedia - Social and Behavioral Sciences, 31*, 522-525.

Hasselgren, A. (2002). Learner corpora and language testing: Smallwords as markers of learner fluency. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.) *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 143-173). Amsterdam, The Netherlands: John Benjamins.

Hegelheimer, V. (2006). Helping ESL writers through a multimodal, corpus-based, online grammar resource. *CALICO Journal, 24*(1), 11-28.

Henriksen, B. (2013). Research on L2 learners' collocational competence and development – A progress report. In C. Bardel, C. Lindqvist, & B. Laufer (Eds.), *Vocabulary acquisition, knowledge and use: New perspectives on*

assessment and corpus analysis (pp. 29-56). Eurosla Monographs Series 2. Retrieved from http://eurosla.org/

Hill, J. (2000). Revising priorities: From grammatical failure to collocational success. In M. Lewis (Ed.), *Teaching collocation: further developments in the lexical approach* (pp. 47-69). London, UK: Language Teaching Publications.

Hoey, M. (1991). *Patterns of lexis in texts*. Oxford, UK: Oxford University Press.

Hoey, M. (2007). Lexical priming and literary creativity. In Hoey, M., Mahlberg, M., Stubbs, M. & Teubert, W. (Eds.) *Text, discourse, and corpora: Theory and analysis*. London, UK: Continuum.

Hou, H. (2014). Teaching specialized vocabulary by integrating a corpus-based approach: Implications for ESP course design at the University Level. *English Language Teaching, 7*(5), 26-37.

Hsu, J.-Y. (2007). Lexical collocations and their relation to the online writing of Taiwanese college English majors and non-English majors. *Electronic Journal of Foreign Language Teaching, 4*(2), 192-209.

Hsu, J.T., & Chiu, C. (2008). Lexical collocations and their relation to speaking proficiency. *The Asian EFL Journal Quarterly, 10*(1), 181-204.

Huang, L. S. (2012). The effectiveness of a corpus-based instruction in deepening EFL learners' knowledge of periphrastic causatives. *TESOL Journal, 6*, 83-108.

Hughes, R. (2010). What a corpus tells us about grammar teaching materials? In A. O'Keeffe and M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 401-412). London, UK: Routledge.

Hulstijn, J.H. (2000). The use of computer technology in experimental studies of second language acquisition: a survey of some techniques and some ongoing studies. *Language Learning and Technology, 3*, 32-43.

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge, UK: Cambridge University Press.

Izumi, E., Uchimotoa, K., & Isahara, H. (2004). SST speech corpus of Japanese learners' English and automatic detection of learners' errors. *ICAME Journal, 28*, 31-48.

Jabbour, G. (2001). *Corpus-based instruction*. (pp.1-20). Opinion paper. New York, NY. (ERIC Document Reproduction Service No. ED452734).

Jablonkai, R. (2009). "In the light of": A corpus-based analysis of lexical bundles in two EU-related registers. *WoPaLP, 3*, 1-27.

Jackson, H., & Zé Amvela, E. (2007). *Words, meaning and vocabulary: An introduction to modern English lexicology*. London, UK: Continuum.

Jaén, M. M. (2007). A corpus-driven design of a test for assessing the ESL collocational competence of university students. *International Journal of English Studies, 7*(2), 127-147.

Jafarpour, A. A., Hashemian, M, & Alipour, S. (2013). A corpus-based approach toward teaching collocation of synonyms. *Theory and Practice in Language Studies, 3*(1), 51-60.

James, C. (1998). *Errors in language learning and use: Exploring error analysis*. Harlow, England: Pearson.

Jaworska. S. (2012). On the F word: A corpus-based analysis of the media representation of feminism in British and German press discourse, 1990-2009. *Discourse Society, 23*(4), 401-431.

Jerz, D., G. (1999). *Academic Journals: What are they?* Retrieved from http://jerz.setonhill.edu/writing/academic1/journals/

Johns, T. (1991). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. *English Language Research Journal, 4*, 27-45.

Jones, S., & Sinclair J. McH. (1974). English lexical collocations. *Cahiers de Lexicologie, 24*, 15-61.

Ju, Y., & Yeon, J. (2015). Linguistic Representation of the Sewol Ferry Disaster: A corpus-based analysis of headlines from two Korean newspapers. *SOAS-AKS Working Paper* (pp. 1-11).

Kay, P., & Fillmore, C. J. (1999). Grammatical constructions and linguistic generalizations: The 'What's X doing Y?' construction. *Language 75*, 1-33.

Kendall, T. (2011). Corpora from a sociolinguistic perspective. *RBLA, 11*(2), 361-389.

Kendall, T. and G. van Herk (2011). Corpus linguistics and sociolinguistic inquiry: Introduction to special issue. *Corpus Linguistics and Linguistic Theory, 7*(1), 1-6.

Kenny, D. (2006). Corpus-based translation studies: A quantitative or qualitative development? *Journal of Translation Studies, 9*(1), 43-58.

Keshavarz, M. H., & Salimi, H. (2007). Collocational competence and cloze test performance: A study of Iranian EFL learners. *International Journal of Applied Linguistics, 17*(1), 81-92.

Kezhen, L. (2015). The use of concordance programs in English lexical teaching in high school. *Higher Education of Social Science, 8*(1), 60-65. DOI: 10.3968/6267.

Khittikote, R. (2011). *The ability to use collocations for business purposes by Thai EFL learners*. Unpublished master's thesis, Language Institute, Thammasat University, Thailand.

Khonamri, F., & Roostaee, S. (2014). The impact of task-based extensive reading on lexical collocation knowledge of intermediate EFL learners. *Procedia - Social and Behavioral Sciences, 136*, 265-270.

Kilgariff, A. (2015). *SketchEngine* [Computer Software]. Available from http://www.sketchengine.co.uk/.

Kim, K., H. (2014). Examining US news media discourses about North Korea: A corpus-based critical discourse analysis. *Discourse & Society, 25*(2), 221-244.

Kırkgöz, Y. (2006). Designing a corpus based English reading course for academic purposes. *The Reading Matrix, 6*(3), 281-298.

Kjellmer, G. (1982). Some problems to the study of collocations in the Brown Corpus. In S. Johansson (Ed.), *Computer corpora in English language research* (pp. 25-33). Bergen, Norway: Norwegian Computing Centre for Humanities.

Koo, K. (2006). *Effects of using corpora and online reference tools on foreign language writing: a study of Korean learners of English as a second language*. PhD (Doctor of Philosophy) thesis, University of Iowa, http://ir.uiowa.edu/etd/65

Kozlowski, A. & Seymour, L. (2003). The importance of collocation in English language teaching. *Contact, 29*(3), 41-49.

Krashen, S. D. (1987). *Principles and practice in second language acquisition*. New York, NY: Prentice-Hall International.

Kuiper, K. (1996). *Smooth talkers: The linguistic performance of auctioneers and sportscasters*. Mahwah, NJ: Lawrence Erlbaum.

Lai, L. S. L., & To, W. M. (2015). Social media content analysis: A grounded approach. *Journal of Electronic Commerce Research, 16*(2), 138-152.

Lakshmi, M. V. (2013). Teaching collocations through comics: An exploratory study. *Humanising Language Teaching*, 4. Retrieved from http://www.hltmag.co.uk/aug13/idea.htm

Lee, D., & Swales, J. (2006). A corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self-compiled corpora. *English for Specific Purposes, 22*(1), 56-75.

Leech, G. (1986). Automatic grammatical analysis and its educational applications. In G. Leech and C. Candlin (Eds.), *Computers in English language teaching and research* (pp. 205-214). London, UK: Longman.

Leśniewska, J. (2006). Collocations and second language use. *Studia Linguistica, 123*, 95-105.

Lété, B. (2003). Building the mental lexicon by exposure to print: A corpus-based analysis of French reading books. In P. Bonin (Ed.), *Mental lexicon. "Some words to talk about words"* (pp. 187-214). Hauppauge, NY: Nova Science Publisher.

Lewis, M. (1993). *The lexical approach: The state of ELT and the way forward.* Hove, UK: Language Teaching Publications.

Lewis, M. (1997). *Implementing the lexical approach*. London, UK: Language Teaching Publication

Lewis, M. (2000). There is nothing as practical as a good theory. In M. Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach* (pp. 10-27). London, UK: Language Teaching Publications.

Lewis, M. (2002a). *The lexical approach: The state of ELT and a way forward*. London, UK: Heinle, Thomson Corporation.

Lewis. M. (2002b). *Implementing the lexical approach: Putting theory into practice*. London, UK: Heinle, Thomson Corporation.

Lin, M. H. (2015), Effects of corpus-aided language learning in the EFL grammar classroom: A case study of students' learning attitudes and teachers' perceptions in Taiwan. *TESOL Quarterly*. doi: 10.1002/tesq.250

Liu, D. (2011). Making grammar instruction more empowering: An exploratory case study of corpus use in the learning/teaching of grammar. *Research in the Teaching of English, 45*(4), 353-377.

Luzón-Marco, M. J. (2011). Exploring atypical verb+noun combinations in learner technical writing. *International Journal of English Studies, 11*(2), 77-95.

Martin, M. (1984). Advanced vocabulary teaching: The problem of synonyms. *Modern Language Journal, 68*(2), 130-137.

Mason, D. J., Kennedy, M. S., Schorr, T., & Flanagin, A. (2006). The power of written word: The influence of nursing journals. In L. C. Andrist, P. K.

Nicholas, and K. A. Wolf (Eds.), *A history of nursing ideas* (pp. 333-356). Boston, MA: Jones and Bartlett Publishers.

Massimo, M. (2011). Spoken corpora and pragmatics. *Revista Brasileira de Linguística Aplicada, 11*(2), 479-519.

Mautner, G. (2007). Mining large corpora for social information: The case of elderly. *Language in Society, 36*, 51-72. doi:10.1017/S0047404507070030.

McCarthy, M. (1990). *Vocabulary*. Oxford, UK: Oxford University Press.

McCarthy, M. (2008). Accessing and interpreting corpus information in the teacher education context. *Language Teaching, 41*, 563-574. doi:10.1017/S0261444808005247

McCarthy. M., & O'Keeffe, A. (2010). Historical perspective: What are corpora and how have they evolved? In A. O'Keeffe and M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 3-13). London, UK: Routledge.

McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. London, UK: Routledge.

McNamara, T. (2000). *Language testing*. Oxford; UK: Oxford University Press.

Mel'čuk, I. A. (1995). Phrasemes in language and phraseology in linguistics. In M. Everaert, E. J. van der Linden, A. Schenk, & R. Schreuder (Eds.), *Idioms: structural and psychological perspectives* (pp. 167-232). Hillsdale, NJ: Lawrence Erlbaum Associates.

Menon, S., & Mukundan, J. (2012). Collocations of high frequency noun keywords in prescribed science textbooks. *International Education Studies, 5*(6), 149-160.

Millar, N. (2011). The processing of malformed formulaic language. *Applied Linguistics, 32*(2), 129-148.

Milton, J. (2013). Measuring the contribution of vocabulary knowledge to proficiency in the four skills. In, C. Bardel, C. Linquist & B. Laufer (Eds.), *L2 vocabulary acquisition, knowledge and use*: *New perspectives on assessment and corpus analysis* (pp.57-78). Eurosla Monographs Series 2.

Milton, J., & Meara, P. (1995). How periods abroad affect vocabulary growth in a foreign language. *Review of Applied Linguistics, 107-108*, 17-34

Miščin, B. E. (2013). Verb collocations in medical English. *US-China Foreign Language, 11*(8), 609-618.

Mohajeri, M. & Ketabi, S. (2013). The knowledge and use of collocations and their relation with English speaking proficiency among upper-intermediate to advanced Iranian EFL learners. *ELT Voices – India, 3*(5), 1-21.

Molavi, A., Koosha, M., & Hosseini, H. (2014). A comparative corpus-based analysis of lexical collocations used in EFL textbooks. *Latin American Journal of Content and Language Integrated Learning, 7*(1), 66-81.

Moon, R. (1987). The analysis of meaning. In J. M. Sinclair (Ed.), *Looking up: An account of the Collins COBUILD Project* (86-103). London, UK: Collins ELT.

Mudraya, O. (2006). Engineering English: A lexical frequency instruction model. *English for Specific Purposes, 25*, 235-256. Retrieved from: http://dx.doi.org/10.1016/j.esp.2005.05.002

Namvar, F., Mohd Nor, N. F., Ibrahim, N., & Mustafa, J. (2012). Analysis of collocations in the Iranian postgraduate students' writings. *3L: Southeast Asian Journal of English Language Studies, 18*(1), 11-22.

Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Boston, MA: Heinle & Heinle Publishers.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge, UK: Cambridge University Press.

Nation, I. S. P. (2011). Research into practice: Vocabulary. *Language Teaching, 44*(4), 529-539. doi:10.1017/S0261444811000267

Nation, P. & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt and M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6-19). Cambridge, UK: Cambridge University Press.

Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford, UK: Oxford University Press.

Nelson, M. (2010). Building a written corpus: What are the basics? In A. O'Keeffe and M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 53-65). London, UK: Routledge.

Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam, The Netherlands: John Benjamins.

Oermann, M. H., & Hays, J. C. (2011). *Writing for publication in nursing*. (2nd ed.). New York, NY: Springer Publishing Company, LLC.

Olohan, M., & Baker, M. (2000). Reporting THAT in translated English: Evidence for subconscious processes of explication? *Across Languages and Cultures, 1*(2), 141-158.

Ördem, E. (2013). Retention and use of lexical collocations (verb+noun and adjective+noun) by applying lexical approach in a reading course. *International Journal of Research Studies in Language Learning, 1*, 1-12.

Oxford University Computing Services on behalf of the BNC Consortium. (2007). *The British National Corpus*, *version 3* (BNC XML Edition). URL: http://www.natcorp.ox.ac.uk/

Palese, A., Coletti, S., & Dante, A. (2013). Publication efficiency among the higher impact factor nursing journals in 2009: A retrospective analysis. *International Journal of Nursing Studies, 50*, 543-551.

Paltridge, B. (2006). *Discourse analysis: an introduction*. London, UK: Continuum

Parkinson, J. (2015). Noun-noun collocations in learner writing. *Journal of English for Academic Purposes 20*, 103-113.

Phoocharoensil, S. (2012). Language corpora for EFL teachers: An exploration of English grammar through concordance lines. *Procedia - Social and Behavioral Sciences, 64*, 507-514.

Pirmoradian, M., & Tabatabaei, O. (2012). The enhancement of lexical collocation learning through concordanceing: A case of Iranian EFL learners. *MJAL, 4*(4), 185-200.

Quesada, J. (2011). Creating your own LSA spaces. In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of latent semantic analysis* (pp. 71-88). New York, NY: Routledge.

Rahimi, M., & Momeni, G. (2012). The effect of teaching collocations on English language proficiency. *Procedia - Social and Behavioral Sciences, 31*, 37-42

Rapti, N. (2013). Data-driven grammar teaching and adolescent EFL learners in Greece. In M. Huber and J. Mukherjee (Eds.), *Studies in variation, contacts and   change in English 13*. Helsinki: VARIENG.

Rashtchi, M., & Afzali, M. (2011). Spoken grammar awareness raising: Does it affect the listening ability of Iranian EFL learners? *Studies in Second Language Learning and Teaching, 1*, 515-531.

Renouf, A. (1997). Teaching corpus linguistics to teachers of English. In A. Wichmann, S. Fligelstone, T. McEnery, and G. Knowles (Eds.), *Teaching and language corpora* (pp. 255-266). London, UK: Longman.

Richards, J. (2006). *Communicative language teaching today*. New York, NY: Cambridge University Press.

Robinson, P. C. (1991). *ESP Today: A practitioner's guide*. London, UK: Prentice Hall

Ruth, B. (2007). How personal is this text? Researching writer and reader presence in student writing using Wordsmith tools. *CORELL: Computer Resources for Language Learning, 1*, 14-21.

Sadighi, S., & Sahragard, R. (2013). The effect of lexical collocational density on Iranian EFL learners' reading comprehension. *The Journal of Teaching Language Skills, 5*(1), 111-136.

Saha, S. (2003). Impact factor: A valid measure of journal quality? *Journal of Medical Library Association, 91*(1), 42-46.

Salza. E. (2014). Using NooJ as a system for (Shallow) ontology population from Italian texts. In S. Koeva, S. Mesfar, and M. Silberztein (Eds.), *Formalising natural languages with NooJ 2013: Selected papers from the NooJ 2013 International Conference* (pp. 191-202). Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

Samburskiy, D. (2014). Corpus-informed pedagogical grammar of English: Pros and cons. *Procedia - Social and Behavioral Sciences, 154*, 263-267. doi:10.1016/j.sbspro.2014.10.148

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge, UK: Cambridge University Press.

Schmitt, N. (2002). Using corpora to teach and assess vocabulary. In M. Tan (Ed.), *Corpus studies in language education* (pp. 31-44). Thailand: IELE Press.

Schmitt, N. (2004). (Ed.). *Formulaic sequences: Acquisition, processing and use*. Amsterdam, The Netherlands: Benjamins.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Basingstoke, UK: Palgrave Macmillan.

Scott, M. (2012). *WordSmith Tools (Version 6)*. [Computer Software]. Liverpool, UK: Lexical Analysis Software.

Shokouhi, H., & Mirsalari, G. (2010). TESL-EJ: Collocational knowledge versus general linguistic knowledge among Iranian EFL learners. *TESL-EJ, 13*(4). Retrieve from: http://www.tesl-ej.org/wordpress/issues/volume13/ej52/ej52a7/

Shooshtari, Z. G., & Karami, N. (2013). Lexical collocation instruction and its impact on Iranian non-academic EFL learners' speaking ability. *Journal of Language Teaching and Research, 4*(4), 767-776.

Shou, C., Wang, H., & Wang, Y. (2014). Critical discourse analysis of the appeals in English women's advertisements. *International Journal of Applied Linguistics & English Literature, 3*(1), 79-85.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford, UK: Oxford University Press.

Sidhom, S., & Lambert, P. (2014). Project management in economic intelligence: NooJ as diagnostic tool for nanometrology cluster. In S. Koeva, S. Mesfar, and M. Silberztein (Eds.), *Formalising natural languages with NooJ 2013: Selected papers from the NooJ 2013 International Conference* (pp. 173-190). Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

Silberztein, M. (2003). *NooJ: A linguistic development environment*. Available from www.nooj4nlp.net.

Sinclair, J. (Ed.) (1987), *Looking up. An account of the COBUILD Project in lexical computing*. London, UK: Collins.

Sinclair, J. (1991). *Corpus, concordance, collocation: Describing English language*. Oxford, UK: Oxford University Press.

Sinclair, J. (1997). *Corpus, concordance, collocation*. Oxford, UK: Oxford University Press.

Sinclair, J. McH. (2004). Introduction. In J. M. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 1-13). Amsterdam, The Netherlands: John Benjamins.

Sinclair, J., Jones, S., & Daley, R. (2004). *English collocation studies: The OSTI report*. London, UK: Continuum.

Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *The Canadian Modern Language Review, 64*(3), 429-458.

Sökmen, A. (1997). Current trends in teaching second language vocabulary. In N. Schmitt and M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy*. (pp. 237-257). Cambridge, UK: Cambridge University Press.

Stvan, L. S. (2013). Stress management: Corpus-based insights into vernacular interpretations of stress. *Communication & Medicine, 10*(1), 81-93.

Steuber, L. C. (2011). Disordered thought, disordered language: A corpus-based description of the speech of individuals undergoing treatment for schizophrenia. *Dissertations and Theses*. Paper 63.

Sun, C-T. (2010). *A corpus-based study of Maritime navigational English materials*. Unpublished Master's thesis, Ming Chuan University, Taiwan.

Sung, J. (2003). *English lexical collocations and their relation to spoken fluency of adult non-native speakers*. Unpublished doctoral dissertation, Indiana University of Pennsylvania, Pennsylvania.

Suwitchanphan, P., & Phoocharoensil, S. (2014). Adjective + noun collocational competence of L1 Thai learners: A comparative study of a regular program and an English program. *Asian Social Science, 10*(17), 210-221.

Svensson, G., & Wood, G. (2007). Ethnocentricity in academic marketing journals: a study of authors, reviewers, editorial boards and editors. *Marketing intelligence & planning, 25*(3), 252-270.

Swales, J. M. (1990). *Genre analysis. English in academic and research settings*. Cambridge, UK: Cambridge University Press.

Szudarski, P. (2012). Effects of meaning- and form-focused instruction on the acquisition of verb-noun collocations in L2 English. *Journal of Second Language Teaching and Research, 1*(2), 3-37.

Takač, V. P. (2008). *Vocabulary learning strategies and foreign language acquisition*. London, UK: Multilingual Matters.

Takač, V. P., & Miščin, E. (2013). Exploring the collocational competence of non-native users of medical English. *JAHR, 4*(7), 235-256.

Tight, D. G. (2010). Perceptual learning style matching and L2 vocabulary acquisition. *Language Learning, 60*(4), 792-833.

Tobina, G., & Lyddy, F. (2014). Media representation of depression in young people: a corpus-based analysis of Irish newspaper coverage. *Irish Journal of Psychological Medicine, 31*(1), 21-30.

Tognini Bonelli, E. (2010). Theoretic overview of the evolution of corpus linguistics. In A. O'Keeffe and M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 14-27). London, UK: Routledge.

Torabian, A. H., Maros, M., & Subakir, M. Y. M. (2014). Lexical collocational knowledge of Iranian undergraduate learners: Implications for receptive and productive performance. *Procedia-Social and Behavioral Sciences, 158*, 343-350.

Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics, 17*, 84-119. http://dx.doi.org/10.1093/applin/17.1.84

University of Toronto's libraries. (2015). *What is a journal article?* Retrieved from http://guides.library.utoronto.ca/c.php?g=250471&p=1670761

Usen, A., & Musigrungsi, S. (2015, 5-6 June). *Effectiveness of teaching collocations to primary school students*. Paper presented at the 7th International Conference on Humanities and Social Sciences, "ASEAN 2015: Challenges and Opportunities". Faculty of Liberal Arts, Prince of Songkla University.

Van der Wouden, T. (1997). *Negative contexts: collocations, polarity and multiple negation*. London, UK: Routledge.

Vannestal, M. E., & Lindquist, H. (2007). Learning English grammar with a corpus: Experimenting with concordancing in a university grammar course. *ReCALL, 19*(3), 329-350.

Vasiljevic, Z. (2014) Teaching collocations in a second language: Why, what and how? *ELTA Journal, 2*(2), 48-73.

Vaughan, E., & Clancy, B. (2013). Small corpora and pragmatics. In J. Romero-Trillo (Ed.) *Yearbook of corpus linguistics and pragmatics 2013 Volume 1* (pp. 53-73). New York, NY: Springer.

Walsh, S. (2010). What features of spoken and written corpora can be exploited in creating language teaching materials and syllabuses? In A. O'Keeffe and M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 333-344). London, UK: Routledge.

Wang, H., Zheng, Y., & Cai, Y. (2015). Application of corpus analysis methods to the teaching of advanced English reading and students' textual analysis skills. In B. Zou, S. Smith, and M. Hoey (Eds.). *Corpus linguistics in Chinese contexts* (pp. 158-177). Hampshire, UK: Palgrave Macmillan.

Wang, J. T., & Good, R. L. (2007). *The repetition of collocations in EFL textbooks: A corpus study*. Paper presented at The Sixteenth International Symposium and Book Fair on English Teaching in the Republic of China, Taipei. Retrieved from http://0-files.eric.ed.gov.opac.msmc.edu /fulltext/ED502758.pdf

Wang, J., Liang, S., & Ge, G. (2008). Establishment of a medical word list. *English for Specific Purposes, 27*, 442-458.

Ward, J. (2009). A basic engineering English word list for less proficient foundation engineering undergraduates. *English for Specific Purposes, 28*, 170-182.

Warren, B. (2005). A model of idiomaticity. *Nordic Journal of English Studies, 4*, 35-54.

Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning. 63*(1), 91-120. doi: 10.1111/j.1467-9922.2012.00729.

Widdowson, H. (1983). *Learning purpose and language use*. Oxford, UK: Oxford University Press.

Wilkinson, M. (2012). The best freeware corpus analysis program for translators? *Translation Journal, 16*(2).

Willis, Y. (1998). *Teaching collocations for productive vocabulary developmen*t. (pp.25-33). Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages. New York, NY. (ERIC Document Reproduction Service No. ED457690).

Willis, D., & Willis, J. (2006). *Doing task-based teaching*. Oxford, UK: Oxford University Press.

Wood, D. (2010). *Perspectives on formulaic language: Acquisition and communication*. London, UK: Continuum.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge, UK: Cambridge University Press.

Yang, C. C. R. (2012). Is gender stereotyping still an issue? An analysis of a Hong Kong primary English textbook series. *Hong Kong Journal of Applied Linguistics, (13)*2, 33-49.

Yang, M.-N. (2015). A nursing academic word list. *English for Specific Purposes, 37*, 27-38.

Yazdandoost, Z., AmalSaleh, E., & afipour, R. (2014). The relationship among collocation knowledge and listening, speaking, reading, and writing proficiency of Iranian EFL learners. *Language, Individual & Society, 8*, 408-419.

Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing, 13*, 257-283.

Yoshida, K. (2013). *Teaching vocabulary to develop communicative competence in foreign language classrooms*. Retrieved from on November, 10, 2015 from http://hdl.handle.net/10935/3465

Yumanee, C., & Phoocharoensil, S. (2013). Analysis of collocational errors of Thai EFL students. *LEARN Journal: Language Education and Acquisition Research Network, 6*(1), 88-98.

APPENDICES
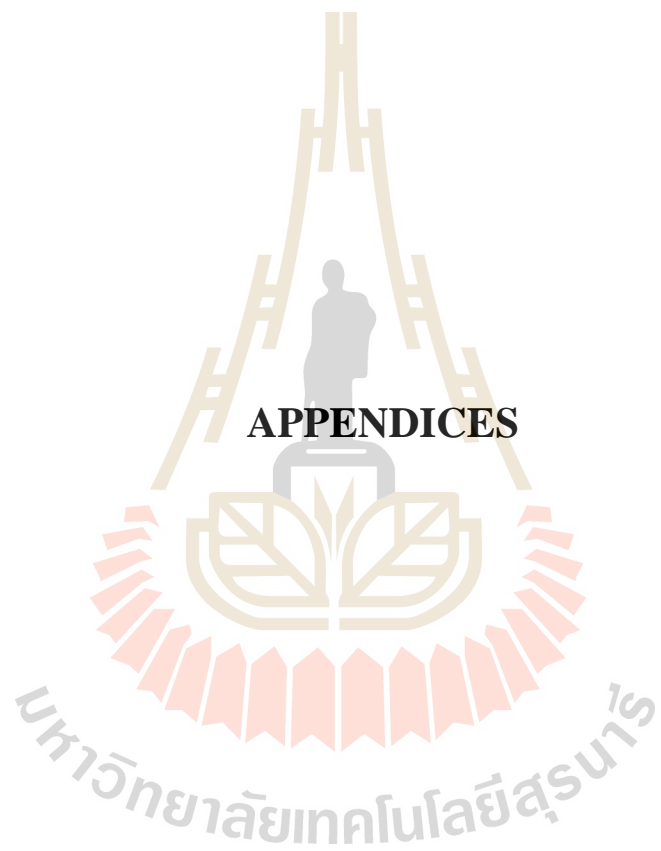
Appendix A

Nursing journals accessible online via SUT's Library Resources

ทุกสาขาวิชา...
ตัวอักษรทั้งหมด... ▼
Search

รูปแบบการให้บริการ
○ Print Only
○ Print+Online
● Online Only

1. British Journal of Community Nursing
2. Clinical nursing reserch
3. Infants and young children
4. International journal of mental heath nursing — Wiley Online Library
5. International journal of nursing practice — Wiley Online Library
6. Journal of advanced nursing — Wiley Online Library
7. Journal of clinical nursing — Wiley Online Library
8. Journal of epidemiology & community health — BMJ
9. Journal of family nursing — SAGE
10. Journal of gerontological nursing
11. Journal of Holistic Nursing
12. Journal of nursing management — Wiley Online Library
13. Journal of nursing research
14. Journal of nursing scholarship — Wiley Online Library
15. Journal of pediatric oncology nursing
16. Journal of psychiatric and mental health nursing — Wiley Online Library
17. Journal of psychosocial nursing and mental health service
18. MCN, American journal of maternal child nursing
19. Medsurg nursing — Academic Search Complete
20. Mother & Care (ภาษาไทย)
21. Nursing and health sciences — Wiley Online Library
22. Nursing BC — ProQuest
23. Nursing economics — ProQuest
24. Nursing education perspectives — ProQuest
25. Nursing ethics — ProQuest
26. Nursing forum — ProQuest
27. Nursing history review — ProQuest
28. Nursing in critical care — Wiley Online Library
29. Nursing inquiry — Wiley Online Library
30. Nursing journal of India — ProQuest

ทุกสาขาวิชา...
ตัวอักษรทั้งหมด... ▼
Search

รูปแบบการให้บริการ
○ Print Only
○ Print+Online
● Online Only

31. Nursing leadership forum — ProQuest
32. Nursing management, Chicago — ProQuest
33. Nursing management, Harrow-on-the-Hill — ProQuest
34. Nursing older people — ProQuest
35. Nursing philosophy — Wiley Online Library
36. Nursing standard — ProQuest
37. Public health nursing — Wiley Online Library
38. Real & Parenting (วารสารภาษาไทย)
39. Scandinavian journal of caring sciences — Wiley Online Library
40. ใกล้หมอ
41. พยาบาลสาร
42. รักลูก
43. รามาธิบดีพยาบาลสาร
44. โรงพยาบาลชุมชน
45. วารสารการพยาบาลและการศึกษา (สถาบันบรมราชชนก)
46. วารสารการพยาบาลและสุขภาพ (มน.)
47. วารสารคณะพยาบาลศาสตร์ มหาวิทยาลัยบูรพา
48. วารสารพยาบาล สมาคมพยาบาลแห่งประเทศไทย
49. วารสารพยาบาลทหารบก
50. วารสารพยาบาลศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
51. วารสารพยาบาลศาสตร์ มหาวิทยาลัยมหิดล
52. วารสารพยาบาลศาสตร์และสุขภาพ มหาวิทยาลัยขอนแก่น(วารสารคณะพยาบาลศาสตร์)
53. วารสารพยาบาลสงขลานครินทร์
54. วารสารระบบบริการปฐมภูมิและเวชศาสตร์ครอบครัว
55. วารสารวิจัยทางการพยาบาล
56. วารสารวิทยาลัยพยาบาลบรมราชชนนี นครราชสีมา
57. วารสารสภาการพยาบาล
58. วารสารสมาคมพยาบาลฯ สาขาภาคตะวันออกเฉียงเหนือ
59. วารสารสมาคมโรคไตแห่งประเทศไทย
60. ศรีนครินทร์เวชสาร — Faculty of Medicine, KKU

|< First   << Back 90   < Prev   (31 - 60 of 61)   Next >   Fwd 90 >>   Last >|

Appendix B

Lexical Collocations in a Sample Corpus of Nursing Research Articles (SCNRA): Pilot Study

I. Objectives

The objectives of this pilot study are:

1. To examine the plausibility of the research project which intends to explore lexical collocations found in a Sample Corpus of Nursing Research Articles (SCNRA).

2. To explore whether any adjustments are necessary in order to conduct the main study successfully.

II. Data Collection Procedures

The pilot study was conducted with a sample corpus of nursing journal articles compiled from 10 journals in the field of nursing intended for the main study. The important procedures in this pilot study involve the selection and compilation of the journal articles.

1. The selection of the journal articles

In the main study, it intends to investigate lexical collocations found in 10 journals in the field of nursing accessible online via Suranaree University of Technology's (SUT) Library Resources in which the 30 latest journal articles of each selected journal will be used. These journals are as follows:

1) Journal of Epidemiology & Community

2) Health International Journal of Mental Health Nursing

3) Journal of Nursing Management

4) Nursing Inquiry

5) Journal of Family Nursing

6) Clinical Nursing Research

7) Journal of Clinical Nursing

8) Journal of Pediatric Oncology

9) Nursing Journal of Psychiatric and Mental Health

10) Nursing International Journal of Nursing Practice

For this pilot study, five research articles from each of the journals listed above were selected to comprise 50 journal articles. With the articles of at least 3,000 word length, the size of the sample corpus should be at least 150,000 words. This corpus size should give sufficient richness of the sample corpus as well as sufficient keywords and their collocates to be examined.

Criteria for selecting journal articles

1) The selected articles were those in the IMRD format since it is commonly used in quantitative and experiment-based research and found used in the journal in the field of nursing.

2) The selected articles were those with the length not less than 3,000 words.

3) As the main study intends to conduct with the 30 latest journal articles from each of the journals, the five articles selected from each journal for this pilot study were those published earlier than the 30 latest articles.

2. Procedures in compiling the articles

1) Once the articles that met the criteria had been identified, the IMRD parts of each article were saved into text files. This means that the abstract, figures, tables, graphs, references and footnotes were left out. The name of each file was given for the purpose of identification and management.

2) Once the selected articles had been compiled, they were uploaded onto the corpus analysis tool, AntConc version 3.4.4, for the analysis. The AntConc is used as to try out its appropriateness to be used as the analysis tool in the main study.

III. Data Analysis Procedures

Data analysis at this stage involves two main steps: the identification of keyword and the identification of the collocations in the sample corpus.

1. Procedures in identifying keywords

Once the selected files were complied, the following procedures were applied to identify the keywords of the corpus being studied.

1) Upload the text files to be studied onto the analysis tool, the AntConc. The latest version 3.4.4 was used. For this pilot study, 50 text files compiled from 10 journals were uploaded.
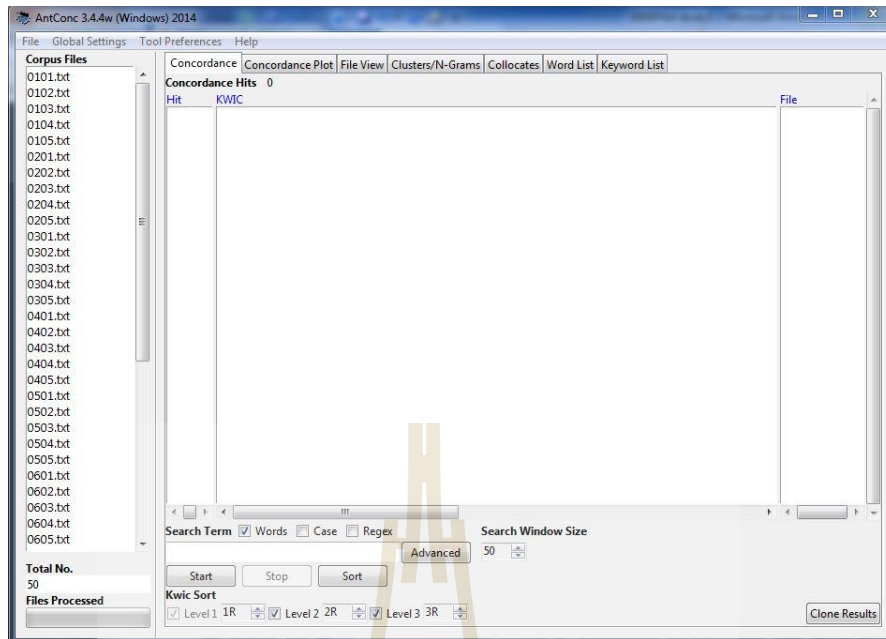
2) Set preferences in 'Global Settings'.
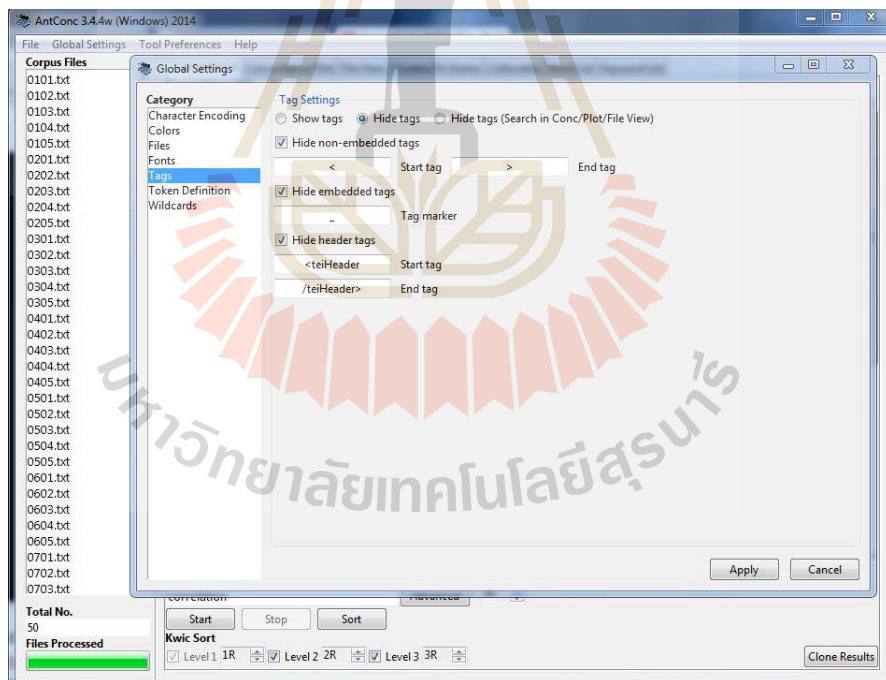
3) Set preferences in 'Tool Preferences'. At this stage, under 'Collocates', MI value is selected as a statistic measure for the association strength of a collocational pair.
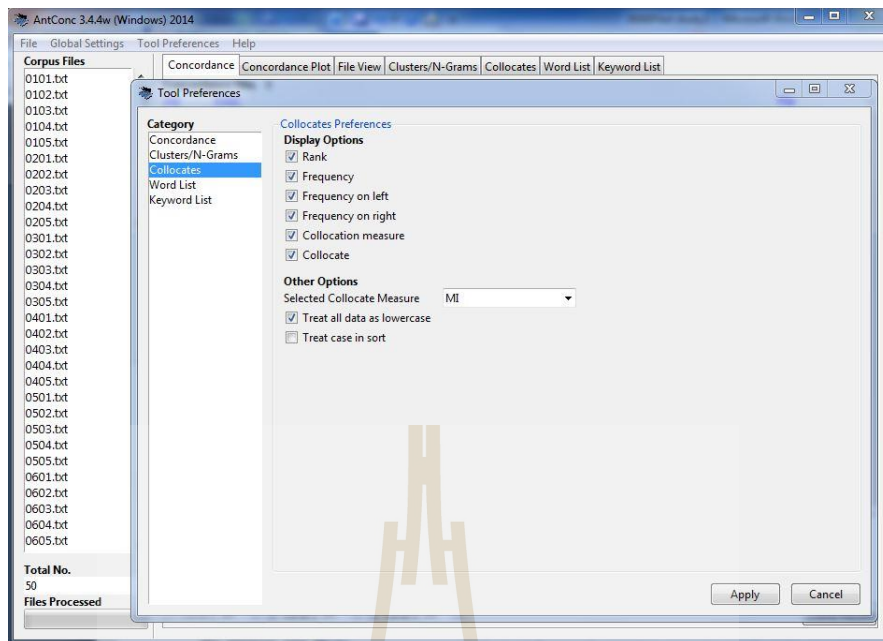
Also, at this stage, under 'Keyword List', Reference Corpus is uploaded. For this study, the British National Corpus is used as the reference corpus.
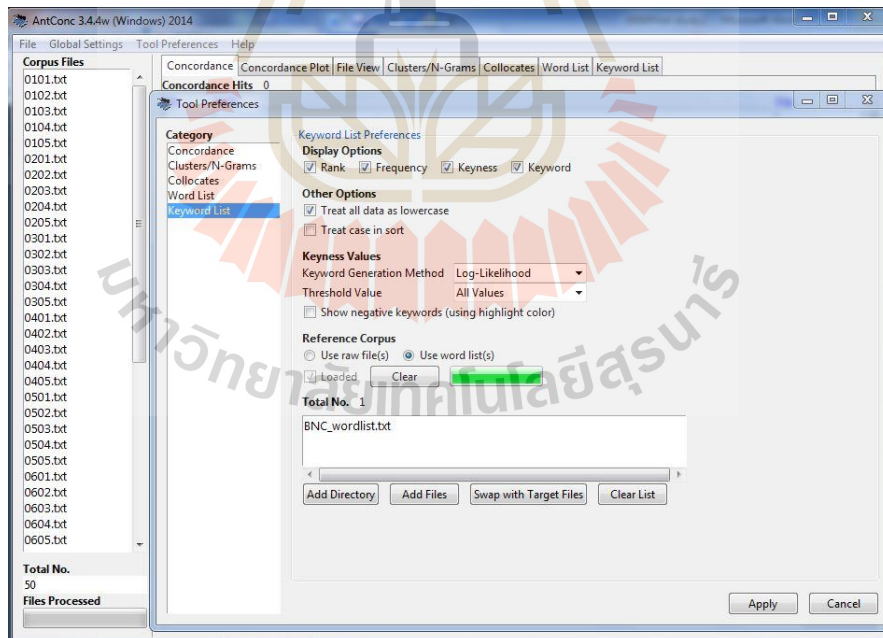
4) Once all the preferences were set, the analysis started. By clicking on the 'Word List' tab followed by clicking on 'Start' button, the results showed the numbers of types and token of the corpus being studied. The words in the corpus were also displayed according to their number of appearance.
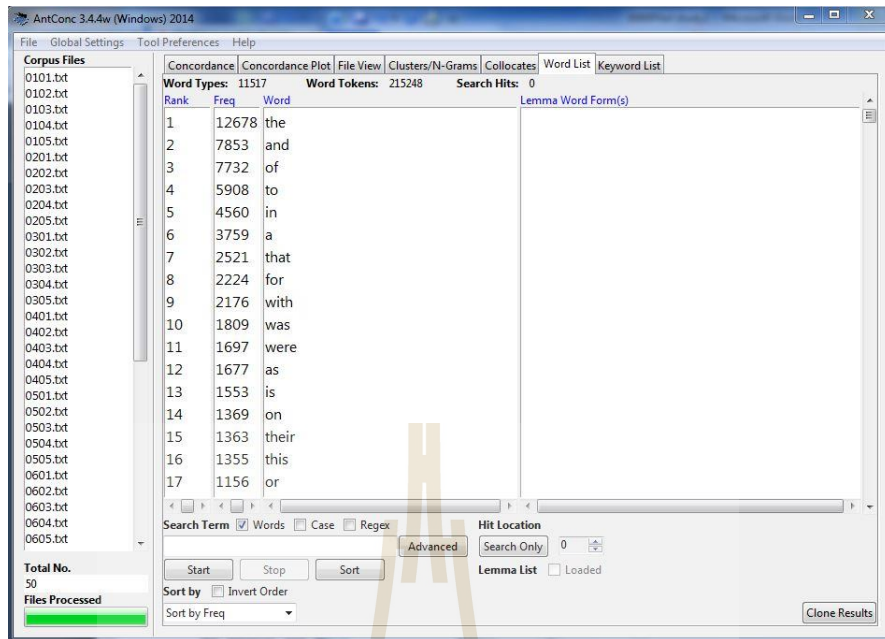
5) By clicking on a 'Keyword List' tab, set 'sort by keyness', and follow by the 'Start' button, the list of keywords of the corpus being studied revealed.
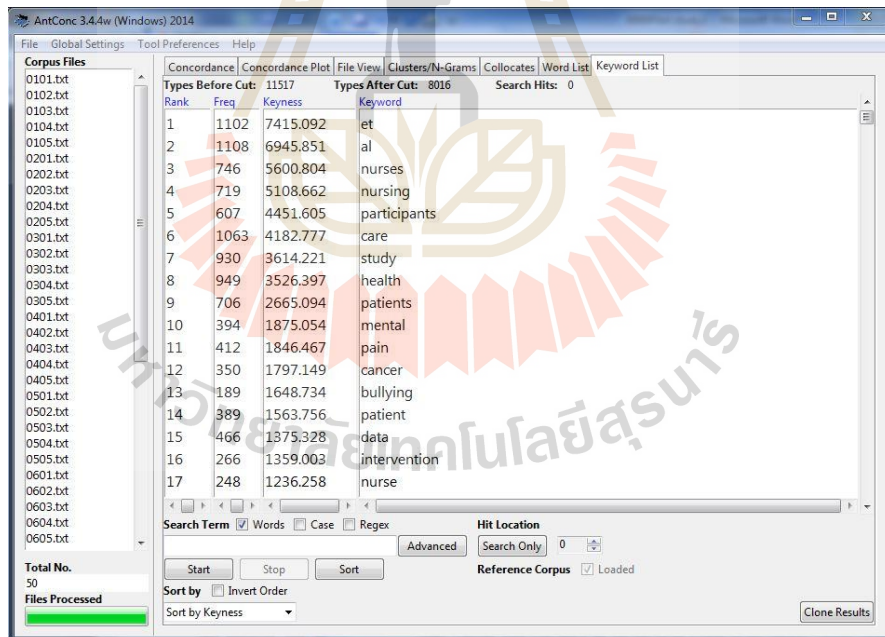
As seen on the picture above, not every word on the list is acceptable as keywords. For example, the first two words on the list being 'et' and 'al'. These two words were not acceptable as keywords. Therefore, the researcher had to look through the list and manually made the list of 500 keywords to be ready of the following stage

of the study, identifying collocations. The list of the first 500 keywords according to the keyness is shown in the Appendix BA.

2. Procedures in identifying collocations

Once the list of keywords were ready, as they were used as the 'nodes', the process of identifying their pairs or 'collocates' started as follows:

1) Under the 'Keyword List' tab, click on a word acceptable to be the node, for example, the word 'Nurses'. The program displayed under the 'Concordance' tab to show the concordance line the word 'Nurses' appeared in the entire corpus with the number of occurrence.

2) To find the collocates, click on the 'Collocates' tab.
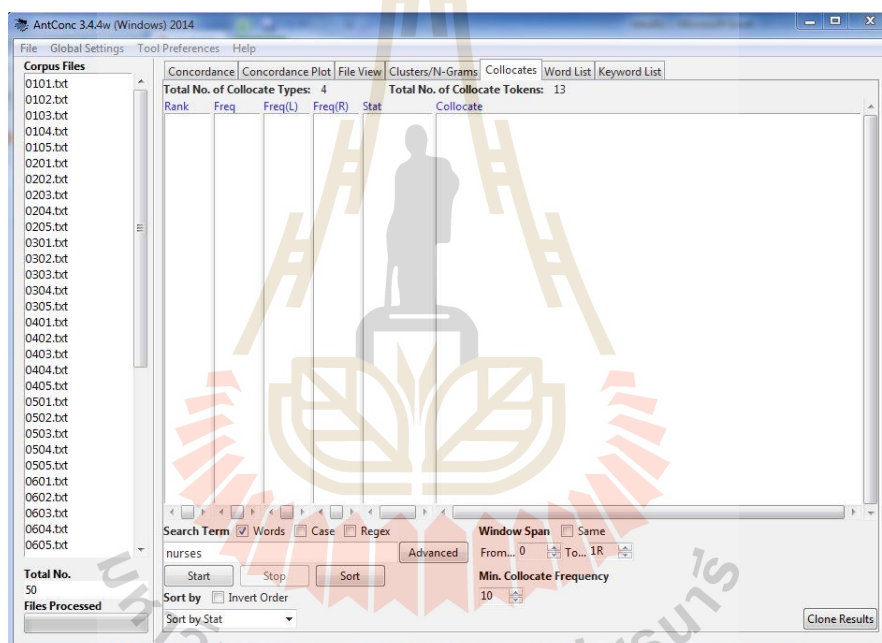
Before clicking on 'Start' button, three preferences were set as follows:

Under 'Window Span', set as 'From 0 to 1R' as the study intends to investigate the two word collocates on the immediate right side of the nodes.

Under 'Sort by', select 'Sort by Stat' as the study give more important to the MI score than the frequency of occurrence.

Under 'Min. Collocate Frequency', set as '10' as the study set the number of the co-occurrence of the pairs at least 10. However, the number of occurrence can be reduced in case of the number of co-occurrence does not meet the criteria. In that case, the intention will be paid on only the MI value of the pair.

Once the preferences were set accordingly, click on 'Start' button.

3) After clicking on 'Start' button, the collocates of the node 'Nurses' revealed. However, as the criteria set for the study that the collocates have to have the MI value at least 3 with the frequency of occurrence at least 10, not all of the words on the list were acceptable collocates. Even those words that meet the criteria are not always eligible to be acceptable pair. In the case of the node 'Nurses', for its collocates shown in the picture below, although the collocate 'working' being the highest ranking in terms of the MI score with the frequency of co-occurrence of 40, they are not considered acceptable pair of collocation. The acceptable collocational pair for the node 'Nurses' in this studied corpus are 'stated' as a Noun+Verb collocation. The collocate 'experiences' was not acceptable collocation as when click on the word to show in the concordance lines it revealed that the pair was in 'nurses' experiences' which with an apostrophe in between them to show possession. Therefore, this pair was not an acceptable pair of collocation in this study.

When reduce the number under 'Min. Collocate Frequency' from 10 to 5, some new collocates appeared on the list. The acceptable collocate on the new list was 'association' as a Noun+Noun combination. However, the collocate 'perceptions' which looked promising was not acceptable as the pair was with an apostrophe in the middle as the case of the collocate 'experiences'.
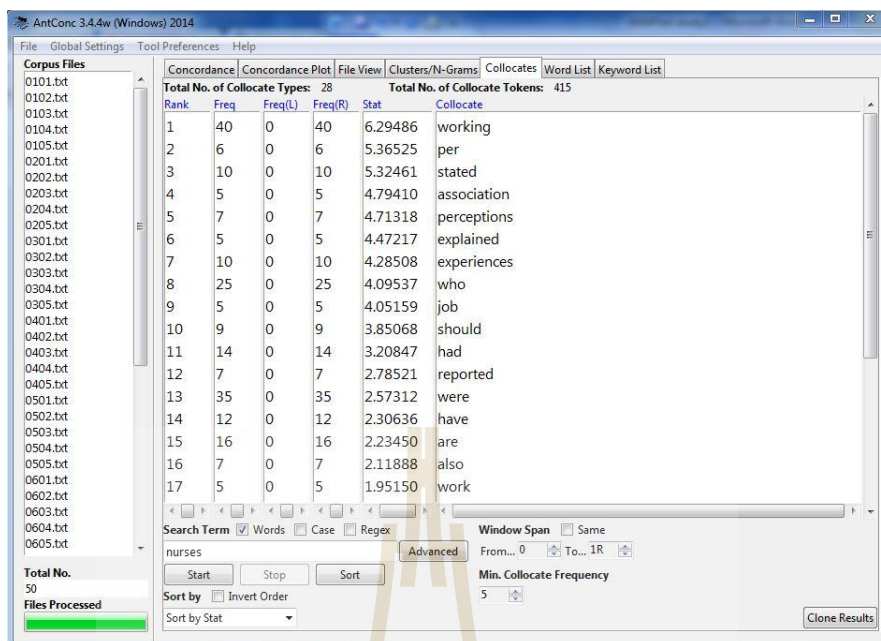
4) The process is repeated for other words to find their collocates.

5) The collocations generated from the 500 keywords or nodes found were grouped together according to the combinations. The combination patterns of the lexical collocations in the study are adapted from that of Benson et al. (2010). That is from the original seven combinations, the present study has adapted to six combinations. This is to create collocations of two immediate word pair on the right side of the nodes which closely related to vocabulary or the extension of the normal single unit vocabulary. The comparison between the combinations given by Benson et al. (2010) and the adapted combinations for the present study is shown in Table 1 below.

Table 1 Lexical Collocation: Benson et al. (2010) and Adapted in Comparison

Types	Combinations by Benson et al. (2010)	Types	Combinations adapted
L1	Verb + Noun		
L2	Verb + Noun	L1	Verb + Noun
L3	Adjective + Noun	L2	Adjective + Noun
L4	Noun + Verb	L3	Noun + Verb
L5	Noun + *of* Noun	L4	Noun + Noun
L6	Adverb + Adjective	L5	Adverb + Adjective
L7	Verb + Adverb	L6	Verb + Adverb

IV. Results

The sample corpus of 50 research articles from the 10 journals in the field of nursing comprises 11,517 types and 225,248 tokens. 500 keywords were selected according to the ranking of their keyness value and frequency of occurrence for further analysis. The list of the keywords is presented in Appendix BA. When classified into categories according to the part of speech, the 500 selected keywords consist of 348 nouns (65.5%), 118 adjectives (22.2%), 58 verbs (10.9%), and 7 adverbs (1.32%) as shown in Table 3.1 below.

Table 3.1 Numbers and percentage of 500 keywords according to parts of speech

No.	Parts of speech	Number	Percent
1	Nouns	348	65.5
2	Adjectives	118	22.2
3	Verbs	56	10.9
4	Adverbs	7	1.32
	Total	531	100

These keywords were then used as 'nodes' for the next step of the study which was to find out the collocates of each keyword. In the present pilot study, as the lexical approach is emphasized, lexical collocations of the nodes will be explored. From the keywords according to their parts of speech presented in Table 3.1 above, the collocations found from the sample corpus can be divided into three groups as follows: 1) lexical collocations according to the set framework; 2) collocations not according to the set framework; and 3) keywords/nodes with no collocates. The numbers and percentage of each group are presented in Table 3.2 below.

Table 3.2 Numbers and percentage of collocations generated from 500 keywords

No.	Parts of speech	Number	Percent
1	Lexical collocations according to set framework	281	52.92
2	Collocations not according to set framework	174	32.77
3	Keywords/nodes with no collocates	76	14.31
	Total	531	100

For the first group, which presents collocations the present study intends to investigate, it reveals the lexical collocations of 281 keywords or nodes that account for 52.92 percent. The numbers and percentage according to the combination types are shown in Table 3.3 below. The complete list of these lexical collocations is presented in Appendix BB.

Table 3.3 Number and percentage of lexical collocations according to types of combination

No.	Types of combination	Number	Percent
1	L1: Verb + Noun	10	3.56
2	L2: Adjective + Noun	115	40.93
3	L3: Noun + Verb	26	9.25
4	L4: Noun + Noun	124	44.13
5	L5: Adverb + Adjective	3	1.07
6	L6: Verb + Adverb	3	1.07
	Total	281	100

The second group, collocations which are not according to the set framework, comprises 174 keywords or nodes with their collocates other than lexical collocations. The majority of these collocations can be categorized as grammatical collocations. Table 3.4 below gives the numbers and percentage of the results in this group. The complete list of collocations in this group is presented in Appendix BC.

Table 3.4 Numbers and percentage of collocations: other types of combination

No.	Types of combination	Number	Percent
1	O1: Noun + Others (e.g. *of+N, prep+N, Vbe+V3*)	127	72.99
2	O2: Verb + Others (e.g. *prep.+N, that+clause*)	41	23.56
3	O3: Adjective + Others (e.g. *prep.+N phrase*)	3	1.72
4	O4: Adverb + Others (e.g. *V, prep.+clause*)	3	1.72
	Total	174	100

The last group, keywords or nodes with no collocates, comprises 76 keywords that account for 14.31 percent. The majority of them are the nouns. Table 3.5 below shows the numbers and percentage of each type of keywords with no collocates. The complete list of this group of keywords is provided in Appendix BD.

Table 3.5 Numbers and percentage of keywords with no collocates

No.	Keywords with no collocates	Number	Percent
1	Nouns	72	93.42

2	Verbs		4	5.26
3	Adverbs		1	1.32
		Total	76	100

V. Discussion

From the findings gained from the pilot study, it is found that from the 500 keywords generated, the majority of them, 65.5 percent, being the nouns. The adjectives come the second place at over 22 percent. The verbs come third at over 10 percent and the adverbs are the least in number, just over 1 percent. Having assigned these keywords as nodes to explore their collocates with one span on the right side to form immediate two word collocations, the results are reported earlier above. It is interesting that these keywords yield the results that can be divided into three groups as shown in Table 3.2 above. Among these three groups, the majority being lexical collocations with the combinations according the set framework of over 52 percent, generated from 281 keywords. The rest of the keywords give the results not according to the set framework that can be divided into two groups: grammatical collocations and no collocations.

As the main focus of the study is on the group of lexical collocations to further apply in the assessment of SUT nursing students' knowledge of lexical collocations, the proportion of the group with over 52 percent and 281 in number is considered large enough to meet the purpose. Although the results reveal that there are combinations of words both within the framework and outside of the framework, all of the combinations are worth investigation as they actually occur in the real use of the language. Thus, despite the focus of the study is on the lexical collocations of the set framework, knowing other possible collocates or no collocates is still useful in effective communication and use of the language as well as the EFL/ESL learning and teaching. With the larger sample size in the main study, it is believed that the results could be slightly different and more statistically significant.

VI. Conclusion

As the objectives of this pilot study are to examine the plausibility of the research project as well as to explore whether any adjustments are necessary in order to successfully conduct the main study, the results of the pilot study assure that this research project is plausible with no adjustment needed.

APPENDIX B.1
List of 500 Keywords generated from the sample corpus

Rank	Freq	Keyness	Keywords	Rank	Freq	Keyness	Keywords
1	746	5600.804	nurses	250	18	146.602	oncology
2	719	5108.662	nursing	251	84	146.536	internal
3	607	4451.605	participants	252	49	146.264	topics
4	1063	4182.777	care	253	23	146.23	logistic
5	930	3614.221	study	254	23	145.631	interviewees
6	949	3526.397	health	255	97	144.715	responsibility
7	706	2665.094	patients	256	64	143.725	concerns
8	394	1875.054	mental	257	75	143.423	aged*
9	412	1846.467	pain				preschool-aged
10	350	1797.149	cancer	258	41	142.862	fathers
11	189	1648.734	bullying	259	85	142.685	units
12	389	1563.756	patient	260	29	142.577	grounded
13	466	1375.328	data	261	52	142.36	guidelines
14	266	1359.003	intervention	262	58	141.392	perspective
15	248	1236.258	nurse	263	51	141.109	explore
16	555	1223.099	family	264	26	140.433	sectional*
17	143	1180.048	siblings				cross-sectional
18	331	1133.696	related*	265	30	140.1	coding
			heart-related,				
			cancer-related	266	30	138.269	questionnaires
19	134	1076.232	adolescents	267	109	136.959	specific
20	220	1041.702	findings	268	23	136.763	weekday
21	155	1030.86	workplace	269	135	136.669	personal
22	331	1013.287	risk	270	72	136.339	stated
23	210	1001.197	clinical	271	76	135.989	educational
24	420	885.978	research	272	22	135.836	transcribed
25	148	876.57	restraint	273	38	135.125	disclosure
26	117	852.836	interventions	274	86	135.064	overall
27	95	848.83	inpatient	275	141	134.319	individual
28	137	843.817	outcomes	276	302	134.068	used
29	293	842.399	reported	277	42	133.263	environments
30	259	841.417	factors	278	38	133.213	lifestyle
31	174	792.594	satisfaction	279	60	133.118	consent
32	328	747.007	self*	280	108	132.581	previous
			self-sacrificing,				
			self-sacrifice,	281	14	131.846	behavioral
			self-efficacy,				
			self-funding,				
			self-care	282	85	131.837	values
33	241	711.982	physical	283	30	131.713	adolescent
34	88	707.663	clinicians	284	91	131.042	context
35	146	707.407	caring	285	17	129.194	analgesics

Rank	Freq	Keyness	Keywords	Rank	Freq	Keyness	Keywords
36	102	692.346	medication	286	40	127.796	dimensions
37	137	674.634	scores	287	31	127.194	balancing
38	146	671.578	professionals	288	46	126.893	institutional
39	72	650.267	conjugal	289	37	126.219	shifts
40	259	648.011	analysis	290	63	126.188	funding
41	57	625.088	caregivers	291	24	126.057	stigma
42	445	620.536	children	292	38	124.809	challenges
43	51	606.409	borns	293	116	124.23	current
44	75	606.02	sickle	294	28	123.152	statistically
45	189	605.432	items	295	46	123.002	errors
46	96	604.47	participant	296	37	122.747	engage
47	137	603.299	perceived	297	81	122.473	decisions
48	148	576.452	experiences	298	212	122.348	important
49	246	575.964	studies	299	33	122.074	treatments
50	140	572.446	conducted	300	22	121.16	psychiatry
51	172	547.963	focus	301	69	121.128	processes
52	101	546.731	organisational	302	12	120.034	subcategories
53	97	535.822	functioning	303	72	118.417	communication
54	123	527.017	interviews	304	373	118.284	work
55	44	522.113	antipsychotic	305	107	118.178	increased
56	139	513.869	tasks	306	13	117.277	healthful
57	102	505.858	collaboration	307	107	116.709	population
58	81	502.397	behaviours	308	130	116.411	value
59	95	498.13	psychiatric	309	27	116.007	respondent
60	193	497.448	medical	310	60	115.286	selected
61	215	493.859	treatment	311	38	114.926	clinic
62	206	491.094	professional	312	12	114.74	sociodemographic
63	235	490.436	results	313	51	114.542	falls
64	134	489.516	emotional	314	48	113.769	reporting
65	167	485.195	users	315	45	113.351	depression
66	227	483.947	hospital	316	146	113.132	process
67	90	473.561	settings	317	133	113.117	groups
68	68	471.728	informants	318	11	113.074	dichotomised
69	176	469.289	families	319	75	112.888	measures
70	99	460.883	discharge	320	21	112.743	obesity
71	316	454.483	support	321	30	112.591	evaluated
72	47	449.661	hospitalization	322	20	112.505	governance
73	241	441.355	role	323	14	112.083	demographics
74	108	440.577	consumers	324	119	111.951	quality
75	182	438.777	scale	325	105	111.944	environment
76	123	436.896	healthy	326	77	111.924	daily
77	40	436.668	subscales	327	63	111.799	identify
78	214	425.837	knowledge	328	59	111.725	doctors
79	196	424.156	significant	329	115	111.645	greater
80	133	423.708	sample	330	32	111.143	quantitative

Rank	Freq	Keyness	Keywords	Rank	Freq	Keyness	Keywords
81	126	423.17	score	331	93	110.787	difference
82	117	421.002	illness	332	34	110.469	highlighted
83	84	419.704	respondents	333	79	110.143	collection
84	91	419.121	distress	334	13	109.234	colors
85	113	409.843	profession	335	102	108.799	lower
86	68	404.938	efficacy	336	52	107.59	ward
87	213	401.456	parents	337	37	107.204	disability
88	86	399.141	validity	338	23	107.134	proxy
89	222	399.123	practice	339	120	106.864	provided
90	81	391.715	questionnaire	340	27	106.72	dynamics
91	37	387.233	pediatric	341	165	106.609	level
92	67	380.019	qualitative	342	46	105.841	contribute
93	188	379.281	relationship	343	25	103.465	avoidance
94	290	378.893	service	344	79	103.34	condition
95	32	377.733	caregiving	345	43	102.8	acute
96	119	374.771	significantly	346	97	102.677	issues
97	106	365.871	symptoms	347	215	102.369	social
98	77	364.84	perceptions	348	14	102.174	individualized
99	264	355.694	based*	349	21	102.077	discursive
			web-based	350	65	101.737	cell
100	54	350.611	fasting	351	48	101.006	consistent
101	233	343.754	age	352	32	100.931	evaluate
102	315	343.408	information	353	42	100.753	assessed
103	51	342.843	healthcare	354	29	100.592	sampling
104	83	342.21	participate	355	17	100.034	disclosing
105	102	340.836	adults	356	24	99.049	illnesses
106	73	338.992	ethical	357	24	98.499	correlated
107	93	336.29	researchers	358	27	97.392	tribal
108	172	331.064	included	359	34	97.377	correlation
109	197	330.933	higher	360	60	97.114	tests
110	64	326.399	reliability	361	38	96.951	foster
111	37	322.361	medications	362	67	96.686	institutions
112	88	317.059	oral	363	8	96.673	hematological
113	64	316.369	deprivation	364	8	96.673	website
114	148	315.096	associated	365	54	95.773	describe
115	217	311.032	experience	366	35	95.684	barrier
116	77	307.357	chronic	367	24	95.339	transportation
117	114	306.887	procedures	368	46	95.311	evaluation
118	77	303.644	breast	369	40	95.079	uncertainty
119	78	299.28	focused	370	36	95.045	relevance
120	74	298.241	themes	371	48	94.63	categories
121	44	290.295	undone	372	86	94.544	influence
122	81	288.908	admission	373	49	94.345	testing
123	56	287.093	disabilities	374	23	93.55	collaborative
124	74	286.056	diagnosis	375	20	93.235	exploratory

Rank	Freq	Keyness	Keywords	Rank	Freq	Keyness	Keywords
125	40	285.75	sibling	376	13	93.187	somatic
126	70	280.918	couples	377	15	92.954	internet
127	55	279.809	therapeutic	378	107	92.86	conditions
128	132	277.197	older	379	55	92.576	practices
129	169	272.773	described	380	20	92.501	cervical
130	51	272.73	directives	381	17	92.491	informant
131	132	272.358	compared	382	43	92.001	demonstrated
132	150	271.604	activities	383	114	91.94	total
133	127	270.289	individuals	384	51	91.844	partners
134	80	268.967	multiple	385	66	91.508	initial
135	67	265.7	prevention	386	12	91.196	confirmatory
136	55	262.175	severity	387	23	90.126	baseline
137	127	261.712	positive	388	12	90.082	centered*
138	110	261.564	relationships				family-centered
139	63	260.825	analyses	389	61	89.991	managers
140	137	259.851	lack	390	14	89.808	biomedical
141	22	257.627	subscale	391	12	89.544	multidimensional
142	80	256.627	participation	392	26	89.1	supportive
143	61	256.616	carers	393	31	89.073	codes
144	81	254.174	hospitals	394	49	88.786	manage
145	78	253.452	responses	395	15	88.586	workplaces
146	73	252.667	variables	396	22	88.445	domains
147	47	252.089	fatigue	397	47	87.445	registered
148	120	251.405	differences	398	22	87.331	participated
149	50	249.522	inequalities	399	21	87.284	contextual
150	88	247.167	characteristics	400	78	87.263	primary
151	141	246.6	effects	401	32	87.087	assessing
152	88	246.226	recovery	402	65	87.064	explained
153	130	244.736	status	403	8	86.674	bedrails
154	77	241.153	strategies	404	181	86.664	help
155	93	240.977	interview	405	24	86.64	ongoing
156	32	240.948	color	406	53	86.377	strongly
157	77	239.442	roles	407	76	86.296	method
158	81	236.839	expectations	408	15	85.642	analyzed
159	30	235.921	behavior	409	26	85.01	focusing
160	50	231.617	providers	410	56	85.003	feelings
161	69	229.653	perception	411	96	84.548	questions
162	53	229.622	prevalence	412	115	84.405	food
163	81	228.536	residents	413	54	84.325	prior
164	108	228.271	assessment	414	18	84.134	validated
165	29	225.774	psychosocial	415	8	83.796	semistructured
166	203	224.976	services	416	21	83.369	workload
167	103	224.951	identified	417	68	83.161	setting
168	109	224.769	aspects	418	49	82.606	informed
169	22	223.892	retest*	419	164	82.383	young

Rank	Freq	Keyness	Keywords	Rank	Freq	Keyness	Keywords
			test-retest	420	48	82.337	performed
170	47	221.062	perspectives	421	94	82.036	considered
171	56	220.281	cognitive	422	11	81.949	circadian
172	68	219.991	mortality	423	16	81.914	depressive
173	48	215.943	surgical	424	30	81.318	construct
174	27	210.594	clinician	425	43	81.207	tools
175	35	209.541	chemotherapy	426	32	80.965	sensitivity
176	84	208.79	indicated	427	214	80.629	found
177	71	208.64	psychological	428	64	80.622	reduce
178	19	208.363	spousal	429	15	79.82	narratives
179	268	207.93	group	430	16	79.782	transcripts
180	28	206.849	preconceived	431	10	79.761	professionalization
181	194	206.609	child	432	12	79.705	readmission
182	81	205.196	shift	433	31	79.693	moderate
183	70	204.981	responsibilities	434	39	79.688	influenced
184	18	204.519	hospitalizations	435	76	79.383	review
185	92	204.254	experienced	436	13	79.158	impairments
186	44	204.183	regression	437	26	78.985	researcher
187	45	199.729	decreased	438	25	78.677	detection
188	52	198.739	screening	439	7	78.565	preadmission
189	87	198.667	observed	440	28	78.277	contexts
190	54	198.077	barriers	441	40	77.893	samples
191	30	197.542	socioeconomic	442	46	77.702	examine
192	34	196.177	outpatient	443	37	77.646	informal
193	86	195.195	partner	444	13	77.631	verbatim
194	72	194.864	mothers	445	26	77.482	facilitate
195	84	194.174	negative	446	11	77.225	facilitators
196	35	190.974	restraints	447	68	76.945	holding
197	39	190.024	expressive	448	33	76.772	duration
198	23	189.515	hospitalized	449	10	76.576	coworkers
199	113	189.199	understanding	450	45	76.356	specifically
200	72	188.975	collected	451	20	76.331	textbooks
201	181	186.721	community	452	30	76.011	descriptions
202	82	186.006	attitudes	453	32	75.965	incidence
203	175	185.599	management	454	59	75.515	formal
204	60	184.102	mutual	455	28	74.874	dishes
205	82	182.239	advance	456	8	74.53	healthiness
206	25	181.622	descriptors	457	69	74.47	weight
207	15	181.261	antibullying	458	29	74.189	enhance
208	58	181.044	sessions	459	24	73.161	recruited
209	132	179.957	model	460	35	73.079	tool
210	108	179.729	disease	461	42	72.955	category
211	29	179.269	orthopaedic	462	71	72.783	sector
212	101	178.891	survey	463	88	72.6	rights
213	71	176.891	item	464	51	72.526	advanced

Rank	Freq	Keyness	Keywords	Rank	Freq	Keyness	Keywords
214	62	173.562	sharing	465	6	72.504	healthfulness
215	81	170.706	literature	466	32	72.359	emotions
216	31	170.593	interpersonal	467	51	72.218	shared
217	40	170.439	demographic	468	37	71.934	surgery
218	186	170.377	education	469	84	71.828	feeling
219	56	169.28	gender	470	81	71.235	activity
220	73	168.418	parent	471	44	71.044	reflect
221	114	167.92	fall	472	74	70.84	traditional
222	41	167.308	ethics	473	25	70.523	rated
223	42	166.568	consistency	474	38	70.254	observation
224	41	166.148	coping	475	60	70.206	cultural
225	42	163.836	participating	476	7	70.187	subcategory
226	118	162.513	response	477	7	70.187	subthemes
227	36	161.985	web*	478	12	70.118	specialties
			web-based	479	55	69.977	completed
228	112	160.897	unit	480	14	69.902	malnutrition
229	92	159.513	impact	481	43	69.72	outcome
230	59	159.249	instrument	482	49	69.124	reduction
231	31	159.213	correlations	483	46	69.113	session
232	40	158.224	citizenship	484	40	69.06	measured
233	13	157.093	caregiver	485	15	68.541	cardiovascular
234	21	156.835	partnered	486	36	68.474	commonly
235	152	156.32	needs	487	7	68.312	unshared
236	37	155.779	diagnosed	488	55	67.488	content
237	36	154.104	descriptive	489	249	67.448	use
238	48	153.921	intensity	490	24	67.308	administered
239	27	153.556	negatively	491	53	67.06	opportunities
240	51	151.43	functional	492	63	66.985	fast
241	36	150.778	survivors	493	147	66.941	working
242	57	150.536	assess	494	21	66.927	dietary
243	160	149.423	staff	495	19	66.9	influencing
244	98	148.65	ability	496	6	66.768	generalizability
245	82	148.515	factor	497	36	66.332	medicine
246	19	147.986	psychometric	498	35	66.24	authors
247	21	146.931	aging	499	41	66.024	adequate
248	45	146.819	indicating	500	21	65.916	hygiene
249	34	146.651	ranged				

APPENDIX B.2
List of Lexical Collocations gained from the pilot study

L1: Verb + Noun

	Nodes	Collocate	MI score	Freq
1	conducted	using	5.60908	6
2	included	statements	8.15187	10
3	hospitalized	patients	6.05048	5
4	assess	fatigue	8.65009	5
5	identify	colors	9.03792	2
6	evaluate	pain	5.02914	2
7	describe	topics	7.34604	2
8	analyzed	using	8.24651	4
9	enhance	support	5.55388	2
10	influencing	health	5.16241	3

L2: Adjective + Noun

	Nodes	Collocate	MI score	Freq.
1	nursing-sensitive	outcomes	9.46561	9
2	mental	distress	7.75572	36
		illness	7.43268	37
		health	7.35814	285
3	heart-related	conditions	6.30393	13
4	cancer-related	lymphedema	7.4636	19
5	clinical	practice	7.25137	33
		outcomes	6.90336	16
6	physical	restraints	8.67347	16
		restraint	8.50019	60
		activity	7.85522	21
		disabilities	7.45483	11
		performance	7.10231	10
7	conjugal	relationship	9.31305	40
		relationships	8.57171	14
8	perceived	family	5.58873	17
9	organisational	processes	8.27083	10
		factors	7.50005	22
10	antipsychotic	medication	10.71307	35
11	psychiatric	hospitalization	9.05063	11
		units	8.32136	12
		unit	7.79786	11
12	medical	errors	9.40698	28
		doctors	8.14743	15
13	professional	caregivers	8.28371	17
		nurses	6.21403	53
14	emotional	symptoms	7.24356	10

		support	6.34577	16
15	healthy	diet	9.32567	11
		lifestyle	8.84713	10
		family	6.74424	34
16	significant	difference	8.60617	33
		differences	7.7176	23
17	pediatric	oncology	11.33626	8
		patients	5.62763	6
18	qualitative	findings	6.45315	6
		analysis	6.21771	6
		data	5.78536	8
		research	5.25723	5
		study	4.78846	8
19	web-based	patient	4.52705	11
20	ethical	considerations	11.04039	10
21	higher	score	6.81675	13
		scores	6.80291	13
		level	6.42771	13
22	oral	mucositis	10.43992	46
		hygiene	9.67125	7
23	chronic	illnesses	11.25621	21
		diseases	10.50544	13
		illness	8.03792	11
24	focused	ethnography	10.84528	6
25	undone	items	6.69397	4
26	therapeutic	alliance	11.82737	13
		conversation	11.15667	14
27	older	adults	8.45925	22
		people	7.29901	31
28	multiple	informants	8.76568	11
29	positive	attitude	8.72696	7
		affect	7.58945	5
		attitudes	6.95437	6
30	psychosocial	health	4.9674	4
31	cognitive	impairment	10.00139	4
32	surgical	wards	11.13068	10
33	clinician	respondents	8.56844	4
34	psychological	distress	8.86545	14
35	spousal	care	5.99875	6
36	preconceived	expectations	11.21229	25
37	socioeconomic	status	8.78638	8
38	negative	assumptions	10.41643	8
39	expressive	family	8.22078	30
40	collected	samples	8.80875	6
		data	5.00346	5
41	mutual	caring	8.52582	15

		care	6.50971	27
42	advance	directive	11.35809	18
		directives	11.30037	49
		care	4.76359	11
43	orthopaedic	nurses	7.48455	18
44	interpersonal	relations	11.061	8
45	demographic	characteristics	9.10421	9
46	descriptive	statistics	11.40821	10
47	functional	limitations	10.47493	29
48	psychometric	properties	12.92022	13
49	internal	consistency	10.8379	30
		medicine	10.32332	18
50	logistic	regression	11.90257	18
51	preschool-aged	children	5.01109	5
52	grounded	theory	10.83974	20
53	cross-sectional	survey	9.16434	7
		study	5.96147	7
54	specific	questions	6.36249	4
55	age-specific	performance	6.92509	4
56	personal	barriers	7.20587	5
		support	4.92001	6
		time	4.4804	7
57	educational	intervention	8.27041	29
58	overall	scale	6.10351	5
59	individual	needs	6.13552	7
60	previous	studies	7.87622	29
		research	6.70594	22
61	behavioral	domains	10.44885	2
		interventions	8.03792	2
62	adolescent	years	5.77533	2
63	institutional	review	9.26608	10
64	current	study	5.80393	28
65	important	implications	6.98772	5
		factor	5.9521	5
		part	5.27622	5
		role	4.65979	6
66	increased	risk	5.77341	9
67	harmful	characteristics	9.8777	5
		workplace	9.32404	6
68	selected	areas	8.37579	5
69	soicodemographic	characteristics	9.67125	4
70	daily	life	7.75636	14
		practice	6.97637	10
71	greater	understanding	6.63493	6
		risk	5.08443	6
72	quantitative	data	6.65881	7

73	lower	level	6.99882	10
74	acute	care	5.93589	13
75	social	recognition	8.55241	12
		support	5.36411	13
76	individualized	feedback	10.58636	2
77	discursive	construction	12.00139	4
78	consistent	pattern	9.04321	2
79	tribal	members	9.16374	10
80	hematological	findings	8.51924	3
81	collaborative	behaviours	9.85223	8
82	exploratory	factor	9.84352	7
		study	5.53263	4
83	somatic	units	10.77573	9
		nurses	6.05713	3
84	cervical	cancer	7.26443	5
85	total	score	6.90547	8
		scale	6.37496	8
		number	6.25548	7
86	initial	number	7.23662	8
87	confirmatory	factor	11.35809	12
88	family-centered	interventions	8.84528	3
		care	5.66171	3
89	biomedical	model	9.18582	5
90	multidimensional	model	8.67125	3
91	supportive	care	3.96127	2
92	registered	nurses	7.7473	35
93	contextual	barriers	11.02787	11
94	primary	caregivers	8.59735	8
		care	4.54624	9
95	ongoing	decisions	8.37579	3
		process	7.94085	4
96	prior	research	5.56844	5
97	semistructured	interview	10.17648	4
98	informed	consent	10.19404	16
99	young	adults	9.20922	46
		people	7.86455	57
		person	7.47273	18
100	depressive	symptoms	10.79507	14
101	moderate	quality	7.45159	3
102	informal	care	5.77419	10
103	formal	help	7.91811	12
		care	5.58643	14
104	weight-related	behaviours	5.59007	6
105	cancer-related	stressors	9.34495	6
106	advanced	practice	7.2488	8
		nursing	6.46026	15

107	shared	decision	8.58378	9
108	traditional	dishes	10.02076	10
		food	8.36114	13
109	cultural	differences	6.48682	3
		support	5.50497	4
110	cardiovascular	diseases	10.74986	3
111	unshared	topics	11.6155	5
112	administered	surveys	10.13068	2
113	working	time	6.75964	37
114	dietary	habits	12.67125	7
115	adequate	patient	6.33942	6

L3: Noun + Verb

	Nodes	Collocate	MI score	Freq.
1	nurses	stated	5.32461	10
2	participants	expressed	6.03713	10
		described	4.87656	14
		reported	4.59726	20
3	study	showed	5.26959	11
		reported	3.11923	11
		used	3.07378	11
4	patients	admitted	5.37765	6
5	intervention	compared	7.57016	31
6	siblings	described	6.96224	14
		reported	6.2679	15
7	findings	suggest	7.76436	12
8	factors	influencing	8.62083	9
9	scores	indicated	7.03265	7
10	professionals	need	5.19669	5
11	tasks	left	9.68416	34
12	results	indicated	7.25316	14
		showed	7.03177	10
13	informants	spoke	9.82082	4
		indicated	7.23586	4
		identified	6.94168	4
14	scale	developed	6.18992	5
15	respondents	expressed	8.1534	6
		described	6.7298	7
16	themes	emerged	9.7513	8
17	individuals	avoided	9.06399	6
18	challenges	associated	6.84322	3
19	errors	made	7.49164	4
20	population	included	5.13287	3
21	respondent	detailed	10.75912	5
22	process	involved	5.33599	2

23	informant	stated	9.45825	4
24	setting	located	8.16875	2
25	coworkers	think	9.80875	3
26	authors	concluded	9.779	2
		reported	5.3916	2

<div align="center">L4: Noun + Noun</div>

	Nodes	**Collocate**	**MI score**	**Freq.**
1	nurses	association	4.7941	5
2	nursing	tasks	7.5987	90
		homes	7.35132	24
		profession	6.26359	29
		home	6.23613	35
		staff	5.90386	32
		activities	5.45641	22
		practice	5.18626	27
		interventions	4.81486	11
		research	4.36953	29
		work	3.93069	19
		care	3.6313	44
3	care	providers	7.01786	32
		plans	6.71418	14
		behaviors	5.79922	11
		planning	5.42905	10
		settings	5.33979	18
		funding	5.00636	10
		professionals	4.93132	22
		education	4.44448	20
4	study	design	5.6008	13
5	health	promotion	7.01802	12
		professionals	6.67995	66
		condition	6.47579	31
		services	6.38886	75
		behaviours	6.34351	29
		organization	6.20089	12
		problems	5.90977	22
		consumers	5.65545	24
		issues	5.22546	16
		care	5.05685	156
		information	4.95243	43
		outcomes	4.72734	16
		service	4.50345	29
6	pain	intensity	7.6141	18
		management	6.86333	39
		scores	6.63155	26
7	cancer	survivors	9.00139	30

			diagnosis	6.51441	11
8	bullying		acts	9.51597	9
9	patient		safety	6.897	14
			education	6.66032	34
			satisfaction	6.19263	23
			outcomes	5.92087	15
			care	3.14555	17
10	data		collection	8.6562	69
			analysis	6.4785	50
11	nurse		satisfaction	4.95409	6
12	family		functioning	7.78074	55
			member	7.5993	10
			members	7.33834	58
			life	5.68441	24
			support	5.46544	36
13	workplace		environments	8.74764	13
			bullying	8.65863	55
			environment	7.63216	15
14	risk		reduction	8.43068	26
			factors	6.65007	40
			assessment	6.39742	14
15	research		question	7.48355	22
			ethics	6.96577	10
			team	6.67947	12
16	restraint		prevalence	8.10019	10
			use	7.30107	27
17	inpatient		unit	8.86199	23
			units	8.19583	11
			experience	7.08467	13
18	self-report		measures	7.37134	6
19	self-care		behaviors	5.79922	11
20	self-sacrificing		behaviors	11.07178	6
21	self-funding		residents	8.20587	7
22	caring		efficacy	9.43835	32
23	medication		side-effects	9.51715	25
24	sickle		cell	11.46445	64
25	participant		observation	8.88275	8
26	focus		group	8.28938	67
			groups	6.69352	11
27	functioning		scale	6.41529	7
28	treatment		course	7.82994	5
29	hospital		stays	9.88909	6
			discharge	5.58166	5
			day	4.97421	6
30	discharge		referrals	9.58378	6
			planning	8.53169	8

31	support	services	5.06845	10
32	role	barriers	7.63283	12
33	knowledge	base	8.48875	5
34	sample	size	9.66036	20
35	efficacy	scale	8.70535	24
36	practice	roles	6.2394	6
37	relationship	status	7.38661	19
38	service	users	9.27097	139
		user	8.97013	25
		providers	7.47684	12
39	fasting	times	9.22379	12
		time	6.3169	10
40	age	groups	7.60353	28
		group	6.87283	34
41	healthcare	agencies	9.85057	7
		institutions	8.78448	7
42	breast	cancer	9.26443	77
43	sibling	adjustment	10.69327	8
44	prevention	reinforcement	11.23451	6
		interventions	7.10111	5
		care	3.91755	5
45	fatigue	scale	7.23822	6
46	recovery	process	6.87374	7
47	interview	questions	7.39887	7
48	test-retest	reliability	10.71564	11
49	mortality	risk	6.25749	8
50	chemotherapy	treatments	9.86389	5
51	group	sessions	8.251	22
52	shift	workers	9.28833	12
53	regression	analysis	8.32686	17
54	screening	tests	8.91566	7
55	outpatient	clinic	10.83968	11
56	restraints	use	6.94828	5
57	community	residents	7.57638	13
58	management	approaches	6.62057	2
59	antibullying	policies	11.93428	12
60	model	fit	8.28658	9
61	disease	control	6.90764	5
62	survey	completion	9.572	5
		items	5.81711	5
63	item	scale	6.86545	7
64	literature	review	8.58729	11
65	education	program	8.59152	11
66	gender	differences	7.00139	4
67	parent	study	4.47208	7
68	fall	prevention	9.06459	19

		group	7.67464	29
		risk	7.09703	24
		patients	4.87866	11
69	ethics	committee	11.87266	20
70	coping	strategies	9.67626	12
71	response	rate	9.11053	20
72	unit	level	5.86389	5
73	staff	turnover	7.02448	3
		members	5.27477	4
74	factor	structure	9.73006	11
		analysis	7.51121	18
75	aging	population	8.58186	4
76	oncology	patients	6.66715	6
77	coding	scheme	12.32332	5
78	weekday	day	9.49959	14
79	disclosure	decision-making	8.16028	5
80	lifestyle	changes	7.31121	3
81	consent	form	9.4102	11
82	balancing	needs	9.97295	22
83	funding	systems	8.79976	6
		issues	8.13845	8
84	psychiatry	ward	8.55577	2
85	communication	efficacy	8.26561	7
86	work	schedule	8.17261	7
		environments	6.58765	7
		environment	5.26572	7
		attitudes	5.13698	5
		life	4.84269	9
87	value-adding	work	7.14286	12
		care	6.50644	22
88	clinic	visit	10.25826	8
89	reporting	pain	5.76611	5
90	depression	subscales	8.48682	3
		scale	6.30096	3
91	obesity	prevention	8.8422	3
92	governance	processes	9.87015	6
93	quality	care	4.08882	10
94	difference	score	8.10609	15
95	collection	method	7.48586	5
96	disability	organizations	9.69883	2
97	proxy	report	8.81357	5
98	cell	pain	8.62148	49
		disease	8.52335	12
99	sampling	strategy	9.53573	2
100	disclosing	health	4.73791	2

101	correlation	coefficients	11.62818	6
102	foster	carers	10.62444	17
103	internet	access	9.02739	2
104	assessing	risk	6.66688	5
105	food	trays	10.87015	4
		restaurant	10.87015	4
		outlet	10.87015	4
		preferences	8.34659	4
106	construct	validity	10.28938	15
107	professionalization	process	9.78885	6
108	readmission	rates	9.3759	2
109	review	board	10.46771	10
110	preadmission	information	8.19404	3
111	holding	children	6.91798	17
112	weight	status	7.75467	9
		outcomes	7.50908	8
113	sector	healthcare	9.06339	9
		institutions	8.30716	7
114	surgery	council	9.50619	3
115	feeling	safe	9.66036	6
116	activity	data	4.09657	3
117	observation	tool	8.33843	2
118	outcome	variables	8.68451	6
119	reduction	initiatives	9.93101	4
120	session	intervention	6.1368	4
121	content	validity	9.67794	18
122	fast-track	programme	11.1534	10
123	fast	food	8.98033	17
124	medicine	wheel	12.54572	2
		units	7.13632	2

L5: Adverb + Adjective

	Nodes	Collocate	MI score	Freq.
1	significantly	higher	8.48417	39
2	statistically	significant	9.93743	25
3	commonly	used	6.89227	6

L6: Verb + Adverb

	Nodes	Collocate	MI score	Freq.
1	reported	significantly	6.62607	16
2	transcribed	verbatim	12.8777	10
3	correlated	negatively	11.69772	10

APPENDIX B.3
List of Collocations with other combinations

Nouns

	Nodes	Collocate	MI score	Freq.
1	adolescents	with N	4.51628	31
2	satisfaction	with N	3.7088	23
3	children	with N	3.85288	65
4	experiences	of N	3.5203	61
5	studies	have V3	5.32191	32
6	focus	on N	5.29673	17
7	collaboration	between N	6.41022	16
		with N	3.27768	10
8	behaviours	about N	6.20301	14
9	knowledge	about N	6.12332	35
10	parents	had V3	4.0168	7
11	validity	of N	3.23073	29
12	questionnaire	was V3	4.01425	11
13	relationship	between N	6.84998	40
14	symptoms	after N	5.29284	5
15	perceptions	of N	3.95849	43
16	information	about N	5.72169	39
		on N	3.08499	17
17	researchers	have V3	4.53263	7
18	reliability	was Adj/V3	4.5951	13
20	medications	on N	5.25721	9
21	deprivation	were Adj/V3	4.4463	11
22	procedures	are V3	3.5296	6
23	couples	who clause	5.18719	5
24	activities	were Adj/V3	3.45849	13
25	individuals	who clause	5.32778	10
26	severity	of N	3.01765	16
27	relationships	between N	5.88625	12
		with N	4.01674	18
28	analyses	were V3	4.33152	10
29	lack	of N	4.67826	126
30	subscale	of N	3.92455	12
31	participation	in N	4.36818	35
32	hospitals	have N	4.24651	5
		in N	3.3909	18
33	responses	were V3	4.0234	10
34	variables	were V3	3.11897	5
35	differences	between N	6.63519	22
		in N	3.97586	40
		were V3	3.53941	11
36	inequalities	in N	5.16489	38

37	characteristics	of N	3.54904	37
38	effects	on N	3.74212	12
		of N	3.3599	52
39	status	as N/Adj	3.56657	12
40	strategies	were V3/Adj/Ving	3.30505	6
41	color	is	4.43673	5
42	roles	as N/Adj	3.05911	5
43	perception	on N	4.64764	11
		of N	4.21405	46
44	prevalence	of N	3.71495	25
45	assessment	is N/Adj	3.16727	7
46	aspects	of N	4.27876	76
47	perspectives	of N	3.15132	15
48	child	with N	4.42058	42
49	barriers	against N	9.90186	12
50	mothers	were V3	3.13887	5
51	understanding	of N	3.78619	56
52	attitudes	towards N	9.5418	23
53	descriptors	for N	4.27477	5
54	sessions	were V3/Adj/N	4.58832	11
55	sharing	everything	10.36289	11
56	consistency	of N	3.49416	17
57	coping	with N	4.73006	11
58	participating	in N	4.81236	25
59	impact	on N	5.90245	35
		of N	3.36292	34
60	instrument	was V3	3.59698	6
		for N	3.52141	7
61	correlations	between N	7.45038	10
62	citizenship	is	4.1148	5
63	intensity	of N	3.30151	17
64	ability	to V1	4.7623	73
65	interviewees	were V3/Adj/N	5.27066	7
66	responsibility	for N phrase	4.75167	27
67	concerns	about N phrase	6.7355	16
68	fathers	of N	3.02642	12
69	guidelines	for N phrase	4.5967	13
70	questionnaires	were V3	5.2499	9
71	consent	was V3	4.57274	12
72	values	were V3	3.89941	10
73	context	of N phrase	3.64877	41
74	dimensions	of N	3.47709	16
75	shifts	were V3/Adj	3.36238	3
76	stigma	towards N	7.79083	2
77	decisions	about N phrase	6.39565	16
78	treatments	were V3/Adj	3.52744	3

79	processes	are V3/N phrase	3.25397	3
80	quality	of N	4.07412	72
81	difference	between N phrase	6.24393	13
		was V3	4.16286	14
		in N	3.60663	24
82	dynamics	within N	6.73675	4
83	level	of N	3.64215	74
84	issues	related to N	5.33	6
85	correlation	between N phrase	7.58014	12
86	tests	were V3/N	3.88733	7
87	institutions	have V3	4.19834	4
88	barrier	for N	4.27477	7
		to N	3.86526	14
89	transportation	to N	3.18719	6
90	evaluation	of N phrase	3.36292	17
91	uncertainty	about N	5.41357	4
92	relevance	of N phrase	3.53598	15
		to N	3.33919	10
93	categories	were V3	3.40191	4
94	influence	on N phrase	4.57091	13
95	testing	is V3/N/Adj	4.08505	6
96	conditions	are V3/Adj	3.35799	5
		for N	3.02516	9
97	practices	are Adj	4.3181	5
98	partners	are N/Adj	4.10511	4
99	managers	were V3	3.05613	4
100	codes	were V3	4.84003	7
101	domains	of N	3.33958	8
102	bedrails	are N/Adj	5.77753	2
103	help	from N	3.63837	10
104	focusing	on N phrase	7.11986	23
105	feelings	of N	3.79901	28
106	questions	were V3	3.98687	12
107	tools	are V3/N/Adj	5.15862	7
108	narratives	of N phrase	3.47709	6
109	transcripts	were V3	4.57183	3
110	researcher	was Adj	3.19422	2
111	detection	of N	4.40309	19
112	samples	were V3	3.2499	3
113	facilitators	and N	4.31718	8
114	duration	of N	4.56198	28
115	textbooks	from N phrase	5.07932	3
116	descriptions	of N	3.89212	16
117	incidence	of N	3.88648	17
118	dishes	from N phrase	4.59389	3
119	healthiness	of N phrase	4.60637	7

120	tool	was V3/N	4.08731	5
121	healthfulness	of N	4.53598	5
122	emotions	were V3/Adj	3.57183	3
123	feeling	of N	3.16158	27
124	reduction	in N	3.64655	13
125	use	of N	3.58048	107
126	opportunities	for N	4.95624	17
127	generalizability	of N	4.53598	5

Verbs

	Nodes	**Collocate**	**MI score**	**Freq.**
1	focus	on N	5.29673	26
2	results	from N	3.64021	13
3	participate	in N	4.70934	46
4	associated	with N	6.46312	132
5	focused	on N	6.22078	37
6	described	as N/Adj	4.50998	30
		by N	4.22588	15
7	compared	with N/clause	6.07564	90
8	identified	by N	4.2033	9
		as N	3.90244	12
9	indicated	that clause	5.57813	47
10	experienced	by N	4.89675	13
		as N/Adj	4.06537	12
11	decreased	after N	6.52891	5
12	observed	that clause	3.14284	9
13	diagnosed	with N	6.32561	30
14	indicating	that clause	4.92401	16
15	ranged	from	7.5868	29
16	stated	that clause	5.53134	39
17	used	to V1	3.60699	101
18	engage	in N phrase	5.15872	28
19	dichotomised	at	7.8109	7
20	highlighted	how clause	6.07871	3
		that clause	3.3284	4
21	provided	by N	5.13491	20
22	contribute	to N phrase	5.02118	41
23	assessed	by	5.49748	9
24	correlated	with N phrase	4.85057	7
25	describe	their N	3.87011	5
26	demonstrated	that clause	4.89649	15
27	manage	their N	5.49572	14
28	participated	in N	5.34932	19
29	explained	that clause	4.91705	23
30	validated	for N	4.01174	3
31	performed	by N	5.59434	11

	Nodes	Collocate	MI score	Freq.
32	considered	as N phrase	4.14982	13
33	found	that clause	4.65167	63
34	reduce	their N	3.88803	6
35	influenced	by N	7.01943	24
36	examine	the N phrase	3.147	24
37	recruited	through N	7.06459	3
		from N	5.81628	6
38	reflect	on clause	4.8373	8
39	rated	on N phrase	4.65288	4
40	completed	by N	5.10844	9
41	measured	by N	5.71987	10

Adverbs

	Nodes	Collocate	MI score	Freq.
1	negatively	with clause	5.19522	10
2	strongly	disagree	11.68816	13
		agree	11.13972	15
3	specifically	examined	8.0945	2

Adjectives

	Nodes	Collocate	MI score	Freq.
1	greater	than	6.13596	13
2	consistent	with N phrase	6.13068	34
3	prior	to N/N phrase	4.85856	43

APPENDIX B.4
List of Keywords with no collocates

Verb with no collocates

1	partnered	3	evaluated
2	explore	4	facilitate

Adverb with no collocates

1 verbatim

Noun with no collocates

1	interventions	25	disabilities	49	demographics
2	outcomes	26	diagnosis	50	environment
3	self-sacrifice	27	directives	51	doctors
4	self-efficacy	28	carers	52	colors
5	clinicians	29	expectations	53	ward
6	analysis	30	behavior	54	avoidance
7	caregivers	31	providers	55	condition
8	borns	32	residents	56	illnesses
9	items	33	services	57	website
10	interviews	34	responsibilities	58	baseline
11	users	35	hospitalizations	59	workplaces
12	settings	36	partner	60	method
13	families	37	caregiver	61	workload
14	hospitalization	38	needs	62	sensitivity
15	consumers	39	survivors	63	impairments
16	subscales	40	topics	64	contexts
17	score	41	units	65	category
18	illness	42	perspective	66	rights
19	distress	43	environments	67	subcategory
20	profession	44	analgesics	68	subthemes
21	caregiving	45	subcategories	69	specialties
22	adults	46	falls	70	malnutrition
23	experience	47	groups	71	hygiene
24	admission	48	measures		

Appendix C

Nursing Collocations' Lesson Plan

Tentative schedule

No.	Contents	Hours
1	- Introduction to the project and the workshop program	1
	- Pre-test on Collocations in Nursing Research Articles	2
2	Noun + Noun collocations; Adjective + Noun collocations	2
3	Noun + Verb and Verb + Noun collocations	1
4	200 most frequent collocations	1.5
5	200 most frequent collocations according to individual nodes	1.5
6	Summarizing the whole lesson	1
7	Post-test	2
	Total	12

Due to a tight schedule of the nursing students, the collocation tests and the workshop were arranged to fit in a two-day schedule given upon the availability of the students.

From the tentative schedule above the lesson plan can be arranged as follows:

No.	Topics	Duration (hour)
1	Noun + Noun collocations	1
2	Adjective + Noun collocations	1
3	Noun + Verb collocations	0.5
4	Verb + Noun collocations	0.5
5	200 most frequent collocations	1.5
6	200 most frequent collocations according to each node	1.5

The objectives:

There are two main objectives for the lesson. One is to introduce students to lexical collocations in the field of nursing. The other is to administer the pre-test and the post-test with the target students. This is to investigate students' knowledge of collocations from the results of the pre-test and to examine the effect of corpus-based instruction of collocations on the knowledge of the students from the results of the post-test.

Teaching methods:

Since the present study is a corpus-based study with the emphasis on lexical approach particularly focusing on collocations found from the studied sample corpus, the teaching method to be applied in the workshop is a corpus-based instruction applying "Three Is: Illustration, Interaction, and Induction as suggested by McEnery et al. (2006). By showing students examples of how words are used and what words they collocate with in the real contexts as shown in concordance lines is beneficial to the EFL/ESL learners (Kozlowski & Seymour, 2003).

Corpora are useful tools for engaging learners in the interpretive process to create models of their own (Leech, 1986). Corpus-based teaching and learning can capture reality and are able to provide valid models for learners as they represent authentic language (Gavioli & Aston, 2001). Corpus tools show students the frequency of particular features of the language (Coxhead, 2010). A corpus reveals register variation of a language and a complex relationship between lexicon and grammar. It also allows learners to investigate the frequency of formulaic lexical bundles in any register (Samburskiy, 2014). A number of studies showed the advantages of demonstrating words in concordance lines over only traditional gap-filling or matching exercises (Cobb, 1997, 1999; Horst, Cobb, & Nicolae, 2005; Pickard, 1994; Stevens, 1991; Thurstun & Candlin, 1998). Concordance lines also enable learners to see variation of linguistic structures, promoting a process of synthesis and analysis of information on their part, which is the key to the acquisition process (Aston, 1995). Additionally, engaging students in corpus-based activities promotes noticing or consciousness-raising (Conrad, 2005; Thurstun & Candlin, 1998).

Materials and instruments:

The main materials and instruments for the workshop on lexical collocations extracted from the SCNRA are a concordance program and online exercises as well as handouts of each collocation pair.

The concordance program used is the AntConc version 3.4.4 loaded with the Sample Corpus of Nursing Research Articles (SCNRA). This enables the researcher to show the concordance lines of particular collocation pairs to the students.

The online exercises used are generated through facilities provided by the website **Quizlet.com**. With this website, the researcher can generate matching exercises of collocation pairs and use them as a warmer task to draw attention of the students for each lesson. Handouts of collocations of each topic are also provided.

The lesson plan:

After the introduction and the pre-test, the lesson begins. As the lesson is on collocations of different types of combination, the lesson is carried out under the similar pattern and procedures as follows:

Introduction: (Time: 5 – 10 minutes)

For each topic, the lesson starts with a matching exercise generated from the online resource website, **Quizlet.com**. This exercise should draw the attention of the students to the lesson as well as give them some ideas of what collocations are. The interfaces of the exercise are shown below.

This matching exercise is a Noun + Noun collocations exercise. When a word is dragged onto its pair, the pair will disappear. This exercise should take less than ten minutes.

The lesson: (Time: depends on the time allowed for each topic)

After the matching exercise, the lesson begins. Handouts of the collocation lists according to the topics will be distributed for each lesson to show the collocations pairs of each type of combination. Then, the concordance lines for each collocation pair are shown on the corpus tool, the AntConc. By showing the concordance lines, the students get the opportunity to see how the collocation pairs appear in the context of the real language use. The explanations and discussion on the meaning and the context of the collocation pairs are carried out in order for the students to understand the context where the collocation pairs are used. The interface of the concordance lines of the pair '*health + care*' is shown below as the example.

The figure above shows how the collocation pair appears in the concordance lines.

Then the full sentences of the collocation pairs are shown in order to discuss their contexts and meaning.

Participants were not unique in their lack of primary health care. In general, presentation to primary health-care services by young people is low because they do not understand them and associate them with discomfort and minimal confidentiality (Jones & Bradley 2007; NSW Centre for the Advancement of Adolescent Health 2008). It is common for young adults with mental illness to identify they have no GP (Curtis et al. 2012). Additionally, some general practice-based services do not

Tackling inequalities in physical health and life expectancy demands improved access and quality of physical health care for people with mental illness. There is often inconsistent and low access to quality physical health services such as screening, effective referral processes, continuity of care and participation in health-behaviour programmes (Connolly et al. 2015; Hyland et al. 2003; Nasrallah et al. 2006). In recent years, the debate on how to address physical health service gaps has become more prominent (De Hert et al. 2011; Ehrlich et al. 2015; Lawrence & Kisely 2010; Richardson et al. 2005). In a policy environment emphasising collaborative, recovery-focused and cross-sectoral care (British Medical Association 2014; National Mental Health Commission 2012), it is important that all stakeholders views are sought and considered to determine how to best reform health care to improve access to quality physical health care for people with mental illness.

The process is repeated with the collocation pairs according to their frequency order. The number of the pairs to be shown for each combination type depends on the time allowed for each session.

After lesson exercise: (Time: 10 minutes)

Each lesson ends with the exercise to allow the students to evaluate themselves. The exercise is a multiple-choice generated by the tool available from **Quizlet.com**. The interface of the exercise is shown below.
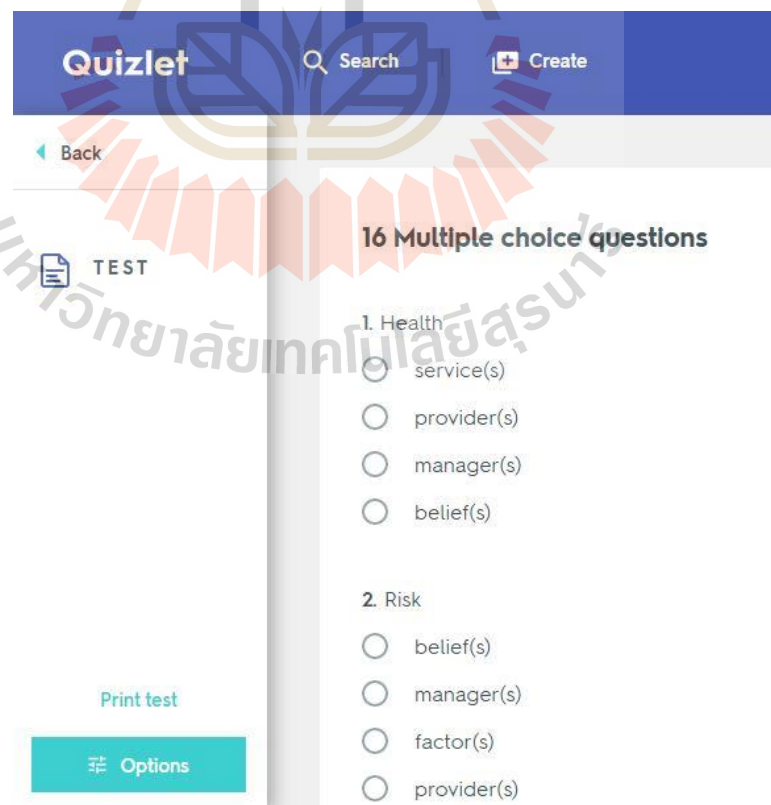
Lesson 1: Noun + Noun collocations

The objectives	to introduce students to lexical collocations with **Noun + Noun** combinations
Target collocations	health (x,xx) care　　　　　family member/s health service/s　　　　　nursing home/s care provider/s　　　　　　service user/s family caregiver/s　　　　health (x) provider/s data collection　　　　　　quality (x,xx) care nursing practice　　　　　job satisfaction focus group/s　　　　　　quality (x) life *full list of Noun + Noun collocations is in Handout #2
Step 1: Introduction	Matching exercise generated through facilities provided by the website **Quizlet.com**
Step 2: The lesson	Show students the concordance lines of each pair of collocations, point out how each pair co-occurs, point out their possible positions, discuss their meaning in contexts
Step 3: After lesson exercise	Multiple-choice exercise generated through facilities provided by the website **Quizlet.com**

Lesson 2: Adjective + Noun collocations

The objectives	to introduce students to lexical collocations with **Adjective + Noun** combinations
Target collocations	mental (x,xx) health　　　physical (x,xx) health mental illness/es　　　　　present study palliative (xx) care　　　　social support physical activity　　　　　mental (x) service/s previous studies　　　　　significant difference/s chronic(x,xx) illness/es　　higher (x,xx) score/s registered nurse/s　　　　older adult/s *full list of Noun + Noun collocations is in Handout #3
Step 1: Introduction	Matching exercise generated through facilities provided by the website **Quizlet.com**
Step 2: The lesson	Show students the concordance lines of each pair of collocations, point out how each pair co-occurs, point out their possible positions, discuss their meaning in contexts
Step 3: After lesson exercise	Multiple-choice exercise generated through facilities provided by the website **Quizlet.com**

Lesson 3: Noun + Verb collocations

The objectives	to introduce students to lexical collocations with **Noun + Verb** combinations
Target collocations	data (x) collected study (x) conducted score/s indicate/ed/ing nurses working participants (x) asked score/s range/s/d study aims/ed results show/ed studies (x,xx) conducted data (x) analysed/zed interviews (x) conducted research (x) needed study (xx) explore findings suggest *full list of Noun + Noun collocations is in Handout #4
Step 1: Introduction	Matching exercise generated through facilities provided by the website **Quizlet.com**
Step 2: The lesson	Show students the concordance lines of each pair of collocations, point out how each pair co-occurs, point out their possible positions, discuss their meaning in contexts
Step 3: After lesson exercise	Multiple-choice exercise generated through facilities provided by the website **Quizlet.com**

Lesson 4: Verb + Noun collocations

The objectives	to introduce students to lexical collocations with **Verb + Noun** combinations
Target collocations	participate (xx) study analys/zed using diagnosed (x,xx) cancer manage (x,xx) health measured using manage (x,xx) care bereaved (x) suicide diagnosed (xx) patients provide (x,xx) support diagnosed (x) schizophrenia improve (x,xx) quality access (x,xx) services provided (x,xx) information assessed using *full list of Noun + Noun collocations is in Handout #4
Step 1: Introduction	Matching exercise generated through facilities provided by the website **Quizlet.com**
Step 2: The lesson	Show students the concordance lines of each pair of collocations, point out how each pair co-occurs, point out their possible positions, discuss their meaning in contexts
Step 3: After lesson exercise	Multiple-choice exercise generated through facilities provided by the website **Quizlet.com**

Lesson 5: 200 most frequent collocations in the SCNRA

The objectives	to introduce students to 200 most frequent lexical collocations found in the SCNRA
Target collocations	mental (x,xx) health health care family member/s health service/s physical (x,xx) health nursing home/s mental illness/es care provider/s service user/s present study palliative (xx) care family caregiver/s social support physical activity *full list of Noun + Noun collocations is in Handout #5
Step 1: Introduction	Matching exercise generated through facilities provided by the website **Quizlet.com**
Step 2: The lesson	Show students the concordance lines of each pair of collocations, point out how each pair co-occurs, point out their possible positions, discuss their meaning in contexts
Step 3: After lesson exercise	Multiple-choice exercise generated through facilities provided by the website **Quizlet.com**

Lesson 6: 200 most frequent collocations in the SCNRA according to each node

The objectives	to introduce students to 200 most frequent lexical collocations found in the SCNRA **according to each node**
Target collocations	mental (x,xx) health health care family member/s physical (x,xx) health nursing home/s care provider/s service user/s present study palliative (xx) care social support data collection quality (x,xx) care job satisfaction focus group/s *full list of Noun + Noun collocations is in Handout #6
Step 1: Introduction	Matching exercise generated through facilities provided by the website **Quizlet.com**
Step 2: The lesson	Show students the concordance lines of each pair of collocations, point out how each pair co-occurs, point out their possible positions, discuss their meaning in contexts
Step 3: After lesson exercise	Multiple-choice exercise generated through facilities provided by the website **Quizlet.com**

REFERENCES

Aston, G. (1995). Corpora in language pedagogy: Matching theory and practice. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics* (pp. 257-270). Oxford: Oxford University Press.

Cobb, T. (1999). Breadth and depth of vocabulary acquisition with hands-on concordancing. *CALL*, 12, 345-360. Retrieved 2016 from: http://www.er.uqam.ca/nobel/r21270/cv/Breadth.htm

Conrad, S. (2005). Corpus linguistics and L2 teaching. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 393-409). Mahwah, NJ: Lawrence Erlbaum.

Coxhead, A. (2010). What can corpora tell us about English for Academic Purposes? In A. O'Keeffe and M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 458-470). London, UK: Routledge.

Gavioli, L., & Aston, G. (2001). Enriching reality: Language corpora in language pedagogy. *ELT Journal, 55*(3), 238-246.

Horst, M., Cobb, T., & Nicolae, I. (2005). Expanding academic vocabulary with an interactive on-line database. *Language Learning and Technology, 9*(2), 90-110.

Johns, T. (1991). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. *English Language Research Journal, 4*, 27-45.

Kozlowski, A. & Seymour, L. (2003). The importance of collocation in English language teaching. *Contact, 29*(3), 41-49.

Leech, G. (1986). Automatic grammatical analysis and its educational applications. In G. Leech and C. Candlin (Eds.), *Computers in English language teaching and research* (pp. 205-214). London, UK: Longman.

McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. London, UK: Routledge.

Pickard, V. (1994). Producing a concordanced-based self-access vocabulary package: Some problems and solutions. In L. Flowerdew & A. K. K. Tong (Eds.), *Entering text* (pp. 215-226). Hong Kong: Hong Kong University of Science and Technology Language Center.

Samburskiy, D. (2014). Corpus-informed pedagogical grammar of English: Pros and cons. *Procedia - Social and Behavioral Sciences, 154*, 263-267. doi:10.1016/j.sbspro.2014.10.148

Stevens, V. (1991). Concordance-based vocabulary exercises: A viable alternative to gap-filling. *English Language Research Journal*, 4, 47-61.

Thurstun, J., & Candlin, C. (1998). Concordancing and the teaching of the vocabulary of academic English. *English for Specific Purposes, 17*(3), 267-280.

Handout #1 **List of 100 Most Frequent Keywords in the SCNRA**

Rnk	Freq	Keyness	Keywords	Rnk	Freq	Keyness	Keywords
1	6600	25519.82	care	51	1169	4567.938	cancer
2	6540	25131.97	health	52	1140	2096.207	hospital
3	5739	21880.57	study	53	1130	1291.218	process
4	5115	34638.35	nurses	54	1118	327.44	different
5	4895	19121.01	patients	55	1094	2835.052	factors
6	3896	9554.67	family	56	1091	1825.716	results
7	3058	19288.67	participants	57	1084	2172.522	included
8	3019	13012.96	patient	58	1077	324.906	high
9	2991	17631.77	nursing	59	1069	2139.128	significant
10	2350	6138.553	data	60	1041	1569.071	quality
11	2284	4441.323	research	61	1022	2394.875	medical
12	2245	9997.944	mental	62	999	1818.681	relationship
13	2029	1204.074	used	63	993	1411.258	higher
14	1946	1864.834	children	64	984	976.429	management
15	1934	5787.167	reported	65	968	2640.939	pain
16	1930	9905.078	nurse	66	967	3786.383	experiences
17	1867	4677.497	parents	67	965	1089.642	needs
18	1859	2645.105	support	68	950	2120.62	associated
19	1841	5300.04	risk	69	933	1036.068	individual
20	1835	4931.322	studies	70	915	2586.099	identified
21	1771	1807.543	information	71	912	3053.136	sample
22	1704	222.11	being	72	902	1874.988	scale
23	1702	211.914	work	73	900	2095.496	positive
24	1611	3395.014	practice	74	895	1282.984	described
25	1579	2421.212	child	75	895	416.266	service
26	1574	535.667	use	76	890	3459.927	intervention
27	1543	529.857	each	77	868	2334.473	items
28	1534	1074.935	social	78	862	913.704	role
29	1520	4545.188	physical	79	847	1379.4	levels
30	1454	3443.88	analysis	80	831	7006.333	caregivers
31	1449	6802.294	illness	81	831	3172.437	symptoms
32	1412	2218.382	experience	82	795	722.911	team
33	1404	2994.305	knowledge	83	794	465.229	person
34	1387	2060.834	staff	84	793	1662.244	older
35	1384	1715.69	education	85	786	2083.571	relationships
36	1382	1609.993	based *hospital-based, evidence-based	86	782	1275.304	professional
37	1372	1825.418	using	87	782	801.966	provided
38	1338	772.635	group	88	779	3577.531	scores
39	1332	1407.442	level	89	769	278.103	mean
40	1292	801.095	important	90	765	500.801	provide
41	1292	294.311	life	91	764	624.705	groups
42	1276	1219.448	members	92	762	5013.064	interventions
43	1267	1699.484	age	93	760	3516.969	diagnosis
44	1256	5571.34	findings	94	759	3161.132	professionals
45	1239	452.568	found	95	756	970.714	questions
46	1204	3437.045	families	96	748	249.32	control
47	1202	303.634	need	97	744	1079.295	previous
48	1200	2597.629	treatment	98	743	491.069	evidence
49	1194	5322.114	clinical	99	727	1842.77	focus
50	1190	1304.372	services	100	725	932.45	population

Lexical Collocations in a Sample Corpus of Nursing Research Articles (SCNRA) Handout #2
100 Most Frequent Noun + Noun Collocations

No.	Nodes	Collocates	Freq.	No.	Nodes	Collocates	Freq.
1	health	(x,xx) care	1516	51	age	(x) gender	94
2	family	member/s	1075	52	coping	strategy/ies	94
3	health	service/s	514	53	workplace	spirituality	94
4	nursing	home/s	366	54	risk	management	93
5	care	provider/s	328	55	adolescent	(x) health	92
6	service	user/s	326	56	work	environment/s	88
7	family	caregiver/s	303	57	aim	(xx) study	87
8	health	(x) provider/s	280	58	health	behaviours	86
9	data	collection	264	59	knowledge	(x) skill/s	86
10	quality	(x,xx) care	261	60	team	members	86
11	nursing	practice	226	61	sleep	(x,xx) problem/s	86
12	job	satisfaction	225	62	staff	member/s	85
13	focus	group/s	222	63	pain	management	85
14	quality	(x) life	221	64	medication	(x) error/s	85
15	patient	education	201	65	caring	behaviors/our/ours	84
16	risk	assessment	187	66	children	(x,xx) adolescents	80
17	nurse	manager/s	185	67	content	analysis	79
18	sample	size/s	172	68	limitation/s	(xx) study	79
19	nursing	student/s	168	69	child	(x,xx) cancer	78
20	children	(x,xx) cancer	161	70	leadership	style/s	78
21	age	(x,xx) years	160	71	ethics	committee/s	78
22	risk	factor/s	157	72	alcohol	consumption	77
23	nursing	staff	155	73	risk	taking	76
24	patients	(x,xx) families	152	74	regression	(xx) analysis/es	76
25	parents	(x,xx) children	150	75	screening	tool/s	73
26	health	problems	145	76	care	units	72
27	health	status	142	77	relationship	quality	71
28	patient	safety	142	78	safety	planning	70
29	research	team	136	79	medication	adherence	69
30	data	analysis	135	80	outpatient	(xx) clinic/s	69
31	health	outcomes	134	81	smoking	cessation	68
32	health	literacy	133	82	health	crisis/es	66
33	providing	(x,xx) care	126	83	research	ethics	66
34	illness	belief/s	123	84	nurse	leader/s	66
35	childhood	cancer	121	85	risk	(x,xx) safety	66
36	medication	administration	115	86	intervention	(xx) group/s	66
37	inclusion	(xx) criteria	114	87	diabetes	(x) education	66
38	assessment	tool/s	112	88	patient	outcomes	64
39	care	settings	109	89	family	(x,xx)friends	62
40	health	system/s	109	90	nursing	interventions	62
41	health	issues	109	91	group	interview/s	62
42	control	group/s	105	92	oncology	patients	61
43	patient	satisfaction	104	93	symptoms	(x,xx) depression	60
44	research	question/s	102	94	training	programs/me/mes	60
45	review	board/s	100	95	literature	review	60
46	emergency	department/s	100	96	consent	form/s	60
47	age	group/s	97	97	diagnosis	(x,xx) treatment	59
48	response	rate/s	97	98	community	setting/s	59
49	anxiety	(x,xx) depression	95	99	discharge	education	59
50	family	functioning	94	100	nurse	staffing	58

Lexical Collocations in a Sample Corpus of Nursing Research Articles (SCNRA)
Handout #3

100 Most Frequent Adjective + Noun Collocations

No.	Nodes	Collocates	Freq.	No.	Nodes	Collocates	Freq.
1	mental	(x,xx) health	1699	51	tactile	massage	77
2	physical	(x,xx) health	385	52	social	capital	76
3	mental	illness/es	352	53	increased	risk	75
4	present	study	306	54	pediatric	(x) patients	75
5	palliative	(xx) care	304	55	mental	(x) problems	74
6	social	support	300	56	high	school	74
7	physical	activity	281	57	evidence-based	practice/s	73
8	mental	(x) service/s	254	58	oral	care	73
9	previous	studies	217	59	qualitative	(x) research	72
10	significant	difference/s	214	60	lower	(x,xx) level/s	72
11	chronic	(x,xx) illness/es	190	61	intensive	(x) unit/s	72
12	higher	(x,xx) score/s	187	62	paediatric	nurses	70
13	registered	nurse/s	186	63	adult	(x) health	69
14	older	adult/s	182	64	severe	(xx) pain	69
15	high	level/s	176	65	sensory	room/s	69
16	older	people	176	66	significant	correlation/s	67
17	mean	score/s	175	67	emotional	exhaustion	67
18	clinical	(x) practice	166	68	low	level/s	66
19	higher	level/s	160	69	ethical	approval	66
20	primary	care	158	70	semi/structured	interview/s	66
21	pediatric	oncology	156	71	educational	(xx) intervention/s	64
22	aged	(x) years	154	72	each	participant	63
23	critical	care	149	73	parental	presence	62
24	previous	research	147	74	educational	(xx) level/s	62
25	total	(x,xx) score/s	147	75	educational	(x) program/s/me/mes	61
26	current	study	141	76	social	network/s	60
27	acute	(x,xx) care	130	77	mean	(xx) years	59
28	intensive	(x,xx) care	125	78	affective	commitment	59
29	chronic	(x) condition/s	121	79	positive	outcomes	58
30	psychological	distress	118	80	emotional	support	58
31	primary	(x) caregiver/s	117	81	daily	living	57
32	systematic	review/s	117	82	cognitive	impairment	57
33	mean	age	110	83	thematic	(x) analysis	57
34	qualitative	(x) study/ies	110	84	social	cohesion	56
35	depressive	symptoms	107	85	negative	(x) effects	56
36	internal	consistency	99	86	mental	(x) crisis/es	55
37	daily	life/ves	93	87	mental	(x) issues	55
38	chronic	(x,xx) disease/s	91	88	different	types	55
39	demographic	(xx) characteristics	91	89	high	(x,xx) rate/s	55
40	everyday	life/ves	91	90	medical	records	55
41	primary	family	90	91	individual	(x) interviews	55
42	each	(x) item	88	92	positive	effect/s	55
43	descriptive	statistics	87	93	significant	relationship/s	54
44	marital	status	86	94	socioeconomic	status	54
45	psychiatric	nurses	83	95	emotional	distress	53
46	institutional	review	81	96	organisational	culture	53
47	clinical	setting/s	80	97	social	worker/s	52
48	institutional	(x) board/s	80	98	negative	(x) emotions	52
49	acute	(x,xx) setting/s	79	99	educational	attainment	52
50	surgical	patients	77	100	mental	(x) triage	51

Lexical Collocations in a Sample Corpus of Nursing Research Articles (SCNRA)

Handout #4

No.	Nodes	Collocates	Freq.	No.	Nodes	Collocates	Freq.
	50 Most Frequent <u>Noun + Verb</u> Collocations				**50 Most Frequent <u>Verb + Noun</u> Collocations**		
1	data	(x) collected	168	1	participate	(xx) study	113
2	study	(x) conducted	138	2	analys/zed	using	73
3	score/s	indicate/ed/ing	122	3	diagnosed	(x,xx) cancer	69
4	nurses	working	116	4	manage	(x,xx) health	59
5	participants	(x) asked	79	5	measured	using	56
6	score/s	range/s/d	78	6	manage	(x,xx) care	56
7	study	aims/ed	66	7	bereaved	(x) suicide	56
8	results	show/ed	65	8	diagnosed	(xx) patients	49
9	studies	conducted	64	9	provide	(x,xx) support	46
10	data	analysed/zed	63	10	diagnosed	(x) schizophrenia	45
11	interviews	(x) conducted	60	11	improve	(x,xx) quality	41
12	research	(x) needed	59	12	access	(x,xx) services	39
13	study	(xx) explore	57	13	provided	(x,xx) information	38
14	findings	suggest	56	14	assessed	using	38
15	scale	ranged/ing	54	15	collected	(x) data	35
16	studies	(x) shown	52	16	reported	feeling	33
17	study	(x) approved	51	17	conducted	(x) using	32
18	factors	(x) influence/d	49	18	provide	(x,xx) evidence	31
19	information	(x,xx) provided	48	19	reduce	(x) risk	30
20	consent	(x) obtained	47	20	participated	(xx) study	30
21	participants	(x) recruited	42	21	performed	(x,xx) using	29
22	studies	(examined	40	22	included	(xx) review	28
23	participants	(x) informed	39	23	received	(x,xx) education	27
24	factors	include/ing	39	24	provide	(x,xx) insight/s	26
25	factors	) associated	38	25	received	(x,xx) training	26
26	questions	(x,xx) asked	38	26	living	(x) chronic	26
27	themes	(x) identified	38	27	completed	questionnaire/s	26
28	methods	(x,xx) used	38	28	examine	relationship/s	25
29	results	suggest	37	29	ranged	(x) years	25
30	patients	(x) admitted	35	30	expressed	(x,xx) concern/s	24
31	analysis	(x) performed	35	31	receiving	(x,xx) treatment	24
32	participants	completed	34	32	included	(x) following	22
33	analysis	showed	34	33	expressed	(x) need	22
34	analysis	(x) conducted	34	34	associated	(x,xx) suicide	21
35	findings	indicate	34	35	living	(xx) illness	21
36	analyses	(xx) using	33	36	conducted	(x,xx) nterviews	20
37	findings	show/ed	32	37	describe	(x,xx) experiences	20
38	factors	(x,xx) affect	30	38	provided	(x,xx) opportunity	19
39	variables	included/ing	29	39	provided	(x,xx) consent	19
40	respondents	(x) reported	29	40	needed	(x) help	19
41	decisions	(x,xx) made	29	41	address	(x,xx) needs	19
42	families	experiencing	28	42	received	(x,xx) treatment	18
43	motivation	(x) manage	28	43	examine	(x,xx) differences	18
44	nurses	play	27	44	described	feeling	17
45	children	diagnosed	27	45	access	(x,xx) resources	17
46	services	(x,xx) provided	27	46	informed	(xx) decisions	17
47	results	indicate	27	47	explore	(x,xx) experiences	17
48	study	(xx) investigate	25	48	experiencing	(x,xx) illness	17
49	patients	(x) undergoing	25	49	translated	(x,xx) English	17
50	scale	(x) developed	25	50	identified	(x,xx) themes	16

Lexical Collocations in a Sample Corpus of Nursing Research Articles (SCNRA)
200 most frequent collocations found in SCNRA (repeated nodes)   Handout#5

No.	Nodes	Collocates	No.	Nodes	Collocates
1	mental	(*ill, and physical*) health	51	previous	research
2	health	care	52	total	(*mean, health literacy*) score/s
3	family	member/s	53	health	problems
4	health	service/s	54	health	status
5	physical	(*ill, and mental*) health	55	patient	safety
6	nursing	home/s	56	current	study
7	mental	illness/es	57	study	(*was*) conducted
8	care	provider/s	58	research	team
9	service	user/s	59	data	analysis
10	present	study	60	health	outcomes
11	palliative	(*and supportive*) care	61	health	literacy
12	family	caregiver/s	62	acute	(*psychiatric, and primary*) care
13	social	support	63	informed	consent
14	physical	activity	64	providing	(*quality, efficient health*) care
15	health	(*care*) provider/s	65	intensive	(*follow-up, support and*) care
16	data	collection	66	illness	belief/s
17	quality	(*of, of nursing*) care	67	score/s	indicate/ed/ing
18	mental	(*health*) service/s	68	childhood	cancer
19	nursing	practice	69	chronic	(*disease*) condition/s
20	job	satisfaction	70	psychological	distress
21	focus	group/s	71	primary	(*family*) caregiver/s
22	quality	(*of*) life	72	systematic	review/s
23	previous	studies	73	nurses	working
24	significant	difference/s	74	medication	administration
25	patient	education	75	inclusion	(*and exclusion*) criteria
26	chronic	(*physical, conditions for*) illness/es	76	participate	(*in the*) study
27	higher	(*and lower*) score/s	77	assessment	tool/s
28	risk	assessment	78	mean	age
29	registered	nurse/s	79	qualitative	(*research*) study/ies
30	nurse	manager/s	80	care	settings
31	older	adult/s	81	health	system/s
32	high	level/s	82	health	issues
33	older	people	83	depressive	symptoms
34	mean	score/s	84	control	group/s
35	sample	size/s	85	patient	satisfaction
36	data	(*were*) collected	86	research	question/s
37	nursing	student/s	87	emergency	department/s
38	clinical	(*nursing*) practice	88	review	board/s
39	children	(*with, diagnosed with*) cancer (*of -, ranged between -*)	89	internal	consistency
40	age	years	90	strongly	agree
41	higher	level/s	91	age	group/s
42	primary	care	92	response	rate/s
43	risk	factor/s	93	anxiety	depression
44	pediatric	oncology	94	age	(*and*) gender
45	nursing	staff	95	coping	strategy/ies
46	aged	(*under -*) years	96	family	functioning
47	statistically	significant	97	workplace	spirituality
48	patients	(*and ,and their*) families	98	daily	life/ves
49	parents	(*of ,and their*) children	99	risk	management
50	critical	care	100	adolescent	(*mental*) health

No.	Nodes	Collocates	No.	Nodes	Collocates
		(*physical, obstructive*			
101	chronic	*pulmonary*) disease/s	151	used	(*to*) measure
102	demographic	characteristics	152	paediatric	nurses
103	everyday	life/ves	153	safety	planning
104	primary	family	154	adult	(*mental*) health
105	each	(*questionnaire*) item	155	diagnosed	(*with, with cervical*) cancer
106	work	environment/s	156	medication	adherence
107	aim	(*of this*) study	157	outpatient	(*heart failure*) clinic/s
108	descriptive	statistics	158	sensory	room/s
109	health	behaviours	159	severe	(*level of*) pain
110	knowledge	(*and*) skill/s	160	strongly	disagree/d
111	marital	status	161	being	able
112	sleep	(*and appetite*) problem/s	162	smoking	cessation
113	team	members	163	emotional	exhaustion
114	medication	(*administration*) error/s	164	positive	(*and*) negative
115	pain	management	165	score/s	(*indicated, indicates a*) higher
116	staff	member/s	166	significant	correlation/s
117	caring	behaviors/our/ours	167	diabetes	(*management*) education
118	psychiatric	nurses	168	ethical	approval
119	institutional	review	169	health	crisis/es
		(*and, and young*)			
120	children	adolescents	170	intervention	(*and control*) group/s
121	clinical	setting/s	171	low	level/s
122	institutional	(*review*) board/s	172	nurse	leader/s
123	acute	(*care hospital*) setting/s	173	research	ethics
124	content	analysis	174	risk	(*and, assessment and*) safety
125	limitation/s	(*of the*) study	175	semi/structured	interview/s
126	participants	(*were*) asked	176	study	aims/ed
		(*with, diagnosed with*)			
127	child	cancer	177	results	show/ed
128	ethics	committee/s	178	significantly	(*associated with*) higher
129	leadership	style/s	179	educational	(*programs and*) intervention/s
130	score/s	range/s/d	180	patient	outcomes
131	alcohol	consumption	181	studies	(*were, have been*) conducted
132	surgical	patients	182	data	(*were*) analysed/zed
133	tactile	massage	183	each	participant
		(*and meditation*)			
134	regression	analysis/es	184	educational	(*and income*) level/s
135	risk	taking	185	family	(*and, members and*) friends
136	social	capital	186	group	interview/s
137	increased	risk	187	nursing	interventions
138	pediatric	(*oncology*) patients	188	parental	presence
139	high	school	189	statistically	(*significant*) difference/s
140	mental	(*health*) problems	190	community	mental
141	analys/zed	using	191	educational	program/s/me/mes
	evidence-				
142	based	practice/s	192	oncology	patients
143	oral	care	193	consent	form/s
144	screening	tool/s	194	interviews	(*were*) conducted
145	care	units	195	literature	review
146	higher	(*scores*) indicate/s/ing	196	social	network/s
147	intensive	(*care*) unit/s	197	symptoms	(*of, such as*) depression
		(*education, baseline*			
148	lower	*energy*) level/s	198	training	programs/me/mes
149	qualitative	(*exploratory*) research	199	affective	commitment
150	relationship	quality	200	community	setting/s

Lexical Collocations in a Sample Corpus of Nursing Research Articles (SCNRA)
200 Collocations ranked by frequency of occurrence according to each node

No.	Nodes	Collocates	No.	Nodes	Collocates
		(*ill, and physical*)			
1	mental	health	51	medication	administration
2	health	care	52	inclusion	(*and exclusion*) criteria
3	family	member/s	53	participate	(*in the*) study
4	physical	(*and mental*) health	54	assessment	tool/s
5	nursing	home/s	55	qualitative	(*research*) study/ies
6	care	provider/s	56	depressive	symptoms
7	service	user/s	57	control	group/s
8	present	study	58	emergency	department/s
9	palliative	(*and supportive*) care	59	review	board/s
10	social	support	60	internal	consistency
11	data	collection	61	strongly	agree
12	quality	(*of, of nursing*) care	62	response	rate/s
13	job	satisfaction	63	anxiety	depression
14	focus	group/s	64	coping	strategy/ies
15	previous	studies	65	workplace	spirituality
16	significant	difference/s	66	daily	life/ves
17	patient	education	67	adolescent	(*mental*) health
18	chronic	illness/es	68	demographic	characteristics
19	higher	score/s	69	each	(*questionnaire*) item
20	risk	assessment	70	work	environment/s
21	registered	nurse/s	71	aim	(*of this*) study
22	nurse	manager/s	72	descriptive	statistics
23	older	adult/s	73	knowledge	(*and*) skill/s
24	high	level/s	74	marital	status
25	mean	score/s	75	sleep	(*and appetite*) problem/s
26	sample	size/s	76	team	members
27	clinical	(*nursing*) practice	77	pain	management
28	children	(*with*) cancer	78	staff	member/s
		(*of -, ranged between*			
29	age	-) years	79	caring	behaviors/our/ours
30	primary	care	80	psychiatric	nurses
31	pediatric	oncology	81	institutional	review
32	aged	(*under -*) years	82	content	analysis
33	statistically	significant	83	limitation/s	(*of the*) study
34	patients	(*and their*) families	84	participants	(*were*) asked
35	parents	(*and their*) children	85	child	(*with, diagnosed with*) cancer
36	critical	care	86	leadership	style/s
37	total	(*mean*) score/s	87	alcohol	consumption
38	current	study	88	surgical	patients
39	study	(*was*) conducted	89	tactile	massage
40	research	team	90	regression	(*and meditation*) analysis/es
41	acute	(*and primary*) care	91	increased	risk
42	informed	consent	92	analys/zed	using
43	providing	(*efficient health*) care	93	evidence-based	practice/s
44	intensive	(*support and*) care	94	oral	care
45	illness	belief/s	95	screening	tool/s
46	score/s	indicate/ed/ing	96	lower	(*education, baseline energy*) level/s
47	childhood	cancer	97	relationship	quality
48	psychological	distress	98	used	(*to*) measure
49	systematic	review/s	99	paediatric	nurses
50	nurses	working	100	safety	planning

No.	Nodes	Collocates	No.	Nodes	Collocates
101	adult	(*mental*) health	151	cancer	treatment
102	diagnosed	(*with cervical*) cancer	152	cross-sectional	(*research, survey research*) design
103	outpatient	(*heart failure*) clinic/s	153	education	programs/me/mes
104	sensory	room/s	154	antipsychotic	medication
105	severe	(*level of*) pain	155	grounded	theory
106	being	able	156	personal	(*and professional*) experience/s
107	smoking	cessation	157	person-centred	care
108	emotional	exhaustion	158	sex	ratio/s
109	positive	(*and*) negative (*management*)	159	statistical	analysis/es
110	diabetes	education	160	factors	(*that*) influence/d
111	ethical	approval	161	mortality	rate/s
112	intervention	(*and control*) group/s	162	information	(*was, could be*) provided
113	low	level/s	163	person	centredness
114	semi/structured	interview/s	164	reliability	(*and, and construct*) validity
115	results	show/ed (*associated with*)	165	self-efficacy	expectations
116	significantly	higher (*programs and*)	166	associated	(*with, with a*) higher
117	educational	intervention/s (*were, have been*)	167	correlation	coefficient/s
118	studies	conducted	168	deterioration	event/s
119	group	interview/s	169	geriatric	team
120	parental	presence	170	majority	(*of, of the*) participants
121	oncology	patients	171	professional	identity
122	consent	form/s	172	provide	(*appropriate, the needed*) support
123	interviews	(*were*) conducted	173	activities	(*of*) daily
124	literature	review (*of, such as*)	174	depression	(*and*) anxiety
125	symptoms	depression	175	factor	structure
126	training	programs/me/mes	176	logistic	(*and linear*) regression
127	affective	commitment	177	longitudinal	(*aging*) study/ies
128	community	setting/s	178	multiple	(*linear*) regression
129	diagnosis	(*specific*) treatment	179	risk-assessment	(*and safety*) planning
130	disagree	(*to strongly*) agree	180	unit	operation
131	discharge	education	181	questionnaire	item/s
132	manage	(*their physical*) health	182	self-care	agency
133	cognitive	impairment	183	therapeutic	conversation/s
134	thematic	(*content*) analysis	184	suicide	death/s
135	bereaved	(*by*) suicide	185	important	role
136	findings	suggest	186	lack	(*of*) knowledge
137	measured	using	187	theoretical	framework
138	negative	(*health*) effects	188	consistent	(*with, with the*) previous
139	practice	environment/s	189	improve	(*the, access and*) quality
140	different	types (*qualitative*)	190	inpatient	care
141	individual	interviews	191	tertiary	medical
142	medical	records	192	caregiver	(*stress and*) burden
143	hospital	stay/s	193	randomized	(*controlled, controlled clinical*) trial/s
144	scale	ranged/ing	194	access	(*the, mental health*) services
145	socioeconomic	status	195	crisis	response/s
146	healthcare	professionals	196	meaning	units
147	organisational	culture	197	online	supplementary
148	transcribed	verbatim	198	perceived	(*social, higher family*) support
149	condition	management	199	psychometric	properties
150	interview	data	200	validity	(*and*) reliability

Appendix D

Analysis of the Tryout Test

The Analysis of the Tryout of the Nursing Collocation Test

The first two parts of the test were tried out with 38 fourth year Nursing students of academic year 3/2016 at SUT. The first part consisted of 30 items of a multiple choice test. The second part consisted of 20 items of a gap filling test in which the items were sub-divided into four groups of five items. The test descriptions were explained in Thai to the test-takers to make sure they have some ideas about the test they were doing. The results of the test were analyzed to find out three important elements of the test. These elements are the test Difficulty Index, Discrimination Index, and Reliability of the test or the internal consistency coefficients using Kuder-Richardson 20 (KR-20). The results are shown in the table below.

The Nursing Collocation Test Part I

			Reliability Statistics			
			Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items	
			.687	.708	30	
			Item-Total Statistics			
Item	Difficulty Index	Discrimination Index	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
1	92.11	0.3	18.1842	15.938	.361	.	.673
2	50.00	0.7	18.6053	14.894	.422	.	.661
3	89.47	**0.1**	18.2105	16.279	.169	.	.682
4	76.32	0.6	18.3421	14.718	.574	.	.651
5	**7.89**	**0.1**	19.0263	16.621	.047	.	.688
6	28.95	0.6	18.8158	15.344	.346	.	.669
7	94.74	0.2	18.1579	16.083	.367	.	.675
8	50.00	0.3	18.6053	15.975	.140	.	.686
9	55.26	0.4	18.5526	16.308	.058	.	.693
10	60.53	0.4	18.5000	15.662	.227	.	.678
11	31.58	**0.1**	18.7895	16.927	-.090	.	**.704**
12	60.53	0.2	18.5000	16.257	.074	.	.691
13	57.89	0.3	18.5263	15.716	.210	.	.680
14	84.21	0.6	18.2632	15.388	.440	.	.665
15	47.37	0.8	18.6316	14.780	.453	.	.658
16	47.37	0.3	18.6316	16.185	.088	.	.691
17	84.21	0.4	18.2632	15.226	.498	.	.661
18	81.58	0.6	18.2895	14.806	.608	.	.651
19	57.89	**0.1**	18.5263	16.526	.006	.	.697
20	65.79	0.6	18.4474	15.119	.388	.	.665
21	78.95	0.3	18.3158	16.222	.122	.	.686
22	39.47	**0.1**	18.7105	16.211	.086	.	.691
23	92.11	**0.1**	18.1842	16.262	.210	.	.681
24	97.37	**0.1**	18.1316	16.604	.128	.	.685
25	34.21	0.2	18.7632	16.402	.043	.	.694
26	68.42	**0**	18.4211	16.899	-.083	.	**.703**
27	55.26	0.4	18.5526	15.335	.307	.	.671
28	65.79	0.6	18.4474	14.903	.449	.	.659
29	71.05	0.3	18.3947	16.245	.093	.	.689

30	84.21	**0.1**	18.2632	16.199	.156	.	.683
Average	**63.68**	**.33**					

Difficulty Index

The tryout test's results of the first part 30 items of multiple choice test reveal that the difficulty index ranges from 7.89 to 97.37 with the average at 63.68. This indicates that the overall test is relatively easy. However, there are four items that the difficulty index values indicate that one of them is a very difficult item and three of them are difficult items. These items are: items no. 5, 6, 11, and 25. The item no. 5 has the lowest difficulty index value at 7.89 indicating that it is very difficult. The items no. 6, 11, and 25 have the difficulty index value at 28.95, 31.58, and 34.21 respectively. Their difficulty indices indicate that they are difficult items. On the other hand, there are also some easy items such as items no. 24 (97.37), 7 (94.74), and 1 and 23 (92. 11). Therefore these items will be examined and revised.

Discrimination Index

There are items that the discrimination index shows that they are needed to be examined and revised. These items are items with the discrimination index value less than 0.2 which indicates poor discrimination ability. The items are item no. 26 (with discrimination index at 0), and items no. 3, 5, 11, 19, 22, 23, 24, and 30 (with discrimination index at 0.1).

Reliability of the Test

The internal consistency coefficients of the test, Kuder-Richardson 20 (KR-20), show the Cronbach's Alpha at .687 which indicates moderate reliability. The Item-total Statistics also show two items which will increase the reliability of the test. These two items are items no. 11 and 26. The two items are clearly problematic as they also appear in both the difficulty index and discrimination index. Therefore, these two items will be examined and revised.

Examining and Revising the Test

According to difficulty index: items no. 5, 6, 11, and 25

5. It is found that *smoking* _____ is difficult because the person is addicted to nicotine in cigarettes.

    a) cessation   b) commotion   c) dissolution   d) suspension

    This item is the most difficulty item with most test takers chosen choices C and D instead of the correct choice A. The appropriate way to revise this item is to change the distractors in C and D.

    a) cessation   b) commotion   c) **expiration**   d) **interruption**

6. Many studies have shown that *job* _____ is directly related to the work's atmosphere. People tend to be happier with their work when they are surrounded with understanding co-workers.

a) ~~attainment~~   b) ~~conclusion~~   c) satisfaction   d) gratification

For this item, the most chosen choices were A and B instead of the correct choice C. Therefore, changes are needed with the choices A and B.

a) **assumption**   b) **completion**   c) satisfaction   d) gratification

11. A supportive *work* _____ is critical to ensure that nurses are given the opportunity to participate in the decision-making processes.

a) environment   b) ~~situation~~   c) surrounding   d) circumstance

This item, choice B was the most chosen instead of the correct choice A. Therefore, choice B is changed.

a) environment   b) **location**   c) surrounding   d) circumstance

25. Parents and children should be offered the choice of *parental* _____ during anesthesia induction. Most children have less anxiety when their parents are there with them.

a) presence   b) ~~arrival~~   c) audience   d) ~~existence~~

Choice B and D were most chosen by the test takers. Therefore, changes are needed.

a) presence   b) **entrance**   c) audience   d) **survival**

According to discrimination index: item no. 26 (at 0); items no. 3, 5, 11, 19, 22, 23, 24, and 30 (at <0.2). Items no. 5 and 11 have been revised according to difficulty index above.

26. A high level of *affective* _____ is related to high employee retention. When the employees feel secure with their job, they gain a sense of belongings to the organization.

a) ~~adherence~~   b) faithfulness   c) commitment   d) ~~deference~~

This item has the discrimination index at 0 which indicates that it does not have ability to discriminate the upper half and the lower half test takers. When look at the item's results, it shows that the lower haft test takers had chosen the correct answer, choice C, more than the upper half test takers. The choices most chosen by the upper half test takers were A and D. Therefore, revision is needed for this item.

a) **observance**   b) faithfulness   c) commitment   d) **admiration**

3. An *evidence-based* _____ is a useful approach for systematic decision-making processes in providing clinical services.

a) tradition   b) ~~mode~~   c) preparation   d) practice

In this item, the test takers in the lower half perform almost as good as those in the upper half by choosing the correct choice D. The distractor B was not chosen by any test takers. Therefore, the change of choice B is needed to make it more attractive.

a) tradition   b) **method**   c) preparation   d) practice

19. Adequate *diabetes* _____ is a useful way for patients to perform better in their self-care.

a) ~~training~~   b) ~~tutoring~~   c) instruction   d) education

Choice A was the most popular choice for both upper and lower groups. Choice B was not chosen. Therefore, they need to be changed.

a) **exercise**   b) **practice**   c) instruction   d) education

22. The lack of support from immediate *family* _____ namely spouses and children can cause stress for caregivers.

a) ~~associates~~   b) fellows   c) members   d) admirers

Choice A was also chosen the most and the upper group chosen more than the lower group although it was the incorrect choice. Therefore it needs to be changed.

a) **companions**   b) fellows   c) members   d) admirers

23. Migraines are a common type of headache that can cause *severe* _____ on one or both sides of the head.

a) anger   b) pain   c) ~~frustration~~   d) irritation

this item, all of the test takers in the lower group chosen the correct choice, while two of the upper group chosen choice C. Therefore, change is needed for the choice C.

a) anger   b) pain   c) **prevention**   d) irritation

24. The foot self-care *educational* _____ by introducing new practices demonstrated effectiveness in increasing foot self-care knowledge.

a) ~~conciliation~~   b) construction   c) ~~intrusion~~   d) intervention

In this item, both groups chosen the correct choice D. Choices A and C were not chosen at all. Therefore, changes are needed for choices A and C.

a) **meditation**   b) construction   c) **involvement**   d) intervention

According to Reliability Index: items 11 and 26.

These two items have been revised above by which item no. 11 has been done along the revision according to the difficulty index and item no. 26 has been revised along with the revision according to the discrimination index.

The Nursing Collocation Test Part II

				Reliability Statistics			
			Cronbach's Alpha		Cronbach's Alpha Based on Standardized Items		N of Items
			.602		.596		20
				Item-Total Statistics			
Item	Difficulty Index	Discrimination Index	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
1	76.32	0.4	10.3158	9.735	.158	.	.596
2	71.05	0.3	10.3684	9.644	.171	.	.594
3	63.16	0.2	10.4474	9.767	.111	.	.603
4	76.32	0.3	10.3158	9.735	.158	.	.596
5	50.00	0.6	10.5789	8.683	.471	.	.549
6	76.32	0.3	10.3158	9.411	.283	.	.580
7	50.00	0.5	10.5789	9.115	.318	.	.573
8	71.05	0.4	10.3684	9.644	.171	.	.594
9	63.16	0.4	10.4474	9.443	.221	.	.588
10	60.53	0.3	10.4737	9.499	.197	.	.591
11	34.21	0.2	10.7368	9.821	.097	.	.605
12	47.37	0.3	10.6053	9.543	.175	.	.594
13	60.53	0.4	10.4737	9.770	.107	.	.604
14	55.26	0.4	10.5263	9.716	.120	.	.602
15	42.11	0.9	10.6579	8.664	.486	.	.547
16	86.84	0.2	10.2105	10.117	.051	.	.606
17	36.84	0.4	10.7105	9.238	.292	.	.577
18	39.47	0.5	10.6842	9.519	.190	.	.592
19	21.05	0.4	10.8684	9.523	.255	.	.584
20	26.32	0.3	10.8158	10.046	.035	.	.612
Average	**55.4**	**0.39**					

Difficulty Index

The try out test's results of the second part 20 items of a gap-filling test reveal that the difficulty index ranges from 21.05 to 86.84 with the average at 55.4. This indicates that the overall test for this part is moderate, neither too difficult nor too easy.

Discrimination Index

The discrimination index shows that there is no item that has the discrimination index value less than 0.2 with the average value at 0.39 which indicates good items.

Reliability of the Test

The internal consistency coefficients of the test, Kuder-Richardson 20 (KR-20), show the Cronbach's Alpha at .602 which indicates moderate reliability. The Item-total Statistics do not show any items which will significantly increase the reliability of test if deleted.

Since the format of the test in this part is gap-filling with collocates provided for each set of five pairs, coupled with the three indices above indicate that the test is in satisfactory level, only minor changes will be made for this part of the test.

Appendix E

Nursing collocation Test

The test consists of three parts:

Part I, 30 items of a multiple-choice format;

Part II, 20 items of a gap-filling format; and

Part III, 10 items of a short sentence writing task.

Part I: Multiple choice (30 items)

Instructions: choose the most appropriate pair of the collocations in each item.

1. People who have *mental* _____ problems such as depression and dementia can face the risk of poor physical condition.

a) situation   b) health   c) fitness   d) wealth

2. Persons with diabetes need support from their health *care* _____ to guide them in their self-management skills.

a) provider   b) contractor   c) sponsor   d) giver

3. An *evidence-based* _____ is a useful approach for systematic decision-making processes in providing clinical services.

a) tradition   b) mode   c) preparation   d) practice

4. Young adults with *chronic* _____ such as cancer and diabetes may experience difficulties with both physical and mental development.

a) syndromes   b) troubles   c) illnesses   d) viruses

5. It is found that *smoking* _____ is difficult because the person is addicted to nicotine in cigarettes.

a) cessation   b) commotion   c) expiration   d) interruption

6. Many studies have shown that *job* _____ is directly related to the work's atmosphere. People tend to be happier with their work when they are surrounded with understanding co-workers.

a) assumption   b) completion   c) satisfaction   d) gratification

7. A core skill required for all *health* _____ workers is being able to have an informative talk with the patients when providing care.

a) attention   b) care   c) caution   d) security

8. Individuals use various *coping* _____ such as self-talk, writing, and consulting others when confronting sudden life-threatening situations.

a) strategies   b) policies   c) campaigns   d) actions

9. At an individual level, *workplace* _____ is related to the combination of personal beliefs and professional growth.

a) holiness   b) religiousness   c) essentiality   d) spirituality

10. Patients' *demographic* _____ such as age, education, gender, and residence patterns and living arrangements may contribute to care availability.

a) potentials   b) landscapes   c) characteristics   d) properties

11. A supportive *work* _____ is critical to ensure that nurses are given the opportunity to participate in the decision-making processes.

a) environment   b) location   c) surrounding   d) circumstance

12. Lack of energy, feeling drowsy, and *sleep* _____ are the most common and distressing symptoms of various illnesses.

a) weaknesses   b) obstacles   c) troubles   d) problems

13. Professional nursing interest in *caring* _____ such as attentive listening, comforting, honesty, and patience has existed since the Nightingale era.

a) actions   b) behaviors   c) conducts   d) operations

14. People's *socioeconomic* _____ such as income, education, and health can act as a critical determinant factor for their living condition.

a) peculiarity   b) honor   c) status   d) prestige

15. The frequency and amount of usual *alcohol* _____ had a linear dose–response relationship with suicide death.

a) consumption   b) absorption   c) obsession   d) ingestion

16. Sufficient information and care can reduce the anxiety of *surgical* _____ both before and after the operation.

a) sufferers   b) circumstances   c) casualties   d) patients

17. Women of child-rearing age had higher *mortality* _____ than males due to the risks they face in the process of giving birth.

a) speeds   b) rates   c) paces   d) scales

18. The use of *assessment* _____ namely the pain scale, fall risk scale, and the depression scale can markedly be beneficial in a process of treatment.

a) machines   b) gadgets   c) tools   d) campaigns

19. Adequate *diabetes* _____ is a useful way for patients to perform better in their self-care.

a) training   b) tutoring   c) instruction   d) education

20. All respondents provided written *informed* _____ accepting that they had been included in the project and were interviewed in a safe and private place.

a) approval   b) consent   c) permission   d) contract

21. Tooth brushing is an effective method of *oral* _____ to reduce dental plaque and bacteria.

a) care   b) repair   c) conservation   d) protection

22. The lack of support from immediate *family* _____ namely spouses and children can cause stress for caregivers.

a) companions   b) fellows   c) members   d) admirers

23. Migraines are a common type of headache that can cause *severe* _____ on one or both sides of the head.

a) anger   b) pain   c) prevention   d) irritation

24. The foot self-care *educational* _____ by introducing new practices demonstrated effectiveness in increasing foot self-care knowledge.

a) meditation   b) construction   c) involvement   d) intervention

25. Parents and children should be offered the choice of *parental* _____ during anesthesia induction. Most children have less anxiety when their parents are there with them.

a) presence   b) entrance   c) audience   d) survival

26. A high level of *affective* _____ is related to high employee retention. When the employees feel secure with their job, they gain a sense of belongings to the organization.

a) observance   b) faithfulness   c) commitment   d) admiration

27. Senior and specialist nurses in *community* _____ are involved in the development and implementation of guidelines and policies.

a) upbringings   b) settings   c) circumstances   d) surroundings

28. Moderate *cognitive* _____ influences the risk of a fall and loss of motor function.

a) disaster   b) weakness   c) infection   d) impairment

29. To adequately inform patients and their relatives, education skills among *healthcare* _____ namely doctors and nurses should be enhanced.

a) institutions   b) professionals   c) establishments   d) analysts

30. Caregivers can experience *negative* _____ such as fear and uncertainty about the future care of their loved one.

a) returns   b) estimations   c) requests   d) emotions

Nursing collocation Test

Part II: Gap-filling (20 items)

Instructions: Fill in each blank with the most appropriate pair of the collocations given in the box.

a. *beliefs*	b. *clinics*	c. *health*	d. *practice*	e. *status*

1. The key things for *physical* _____ are diet, exercise, and sleep.

2. The number of children who attend psychiatric *outpatient* _____ is increasing as a result of attention deficit disorder and autism.

3. Factors such as education, age, and *marital* _____ are essential to job performance.

4. In general, *clinical* _____ needs to focus on building a trustful relationship to the patients.

5. There is a need to understand *illness* _____ on a family level to be able to develop suitable caring interventions.

a. *administration*	b. *criteria*	c. *distress*	d. *risk*	e. *support*

6. Depression is associated with *increased* _____ of problem drinking in later life.

7. The greatest gain in decreased *psychological* _____ occurs in families who have been moved out of overcrowded situations.

8. The nurses' work environment is complex, and *social* _____ is one contributing factor to this complexity.

9. Bar-code-assisted *medication* _____ systems are designed to reduce medication errors.

10. Fifteen mental nurses who met *inclusion* _____ with aged between 33 and 58 years were interviewed.

a. *care*	b. *expectations*	c. *experiences*	d. *planning*	e. *symptoms*

11. Recovery-oriented *safety* _____ focuses on the person's strengths, resources, and capabilities, as well as fostering engagement and a shared responsibility.

12. Barriers to sleep in *critical* _____ are commonly associated with patient monitoring and treatment.

13. Neighborhood-level income inequality is associated with higher levels of *depressive* _____ among adolescent girls.

14. Developmental activities not only widen *personal* _____, but also contribute to better organizational quality.

15. A person who has high *self-efficacy* _____ tends to be a person who wants to achieve goals with high outcomes.

a. *education*	b. *events*	c. *identity*	d. *management*	e. *users*

16. A common strategy for *pain* _____ used by parents was the use of analgesic drugs such as paracetamol and aspirin.

17. Intensive *patient* _____ was the most used strategy to make patients cooperate in the treatment. Once they understand the processes involved, they feel less stressed.

18. Many *service* _____ experience interacting with doctors as discouraging and disempowering. They feel that they know very little about their health problems.

19. The surgical team expressed a strong *professional* _____ and took pride in what they were doing.

20. Nurses are key players in identifying and responding to *deterioration* _____ to escalate the level of care essential to address specific needs of patients.

Part III: Short sentence writing (10 items)

Instructions: Write a meaningful sentence containing the given collocation pair.

1. nursing home

2. primary care

3. nurse manager

4. psychiatric nurses

5. hospital stay

6. sensory rooms

7. emotional exhaustion

8. positive outcomes

9. medical records

10. relationship quality

*********END OF THE TEST*********

Nursing collocation Test

Answer sheet

Part I: Multiple choice (30 items)

Instructions: Mark a cross (X) in each box that corresponds with the most appropriate answer for each item.

No.	a	b	c	d	No.	a	b	c	d	No.	a	b	c	d
1					11					21				
2					12					22				
3					13					23				
4					14					24				
5					15					25				
6					16					26				
7					17					27				
8					18					28				
9					19					29				
10					20					30				

Part II: Gap-filling (20 items)

Instructions: Mark a cross (X) in each box that corresponds with the most appropriate answer for each item.

No.	a	b	c	d	e	No.	a	b	c	d	e
1						11					
2						12					
3						13					
4						14					
5						15					
6						16					
7						17					
8						18					
9						19					
10						20					

Part III: Short sentence writing (10 items)

Instructions: Write a meaningful sentence containing the given collocation pair.

1. nursing home:

_____

_____

2. primary care:

_____

_____

3. nurse manager:

_____

_____

4. psychiatric nurses:

_____

_____

5. hospital stay:

_____

_____

6. sensory rooms:

_____

_____

7. emotional exhaustion:

_____

_____

8. positive outcomes:

_____

_____

9. medical records:

_____

_____

10. relationship quality:

_____

_____

☺☺☺☺☺THANK YOU☺☺☺☺☺

Appendix F

Keyword List of SCNRA

Rank	Freq	Keyness	Keywords	Rank	Freq	Keyness	Keywords
1	6600	25519.82	care	360	239	820.52	ethical
2	6540	25131.97	health	361	237	308.53	describe
3	5739	21880.57	study	362	236	353.28	specifically
4	5115	34638.35	nurses	363	235	453.5	childhood
5	4895	19121.01	patients	364	234	227.68	colleagues
6	3896	9554.67	family	365	233	364.58	registered
7	3058	19288.67	participants	366	232	474.91	risks
8	3019	13012.96	patient	367	227	346.33	healthy
9	2991	17631.77	nursing	368	226	235.41	actions
10	2350	6138.553	data	369	226	484.65	transition
11	2284	4441.323	research	370	225	415.56	approximately
12	2245	9997.944	mental	371	225	535.89	institutional
13	2029	1204.074	used	372	224	1065.4	questionnaires
14	1946	1864.834	children	373	223	1021.6	baseline
15	1934	5787.167	reported	374	223	1635.6	medications
16	1930	9905.078	nurse	375	223	319.57	tasks
17	1867	4677.497	parents	376	222	353.81	approaches
18	1859	2645.105	support	377	222	272.16	program
19	1841	5300.04	risk	378	221	430.4	influenced
20	1835	4931.322	studies	379	220	509.95	practitioners
21	1771	1807.543	information	380	220	354.16	receiving
22	1704	222.11	being	381	219	409.14	demonstrated
23	1702	211.914	work	382	219	469.88	interaction
24	1611	3395.014	practice	383	219	255.24	selected
25	1579	2421.212	child	384	216	308.7	core
26	1574	535.667	use	385	216	470.29	effectiveness
27	1543	529.857	each	386	216	1094.3	physicians
28	1534	1074.935	social	387	216	348.56	promote
29	1520	4545.188	physical	388	216	892.87	therapeutic
30	1454	3443.88	analysis	389	215	275.48	framework
31	1449	6802.294	illness	390	215	755.23	prevalence
32	1412	2218.382	experience	391	215	282.59	reflect
33	1404	2994.305	knowledge	392	214	460.89	perception
34	1387	2060.834	staff	393	212	230.11	consequences
35	1384	1715.69	education	394	212	611.85	correlation
36	1382	1609.993	based *hospital-based, evidence-based	395	212	432.48	discourse

Rank	Freq	Keyness	Keywords	Rank	Freq	Keyness	Keywords
37	1372	1825.418	using	396	212	1380.4	inpatient
38	1338	772.635	group	397	211	205.06	improved
39	1332	1407.442	level	398	211	266.13	theme
40	1292	801.095	important	399	210	463.88	statistical
41	1292	294.311	life	400	209	408.65	excluded
42	1276	1219.448	members	401	208	1033.2	adherence
43	1267	1699.484	age	402	207	698.22	interactions
44	1256	5571.34	findings	403	206	642.48	engage
45	1239	452.568	found	404	205	226.53	involvement
46	1204	3437.045	families	405	204	1647.2	burnout
47	1202	303.634	need	406	204	451.29	everyday
48	1200	2597.629	treatment	407	203	233.39	identity
49	1194	5322.114	clinical	408	201	300.98	alcohol
50	1190	1304.372	services	409	201	596.45	screening
51	1169	4567.938	cancer	410	199	585.81	codes
52	1140	2096.207	hospital	411	198	593.52	explored
53	1130	1291.218	process	412	198	430.49	intensive
54	1118	327.44	different	413	197	323.21	frequency
55	1094	2835.052	factors	414	197	519.85	prevention
56	1091	1825.716	results	415	196	1669.6	subscale
57	1084	2172.522	included	416	194	293.3	evaluation
58	1077	324.906	high	417	193	428.43	therapy
59	1069	2139.128	significant	418	192	301.58	harm
60	1041	1569.071	quality	419	192	351.25	oral
61	1022	2394.875	medical	420	192	459.46	systematic
62	999	1818.681	relationship	421	190	224.39	awareness
63	993	1411.258	higher	422	190	1312	behavior
64	984	976.429	management	423	190	494.68	indicating
65	968	2640.939	pain	424	188	665.37	disorders
66	967	3786.383	experiences	425	188	322.01	improving
67	965	1089.642	needs	426	186	669.34	severity
68	950	2120.62	associated	427	186	1523.6	subscales
69	933	1036.068	individual	428	185	367.07	errors
70	915	2586.099	identified	429	185	235.96	managing
71	912	3053.136	sample	430	185	712.18	regression
72	902	1874.988	scale	431	184	249.4	components
73	900	2095.496	positive	432	184	840.51	discourses
74	895	1282.984	described	433	184	760.63	loneliness
75	895	416.266	service	434	184	254.17	statistics
76	890	3459.927	intervention	435	183	298.07	contribute
77	868	2334.473	items	436	182	461.85	populations
78	862	913.704	role	437	182	1093.7	socioeconomic
79	847	1379.4	levels	438	181	575.74	facilitate
80	831	7006.333	caregivers	439	180	890.83	cohort
81	831	3172.437	symptoms	440	180	699.81	ongoing

Rank	Freq	Keyness	Keywords	Rank	Freq	Keyness	Keywords
82	795	722.911	team	441	180	580.64	researcher
83	794	465.229	person	442	180	526.31	restraint
84	793	1662.244	older	443	180	869.66	vaccine
85	786	2083.571	relationships	444	178	243.68	reducing
86	782	1275.304	professional	445	176	203.78	delivery
87	782	801.966	provided	446	176	366.73	topics
88	779	3577.531	scores	447	175	601.87	experiencing
89	769	278.103	mean	448	175	750.64	statistically
90	765	500.801	provide	449	174	214.31	comfort
91	764	624.705	groups	450	173	265.28	calculated
92	762	5013.064	interventions	451	173	408.61	recruitment
93	760	3516.969	diagnosis	452	172	247.98	addressed
94	759	3161.132	professionals	453	172	335.62	associations
95	756	970.714	questions	454	172	572.61	conversations
96	748	249.32	control	455	172	355.95	emotions
97	744	1079.295	previous	456	172	631.93	perspectives
98	743	491.069	evidence	457	172	566.55	recruited
99	727	1842.77	focus	458	172	424.1	reviewed
100	725	932.45	population	459	171	198.47	category
101	721	2662.786	conducted	460	171	428.07	enhance
102	721	1023.03	specific	461	171	411.17	moderate
103	711	2648.781	satisfaction	462	171	281.2	potentially
104	705	4453.332	medication	463	170	319.37	sessions
105	703	1574.868	assessment	464	170	229.15	variable
106	703	258.09	problems	465	168	481.87	evaluate
107	700	2060.448	interview	466	168	871.74	sectional
108	699	1020.978	effects	467	168	557.66	treatments
109	698	1206.136	understanding	468	167	353.77	analysed
110	695	3779.345	outcomes	469	165	606.11	dementia
111	689	1155.691	review	470	165	445.69	documentation
112	688	438.735	training	471	165	383.51	motivation
113	685	2739.449	interviews	472	165	427.7	positively
114	685	1150.602	status	473	165	379.7	structured
115	676	257.615	present	474	165	208.71	visits
116	662	2578.482	beliefs	475	164	212.89	articles
117	655	2522.972	perceived	476	164	534.77	ethics
118	653	718.217	model	477	164	501.26	participating
119	648	548.88	approach	478	163	201.12	helpful
120	646	1198.734	individuals	479	163	878.25	parenting
121	641	405.455	changes	480	163	201.52	theoretical
122	639	1867.505	parent	481	161	348.75	disorder
123	637	1077.084	context	482	159	310.74	duration
124	633	252.855	community	483	159	435.09	feedback
125	628	745.704	issues	484	159	200.68	resulted
126	625	1210.835	impact	485	158	301.36	networks

Rank	Freq	Keyness	Keywords	Rank	Freq	Keyness	Keywords
127	621	907.798	lack	486	157	526.68	decreased
128	610	642.687	increased	487	157	215.93	sharing
129	610	790.861	potential	488	157	527.03	supportive
130	603	1804.548	characteristics	489	156	577.51	ranged
131	603	2317.803	depression	490	156	271.99	trials
132	601	2402.093	caring	491	155	346.74	competence
133	601	1555.907	negative	492	155	643.6	domains
134	599	968.154	compared	493	155	636.62	participated
135	593	375.711	personal	494	155	550.78	relational
136	592	1376.051	experienced	495	154	359.18	neighbourhood
137	591	904.655	disease	496	154	572.68	routines
138	591	383.61	total	497	153	327.75	identifying
139	590	1023.365	survey	498	152	454.13	consistency
140	588	1681.401	score	499	152	451.09	organizational
141	585	552.748	considered	500	152	479.83	sampling
142	584	249.734	job	501	152	530.6	wards
143	581	1390.983	literature	502	151	437.09	literacy
144	577	834.078	primary	503	151	640.87	provider
145	577	926.806	safety	504	151	244.01	uncertainty
146	576	1041.945	daily	505	150	407.06	lifestyle
147	576	2805.555	questionnaire	506	149	228.42	admission
148	564	2144.325	variables	507	147	329.86	assessing
149	562	952.327	differences	508	147	343.24	bias
150	561	1907.204	strategies	509	147	367.29	challenging
151	559	1644.327	emotional	510	147	540.55	online
152	557	750.32	importance	511	147	717.05	outpatient
153	552	230.44	low	512	146	480.74	evaluated
154	539	405.579	current	513	146	198.38	recommendations
155	539	540.138	response	514	146	752.3	vaccination
156	538	257.835	similar	515	145	378.14	highlighted
157	536	491.626	developed	516	145	464.14	placement
158	534	3031.267	participant	517	142	236.38	functional
159	530	2383.8	parental	518	141	200.54	indicates
160	525	1029.975	communication	519	141	377.48	rated
161	523	373.635	students	520	140	261.5	narrative
162	516	2523.812	settings	521	140	216.13	viewed
163	514	545.798	unit	522	139	282.2	promoting
164	512	1514.481	mothers	523	138	343.36	adverse
165	508	284.533	needed	524	137	715.02	affective
166	500	1180.285	practices	525	137	652.4	bullying
167	498	1537.582	smoking	526	137	828.46	paediatric
168	495	1614.892	anxiety	527	136	465.59	massage
169	495	1556.463	psychological	528	135	723.7	depressive
170	494	2011.732	themes	529	135	377.4	limitation
171	489	448.279	activities	530	134	658.52	empathy

Rank	Freq	Keyness	Keywords	Rank	Freq	Keyness	Keywords
172	489	367.274	environment	531	132	218.13	centred
173	488	2295.174	psychiatric	532	132	309.79	decrease
174	486	601.546	skills	533	132	263.17	disability
175	484	359.323	received	534	132	372.65	documented
176	483	862.708	factor	535	132	229.84	programs
177	482	246.197	living	536	130	474.54	engaging
178	481	2673.069	qualitative	537	130	618.68	validated
179	476	1528.13	responses	538	129	1020.5	behavioral
180	473	469.441	limited	539	129	243.3	oriented
181	472	567.885	ability	540	129	879.45	predictors
182	470	1027.472	stress	541	128	266.89	dimensions
183	464	233.461	greater	542	128	275.48	eligible
184	463	388.631	activity	543	128	244.7	scales
185	462	1783.659	chronic	544	127	389.2	clinics
186	459	1159.071	recovery	545	127	266.93	collaboration
187	456	1676.962	focused	546	127	246.17	descriptions
188	454	964.455	attitudes	547	127	215.56	implemented
189	454	712.259	sleep	548	127	627.31	negatively
190	453	1035.952	indicated	549	125	357.76	complications
191	451	1845.208	distress	550	123	304.7	addressing
192	446	784.609	educational	551	123	251.99	construct
193	445	604.232	overall	552	123	396.71	sensory
194	445	418.948	showed	553	121	219.34	couples
195	442	2890.917	healthcare	554	121	504.12	morbidity
196	442	2002.365	respondents	555	121	717.45	predictor
197	441	631.286	providing	556	120	455.13	coding
198	439	231.355	understand	557	120	304.11	meaningful
199	437	396.235	effective	558	120	431.22	physician
200	437	1643.517	participate	559	120	311.81	quantitative
201	436	486.596	culture	560	119	576.54	rounding
202	435	871.664	identify	561	119	482	schizophrenia
203	431	1899.228	perceptions	562	118	239.58	diverse
204	427	667.906	measure	563	118	469.29	impairment
205	424	729.359	critical	564	118	267.09	indicators
206	419	251.531	lower	565	117	221.22	adjustment
207	418	827.89	leadership	566	117	317.83	focusing
208	416	1010.586	item	567	117	226.73	independently
209	416	2057.116	providers	568	117	244.13	rounds
210	415	343.995	resources	569	117	574.18	spirituality
211	413	339.242	follow	570	116	260.51	shifts
212	413	338.455	influence	571	116	582.94	transcripts
213	412	570.569	measures	572	115	208.55	inform
214	402	273.021	appropriate	573	115	409.65	respondent
215	402	1109.076	hospitals	574	115	631.72	transcribed
216	400	288.834	access	575	114	405.96	nutrition

Rank	Freq	Keyness	Keywords	Rank	Freq	Keyness	Keywords
217	400	617.912	improve	576	114	204.66	translation
218	400	336.847	trust	577	114	401.35	trauma
219	396	1116.442	participation	578	113	403.45	carer
220	391	585.603	content	579	113	220.7	consisted
221	390	618.423	managers	580	113	678.07	delirium
222	390	812.367	significantly	581	113	212.43	interviewed
223	388	1185.442	acute	582	112	596.04	trajectory
224	388	1128.21	researchers	583	111	202	insight
225	387	387.918	exercise	584	111	473.73	longitudinal
226	386	226.496	association	585	109	610.77	analyzed
227	386	1794.802	coping	586	109	365.94	grandparents
228	384	857.86	concerns	587	109	411.05	segregation
229	384	343.691	learning	588	109	435.84	thematic
230	381	924.895	categories	589	109	226.28	translated
231	381	1013.943	implementation	590	108	503.85	cardiovascular
232	381	695.628	prior	591	108	386.99	workload
233	380	1345.851	suicide	592	107	472.79	bereaved
234	377	1783.621	workplace	593	107	409.02	coefficient
235	373	1988.308	behaviours	594	107	374.83	collaborative
236	373	773.124	informed	595	107	424.07	correlations
237	368	1847.589	diabetes	596	106	239.16	administered
238	364	2200.943	adolescents	597	104	218.31	strengths
239	364	522.569	completed	598	104	881.29	thalassemia
240	362	1922.764	adolescent	599	103	371.73	cardiac
241	362	416.681	aspects	600	103	648.88	clinician
242	361	1417.987	cognitive	601	103	435.26	interpersonal
243	359	1734.5	diagnosed	602	102	363.7	coded
244	359	259.891	majority	603	102	281.15	deterioration
245	359	1004.151	multiple	604	102	273.16	staffing
246	358	813.732	residents	605	101	254.7	marital
247	355	295.75	methods	606	101	347.6	simulation
248	354	1663.237	reliability	607	100	517.12	narratives
249	352	380.348	additional	608	97	422.6	impacts
250	352	363.528	setting	609	97	315.44	respiratory
251	350	2885.505	pediatric	610	96	203.91	behavioural
252	348	808.024	adults	611	96	415.43	facilitating
253	348	702.255	criteria	612	96	228.4	implementing
254	348	320.462	version	613	96	274.62	smokers
255	343	349.128	decisions	614	94	305.18	correlated
256	342	280.204	suggest	615	94	307.14	illnesses
257	340	292.951	condition	616	94	401.79	stigma
258	339	425.09	obtained	617	93	638.99	hospitalization
259	338	528.032	adult	618	93	261.33	physiological
260	337	276.654	develop	619	92	238.57	disagree
261	337	1454.865	functioning	620	92	264.42	dynamics

Rank	Freq	Keyness	Keywords	Rank	Freq	Keyness	Keywords
262	337	369.389	sexual	621	92	692.02	triage
263	334	2488.992	palliative	622	91	200.5	cluster
264	331	238.113	responsibility	623	90	249.64	developmental
265	330	681.909	consent	624	90	200.21	perceive
266	329	258.978	sex	625	89	236.11	induction
267	326	725.38	measured	626	89	418.22	multidisciplinary
268	325	1238.207	challenges	627	89	503.97	postoperative
269	325	642.204	manage	628	87	373.24	cessation
270	324	364.653	finding	629	87	199.17	onset
271	317	884.124	explore	630	86	228.78	focuses
272	317	279.153	values	631	86	538.52	preoperative
273	316	249.309	female	632	86	496.66	tattoos
274	314	1425.2	demographic	633	85	296.91	grounded
275	314	305.929	expressed	634	84	400.11	diagnoses
276	314	1267.543	organisational	635	84	399.35	empowerment
277	314	283.786	relation	636	84	676.99	neuroticism
278	313	1187.291	carers	637	84	200.66	problematic
279	313	252.443	relevant	638	84	357.84	spouses
280	312	2677.673	caregiver	639	83	701.51	antipsychotic
281	311	299.215	units	640	82	312.43	coefficients
282	310	960.451	limitations	641	82	227.67	medicines
283	310	229.716	meaning	642	81	264.73	contextual
284	310	447.976	observed	643	81	202.16	disabilities
285	309	1578.611	efficacy	644	81	248.06	exhaustion
286	308	372.045	crisis	645	81	354.9	teamwork
287	308	320.271	initial	646	80	362.32	de-escalation
288	307	1119.553	analyses	647	79	503.44	mentors
289	306	645.762	expectations	648	79	360.22	midwives
290	306	477.911	severe	649	79	351	standardized
291	305	616.94	collected	650	79	392.05	tactile
292	304	734.124	assess	651	79	213.93	undergoing
293	304	1014.056	discharge	652	78	217.91	influencing
294	303	660.823	consistent	653	78	266.07	tertiary
295	303	814.453	tool	654	77	212.24	interviewer
296	302	725.225	instrument	655	77	275.5	stressful
297	301	377.443	violence	656	76	508.18	family-centered
298	299	230.112	collection	657	75	283.08	discursive
299	299	288.963	users	658	74	221.33	pathway
300	297	399.395	feelings	659	74	367.99	subgroups
301	297	782.548	mortality	660	73	233.41	investigator
302	294	1502.113	fatigue	661	73	524.29	tumor
303	294	266.024	reduce	662	72	206.74	cohesion
304	294	1030.049	validity	663	72	310.82	geriatric
305	293	267.789	discussed	664	71	549.07	mentoring
306	291	985.547	barriers	665	71	383.16	verbatim

Rank	Freq	Keyness	Keywords	Rank	Freq	Keyness	Keywords
307	291	505.947	performed	666	71	299.7	yoga
308	290	2475.547	caregiving	667	70	226.99	uptake
309	289	1247.451	surgical	668	69	221.43	culturally
310	287	591.621	tools	669	69	494.97	stressors
311	286	249.788	cultural	670	68	256.39	chemotherapy
312	285	1788.88	clinicians	671	68	252.51	constructs
313	285	350.55	models	672	67	254.92	educators
314	285	665.709	reporting	673	67	219.99	enrolled
315	284	343.705	aged	674	67	538.83	extraversion
316	284	358.433	procedures	675	67	360.01	utilization
317	283	227.484	confidence	676	66	566.43	homebirth
318	283	759.077	gender	677	66	219.7	initiating
319	283	502.482	outcome	678	66	548.44	mentorship
320	281	474.356	emergency	679	66	249.46	prognosis
321	281	347.734	stated	680	66	407.57	psychometric
322	280	481.452	examine	681	66	315.68	seclusion
323	280	499.772	examined	682	66	409.21	somatic
324	280	281.89	noted	683	66	228.17	systemic
325	279	573.484	perspective	684	65	220.55	facilitated
326	278	263.333	aim	685	64	351.35	biomedical
327	278	653.291	authors	686	64	513.53	dyads
328	277	696.204	consumers	687	64	234.82	holistic
329	275	442.185	indicate	688	63	220.11	familial
330	274	626.337	surgery	689	63	486.68	interprofessional
331	273	236.398	address	690	63	204.45	rationing
332	273	593.831	roles	691	63	482.64	sociodemographic
333	273	431.267	shift	692	62	199.85	clinically
334	272	234.009	administration	693	62	290.43	subgroup
335	270	347.915	affect	694	60	213.26	mentor
336	270	227.477	difficulties	695	59	363.68	facilitators
337	270	228.07	internal	696	59	265.62	influenza
338	269	964.57	inclusion	697	59	476.67	tumors
339	267	436.739	situations	698	57	254.09	utilized
340	266	1251.305	symptom	699	56	212.16	psychiatry
341	266	481.174	ward	700	56	339.9	randomized
342	264	648.788	assessed	701	55	286.77	competencies
343	263	285.016	processes	702	55	238.32	neonatal
344	263	1146.571	survivors	703	55	204.13	obesity
345	260	623.65	guidelines	704	55	198.77	suicidal
346	259	1992.614	oncology	705	54	463.44	dyspnea
347	256	2155.842	behaviors	706	54	267.5	internet
348	256	279.735	birth	707	53	220.27	accessing
349	254	833.917	aggression	708	53	364.7	dyadic
350	252	381.849	determine	709	52	331.79	centredness
351	251	536.933	burden	710	52	446.28	hemodialysis

Rank	Freq	Keyness	Keywords	Rank	Freq	Keyness	Keywords
352	250	781.789	clinic	711	52	288.48	impacted
353	249	316.41	strongly	712	52	324.75	inhaler
354	247	1763.773	psychosocial	713	52	226.46	logistic
355	246	217.287	affected	714	52	388.23	test-retest
356	245	271.236	shared	715	51	437.7	telehealth
357	244	1053.708	descriptive	716	50	296.03	hospitalized
358	243	814.356	engagement	717	50	197.98	understandings
359	243	1358.745	siblings				

Appendix G

Lexical Collocations with Combination Types in accordance with the Set Framework

L1 Verb + Noun nodes=75, 128 pairs

No.	Nodes	Collocates	No.	Nodes	Collocates
1	reported	feeling	30	explore	(x,xx) experiences
		(x,xx) advice			(x) relationship
2	use	lubricants			(x,xx) factors
3	need	assistance	31	expressed	(x,xx) concern/s
4	included	(x) gender			(x,xx) interest
		(xx) review			(x) need
		(x) following	32	observed	(x,xx) gradient
5	associated	(x,xx) suicide	33	collected	(x) data
6	identified	(x,xx) key			(x,xx) using
		(x,xx) themes	34	assess	(x) association
7	described	(x,xx) detail	35	reduce	(x) risk
		feeling			(xx) medication
8	provided	(x,xx) insight	36	performed	(x,xx) using
		(x,xx) opportunity	37	examine	(x,xx) differences
		(x,xx) consent			(x,xx) relationship/s
		(x,xx) information	38	examined	(x) relationship
9	provide	(x,xx) insight/s			using
		(x) opportunities	39	address	(x,xx) concerns
		(x,xx) evidence			(x,xx) needs
		(x,xx) support			(x,xx) issues
10	conducted	(x) examine	40	affect	(x) quality
		(x,xx) interviews	41	assessed	using
		(x) using	42	determine	(x) expectations
11	compared	(x,xx) women	43	affected	(x) results
12	experienced	(x,xx) stress	44	describe	(x,xx) experiences
13	needed	(x) help			(x,xx) characteristics
14	received	(x) approval	45	influenced	(x,xx) factors
		(x,xx) attention	46	receiving	(x,xx) treatment
		(x,xx) training			(x,xx) diagnosis
		(x,xx) education	47	selected	(x,xx) hospitals
		(x,xx) treatment	48	promote	(x) change
15	living	(x) chronic	49	improved	(x,xx) quality
		(x,xx) areas			(x) knowledge
		(xx) illness	50	excluded	(xx) sample
		(x,xx) cancer	51	engage	(x,xx) activities
15	understand	(x,xx) experiences	52	explored	(x,xx) experiences
17	participate	(xx) study	53	indicating	(x,xx) levels
18	identify	(x) areas	54	contribute	(x,xx) development
19	measure	(x,xx) beliefs	55	facilitate	(x) uptake

No.	Nodes	Collocates	No.	Nodes	Collocates
20	access	(x,xx) resources	56	experiencing	(x,xx) illness
		(x,xx) services	57	calculated	(x) using
21	improve	(x,xx) quality	58	enhance	(x,xx) understanding
		(x) understanding	59	evaluate	(x,xx) interventions
		(x,xx) outcomes	60	analys/zed	using
22	informed	consent	61	ranged	(x) years
		(xx) decisions			(x) age
23	completed	(x,xx) survey	62	participated	(xx) study
		(x,xx) questionnaire/s			(x,xx) research
24	diagnosed	(x) schizophrenia	63	evaluated	using
		(x,xx) cancer	64	highlighted	(x,xx) importance
		(x,xx) diabetes			(x) need
		(xx) patients	65	rated	(xx) point
		families	66	decrease	(x,xx) stress
25	suggest	(x) stigma	67	inform	(x,xx) development
26	obtained	(x) permission			(x,xx) practice
		(x,xx) consent	68	consisted	(xx) items
27	develop	(x,xx) strategies	69	translated	(x,xx) English
		(x,xx) interventions	70	bereaved	(x) suicide
		(x,xx) understanding	71	coded	(x,xx) transcripts
28	measured	using	72	undergoing	(x) treatment
29	manage	(x,xx) condition	73	enrolled	(xx) study
		(x) body	74	facilitated	(x,xx) sessions
		(x) child	75	impacted	(x) results
		(x,xx) health			
		(x,xx) care			

L2 Adjective + Noun

No.	Nodes	Collocates	No.	Nodes	Collocates
1	mental	(x) triage	39	parental	rearing
		(x) crisis/es			presence
		(x) commission			absence
		(x,xx) health			(x) behavior/s
		illness/es			vaccine
		(x) facilities			reports
		(x) service/s			self-efficacy
		(x,xx) workforce			(x,xx) attitudes
		(x) disorder/s			di/stress
		(x) practitioners			(x,xx) satisfaction
		(x) practitioners			role
		(x) teams	40	psychological	contract
		(x) simulation			(x) fulfillment
		(x) consumers			distress
		(x) problems			parameters
		(x) issues			trauma
		state			(xx) impact
		(x) settings			well-being
		(x) professionals	41	psychiatric	inpatient
		(x) clinicians			wards
2	self-reported	(x,xx) adherence			disorders
3	nurse-led	clinic			hospitalization

No	Word	Collocates	No	Word	Collocates
4	each	(x) transcript			(x) units
		(x) item			hospitals
		(x,xx) dimension			(x) association
		(x,xx) site			symptoms
		(x) country			(x,xx) settings
		(x,xx) indicator			nurses
		subscale	42	qualitative	(x,xx) design
		(x) component			(x) methods
		participant			(x) content
		(x) category			(x) research
		session			(x) analysis
		domain			(x) interview/s
		(x) theme			(x) study/ies
		(x) member			(x) data
		(x) source			findings
		(x) variable			(x) approach
		(x) interview	43	limited	(x,xx) resources
		(x,xx) question			(x) evidence
		year			number
5	social	cohesion	44	greater	dyspnea
		gradient			(x,xx) depression
		science/s			understanding
		capital			(x,xx) satisfaction
		isolation	45	chronic	(x,xx) illness/es
		network/s			(x,xx) disease/s
		support			(x) condition/s
		norms	46	customer-oriented	behaviour/s
		worker/s	47	educational	attainment
		class			(x) program/s/me/mes
		relations			(xx) intervention/s
		(xx) contexts			(x) content
		media			(xx) level/s
		(x) theory	48	overall	(x) satisfaction
		activities			(x,xx) rate
		interaction/s			(x,xx) score
		functioning	49	effective	(x,xx) reducing
6	physical	activity			(x,xx) improving
		fitness			(x,xx) communication
		assault			(x,xx) strategies
		disorder			(x,xx) interventions
		performance			(x,xx) management
		restraint	50	critical	(x) setting
		illnesses			care
		functioning			role
		(x) tests			(x) analysis
		abuse	51	lower	(x,xx) level/s
		function			(x) income
		(x,xx) health			(x,xx) scores
		(x,xx) problems			(x,xx) rates
		aggression			(x) status

		difficulties	52	appropriate	(x) interventions
		condition/s			(x,xx) treatment
		(x,xx) violence	53	acute	(x) leukemia
7	hospital-based	oncology			(x,xx) inpatient
8	evidence-based	practice/s			(x,xx) setting/s
		interventions			(x,xx) units
9	important	contribution			(x,xx) hospital/s
		predictor			(x,xx) community
		aspect			(x,xx) care
		component			(x,xx) illness
		implications	54	prior	(x) death
		(x,xx) issue	55	completed	questionnaire/s
		role	56	cognitive	impairment
		(x) factor			(xx) function
		(x) step			(xx) functioning
10	clinical	pathway/s			theory
		excellence			(x,xx) fatigue
		deterioration			depression
		judgement			interview
		(x) specialists	57	multiple	(x,xx) comparisons
		supervision			(x) regression
		trial/s			(x,xx) analysis
		setting/s	58	additional	(x) topics
		(x) practice			(x,xx) training
		populations			(x) information
		(x) guidelines	59	pediatric	oncology
		characteristics			cancer
11	different	types			(x) unit
		styles			(x) patients
		ways			(xx) services
		approaches	60	adult	(xx) caregivers
		levels			(x,xx) providers
		(x) points			(xx) population
		countries			(x) health
		(x) settings	61	sexual	intercourse
		roles			intimacy
		aspects			orientation
12	included	trials			abuse
13	high	secure			partners
		(x,xx) neuroticism			communication
		(x) turnover			activity
		(x,xx) extraversion			practices
		school			risk
		profile	62	palliative	(x) provision
		level/s			(xx) care
		workload			(x) needs
		priority			(x) team
		(x,xx) rate/s			(x) services
		prevalence	63	female	(x) worker
		(x) demands			(x,xx) years
		degree	64	demographic	(xx) characteristics

(x,xx) burnout
(x) score/s
reliability

14 high-risk behaviors

15 significant difference/s
predictor/s
correlation/s
improvement/s
reduction
association/s
effect/s
relationship/s
impact
others
increase

16 medical comorbidity
records
center
history
condition/s
discourse
unit
tasks
association
(x) staff

17 higher (x,xx) score/s
prevalence
level/s
proportion
(x,xx) mobility
(x) degree
(x) rate/s
(x,xx) income
(x) efficacy
(x,xx) suicide
(x) education
deprivation

18 individual (x) interviews
(x,xx) characteristics

19 positive emotions
correlation
risk-taking
attitude/s
outcomes
feedback
aspects
effect/s
association
experiences
impact
relationship/s

(xx) variables
(x) questionnaire
differences
(xx) information
(xx) data
(xx) factors
(x,xx) age

65 organisational injustice
justice
culture
commitment
performance
levels

66 relevant (x) literature
67 initial coding
codes
(x,xx) interview
(xx) diagnosis
(x) assessment

68 severe (xx) pain
(x) symptoms
(x) illness

69 consistent (x,xx) literature
(x,xx) findings
(x,xx) studies
(x,xx) results

70 surgical (x,xx) wards
procedure/s
(xx) unit
patients
intervention

71 cultural (xx) capital
(xx) values
(xx) differences
(xx) contact
(x,xx) practices

72 aged (x) years
73 internal consistency
(x) reliability
(x) models
influence
(xx) trust
working

74 psychosocial climate
aspects
(x,xx) needs
(x,xx) support
(x,xx) factors

75 shared decision-making
76 descriptive statistics
(x) design

20	older	adult/s
		people
		women
		(x,xx) member
		residents
		(x) discharge
		person
		age
21	professional	identity
		(x,xx) accountability
		standards
		attitude
		autonomy
		development
		codes
		role
		(x,xx) culture
		(x) environment
		(xx) groups
		practice
22	mean	score/s
		(x) duration
		age
		(xx) years
		(x,xx) item
23	previous	studies
		research
		(x) months
		researchers
		findings
24	specific	aspects
		communication
		(x,xx) content
		(x,xx) interventions
		(x) strategies
		(x) needs
		(x) context
		(x) questions
25	present	study
26	perceived	(x,xx) barriers
		(xx) benefits
		(xx) confidence
		(x,xx) status
		(x,xx) stress
		(x,xx) support
27	increased	risk
		(x,xx) rates
		(x,xx) confidence
		(x) activity
28	potential	confounders/ing
		(x,xx) bias

77	ethical	analysis
		approval
		committee
		(x,xx) principles
		(x) board
		review
78	registered	nurse/s
79	healthy	lifestyle
		diet
80	institutional	(x) board/s
		review
81	selected	(x) hospitals
82	core	category
		beliefs
		(x,xx) illness
83	therapeutic	conversation/s
		relationship/s
		(xx) risk
84	statistical	significance
		analysis/es
		power
85	everyday	life/ves
		(x) practice
		work
86	intensive	(x) unit/s
		(x,xx) care
		(xx) hygiene
87	oral	(x) status
		care
88	systematic	review/s
89	socioeconomic	inequalities
		status
		(xx) factors
90	ongoing	(x,xx) process
91	moderate	level/s
92	cross-sectional	(x,xx) design
		(x) survey
		(x,xx) study/ies
93	semi/structured	interview/s
94	participating	countries
95	helpful	(x,xx) services
96	theoretical	framework
97	decreased	(xx) quality
98	supportive	(xx) relationships
99	relational	closeness
		issues
100	organizational	(xx) policies
		culture
		(xx) factors
		(xx) support
101	challenging	situations

Left column

(x,xx) benefits
impact
(x,xx) effect/s

29 negative
(x) emotions
consequences
(x,xx) impact/s
(x,xx) attitudes
(xx) correlation
(x) effects
(x) outcomes
(x,xx) coping
association
aspects
(x) perceptions
(x,xx) experiences

30 personal
accomplishment
growth
contact
(x,xx) characteristics
responsibility
(xx) experience/s
(xx)development

31 total
(x,xx) score/s
variance
number
(x) range
(x) scale
(x) sample
(x,xx) population

32 primary
(x) caregiver/s
(x,xx) prevention
aim
(xx) outcome
(x,xx) services
(x,xx) settings
care
family
(xx) caring

33 daily
(xx) bundle
(x) hemodialysis
maintenance
(xx) shower
bath
(x) operation
life/ves
living
routines
basis
unit
(x,xx) activity/ies
step-count/s

Right column

102	functional	(x) decline
		limitations
		(x,xx) status
103	narrative	(x) notes
104	adverse	(x) events
		(x) outcomes
		(x) effects
105	affective	commitment
		(xx) depression
		(xx) staff
106	paediatric	nurses
107	depressive	symptoms
108	person-centred	care
109	documented	(x) symptoms
110	validated	(x) tools
		(x) assessment
111	behavioral	beliefs
112	eligible	participants
		patients
113	sensory	room/s
114	meaningful	coping
		work
115	quantitative	studies
		(x,xx) data
116	diverse	(x,xx) settings
117	longitudinal	(x) study/ies
		(x) data
118	thematic	(x) analysis
119	bereaved	siblings
120	collaborative	(x,xx) relationships
121	self-administered	(x,xx) questionnaire
122	cardiac	(x,xx) program
123	interpersonal	dynamics
		relationships
124	marital	status
125	respiratory	rate
126	behavioural	problems
127	physiological	parameters
128	developmental	disabilities
129	multidisciplinary	(x) team
		(x,xx) approach
130	postoperative	pain
		(x) management
131	preoperative	anxiety
		experiences
		(x,xx) patients
132	grounded	theory
133	problematic	(x) behaviours
134	antipsychotic	polypharmacy
		(x) side-effect/s
		medication

			No.	Nodes	Collocates
		body	135	contextual	(x) factors
		(x) condition	136	standardized	assessment
		home	137	tactile	massage
		(xx) management	138	influencing	factors
34	emotional	exhaustion	139	tertiary	(x,xx) hospital
		(xx) depersonalisation	140	stressful	(x,xx) events
		warmth	141	family-centered	care
		reactions	142	discursive	practice/s
		distress	143	geriatric	team
		response/s	144	verbatim	transcripts
		support	145	psychometric	properties
		(x,xx) functioning	146	somatic	(xx) depression
		(x,xx) problems	147	systematic	perspective
35	low	(x,xx) extraversion	148	biomedical	discourse
		(x,xx) neuroticism			(x) care
		fall	149	holistic	(x,xx) care
		income	150	familial	(x) boundaries
		(x) literacy			(x) expectations
		level/s			(xx) care
		(x) efficacy	151	interprofessional	trust
		(x) scores	152	sociodemographic	(xx) characteristics
		risk	153	randomized	(x,xx) trial/s
36	current	literature	154	neonatal	nurses
		study			(x) care
		(x,xx) evidence	155	suicidal	ideation
37	similar	(x,xx) findings	156	dyadic	(xx) interview/s
		(x,xx) results	157	logistic	(xx) regression
38	developed	(x) countries			(x) models
					(x) analysis
			158	hospitalized	(x,xx) patients

L3 Noun + Verb nodes=117, 240 pairs

No.	Nodes	Collocates	No.	Nodes	Collocates
1	study	(x) approved	28	interventions	(x,xx) designed
		aims/ed			(x,xx) offered
		adds			(x,xx) reduce
		(x) granted			(x,xx) improve
		highlights	29	questions	(x,xx) asked
		contributes	30	evidence	suggests
		(xx) investigate	31	interview	(x,xx) conducted
		(x) conducted	32	interviews	(x) audio-recorded
		(xx) explore			lasted
		investigated			(x) transcribed
		(x) consisted			(x) conducted
		sought			took
2	nurses	working	33	beliefs	(x,xx) held
		play	35	approach	(x) involves
3	patients	(x) hospitalised	36	changes	(x,xx) made
		(x) admitted	37	issues	(x) raised
		(x) undergoing			related
		(x) discharged	38	survey	(x,xx) completed

4	participants	(x) recruited			(x,xx) conducted
		(x) asked	39	literature	(x) suggests
		responded	40	questionnaire	(x,xx) developed
		spoke			(x) included
		(x) informed	41	variables	(x) included/ing
		mentioned	42	differences	(x,xx) found
		believed	43	strategies	(x,xx) promote
		gave			(x,xx) reduce
		completed			(x,xx) improve
		(x) invited	44	participant	said
		(x) interviewed			stated
5	data	(x) collected			described
		(x) gathered	45	themes	(x) emerged
		(x) analysed/zed			(x) identified
		(x) occurred	46	skills	required
		(x,xx) obtained			(x,xx) needed
6	research	exploring	47	ability	(x) engage
		(x) needed			(x,xx) manage
		suggests			make
		indicates			provide
		(x,xx) guided	48	respondents	(x) asked
		(x,xx) demonstrated			(x) indicated
		(x,xx) reviewed			(x) reported
7	children	undergoing			working
		(x,xx) allocated	49	item	(x) rated
		(x,xx) diagnosed			(xx) measure
8	nurse	stated	50	resources	(x) needed
9	parents	(x) expressed	51	concerns	(x) expressed
		(x) agreed	52	methods	(x,xx) used
10	studies	investigating	53	criteria	(x,xx) included
		(x,xx) examined	54	decisions	(x,xx) made
		examining	55	consent	(x) obtained
		(x) shown			(x) participate
		(x) investigated	56	challenges	(x) faced
		(x,xx) explored	57	finding	(x,xx) supported
		exploring	58	carers	(x,xx) described
		focusing	59	caregiver	stated
		involving	60	analyses	(x) performed
		(x,xx) conducted			(x) conducted
		show			(xx) using
		(x) published	61	tool	(x,xx) developed
		(x,xx) evaluated	62	instrument	(x,xx) measure
		reporting	63	collection	(x,xx) included
		(x) focused	64	tools	(x,xx) used
11	information	(x,xx) provided	65	procedures	(x) performed
		(x,xx) collected	66	authors	argue
12	being	assaulted	67	surgery	(x) performed
		treated	68	difficulties	(x,xx) experienced
		held			(x,xx) related
		diagnosed	69	situations	(x,xx) require
13	analysis	(x,xx) restricted	70	processes	involved

		(x) performed
		revealed
		showed
		(x) conducted
		involved
14	education	(xx) diagnosed
15	findings	suggest
		indicate
		highlight
		show/ed
		revealed
		presented
16	families	experiencing
		(x,xx) live
		living
17	need	(x) consider
		(x) develop
		(x) explore
		(x) understand
18	services	(x,xx) provided
19	process	(x,xx) requires
20	factors	influencing
		affecting
		(x,xx) affect
		(x) influence/d
		(x) contribute
		(x,xx) associated
		include/ing
21	results	show/ed
		highlight
		suggest
		indicate
		revealed
		(x) presented
		indicated
		(x) obtained
22	sample	(xx) consisted
23	scale	ranged/ing
		(x) developed
24	items	(x,xx) loaded
		(x,xx) scored
		(x,xx) measure
25	role	(x,xx) play
26	caregivers	(xx) living
		(x,xx) caring
27	score/s	indicate/ed/ing
		range/s/d
		(x,xx) calculated
		(x,xx) obtained

71	colleagues	found
72	risks	posed
		(xx) associated
73	actions	(x,xx) taken
74	questionnaires	(x,xx) returned
		(x,xx) completed
75	tasks	(x,xx) performed
76	framework	(x) developed
77	resources	(xx) needed
78	providers	(x) reconcile
79	program	(x) improve
80	theme	describes
81	errors	(x,xx) occur
81	regression	(x,xx) performed
		(x,xx) used
82	discourses	(x,xx) produce
83	statistics	(x) used
84	researcher	(x,xx) contacted
85	category	(x) abstracted
86	sessions	(x) conducted
87	motivation	(x) manage
88	articles	(x,xx) excluded
89	feedback	(xx) provided
90	trials	included
91	routines	(x,xx) used
92	sampling	(x,xx) used
93	bias	(xx) given
94	recommendations	(x) include
95	programs	(x) provide
96	scales	(x,xx) demonstrated
97	complications	associated
98	couples	(x,xx) reported
99	schizophrenia	living
100	spirituality	(x,xx) defined
101	transcripts	(x) analysed
102	segregation	(x) measured
103	coefficient/s	(x) calculated
104	correlations	(x) calculated
105	stigma	associated
106	tattoos	(x,xx) analyze
107	diagnoses	(x,xx) included
108	neuroticism	(x) predicted
109	spouses	(x) associated
110	medicines	(x,xx) prescribed
111	midwives	attending
112	interviewer	(x) see
113	pathway	(xx) undergoing
114	investigator	explained
115	educators	(x,xx) teach
116	competencies	(xx) needed
117	professionals	working

(x,xx) involved

L4 **Noun + Noun**   nodes=363, 889 pairs

No.	Nodes	Collocates	No.	Nodes	Collocates
1	care	bundle	100	smoking	ban/s
		provider/s			cessation
		recipient/s			(x,xx) alcohol
		(x) recommendation			(x) rates
		plans			behaviour
		provision	101	anxiety	(x,xx) depression
		coordination			(x,xx) fatigue
		units			(x,xx) sleep
		settings			(xx) levels
		facilities			(x,xx) symptoms
		delivery			(xx) scale
2	self-care	agency	102	themes	(x,xx) subthemes
3	health	literacy	103	activities	(xx) living
		professionals	104	environment	(x,xx) rationing
		(x) provider/s	105	skills	training
		promotion			(x) knowledge
		centres	106	factor	loadings
		check/s			structure
		professions			solution
		crisis/es			model
		commission			analysis
		service/s	107	living	(xx) edge
		insurance			arrangements
		facility/ies	108	stress	disorder
		complaints			(x,xx) anxiety
		(x,xx) recommendation			(xx) management
		sciences			(x) symptoms
		(x,xx) excellence	109	activity	levels
		workforce	110	recovery	(x,xx) practice/s
		behaviours			(x) approach
		(x,xx) care			process
		status	111	attitudes	(x,xx) smoking
		problems			(x,xx) beliefs
		system/s	112	sleep	disturbances
		outcomes			deprivation
		practitioners			(x,xx) problem/s
		issues			(x,xx) anxiety
4	study	protocol			(x,xx) fatigue
5	nurses	(x) assistant			(x,xx) quality
		(xx) assistants	113	distress	(x,xx) follow-up
		(x) physicians	114	providing	tactile
		(xx) commitment			(x) massage
6	patients	(x) spouses			(x,xx) information
		(x,xx) relatives			(x,xx) care
		(x) carers			(x) education
		(x) dementia			(x) support
		(x,xx) families	115	culture	(x,xx) justice

7	family	member/s
		normalcy
		harmony
		advocate
		(x,xx)friends
		caregiver/s
		functioning
		empowerment
		involvement
		systems
		strengths
		dynamics
		conversations
8	patient	acuity
		(x) advocate
		safety
		satisfaction
		education
		teaching
		(x) carer
		outcomes
		participation
		interactions
9	nursing	curricula
		home/s
		shortage
		(x) placement
		assistants
		profession
		pathway
		student/s
		discipline
		educators
		(x) environments
		practice
		workforce
		notes
		science
		teamwork
		staff
		rounds
		leaders
		(x) adjustment
		documentation
		competence
		interventions
10	data	collection
		set/s
		saturation
		sources
		analysis

116	perceptions	(x,xx) behaviors
		(x) caring
117	leadership	style/s
		behaviours
118	item	measure
		scale
		responses
		(x) questionnaire
119	follow-up	(x) interviews
120	influence	(x,xx) outcomes
121	measures	(x,xx) distress
		(x,xx) quality
122	hospitals	(x) clinics
123	access	viewing
		(x,xx) resources
		(x,xx) services
124	trust	(x) solidarity
125	participation	(x) empowerment
		rate
126	content	(xx) validity
		(x) index
		analysis
		areas
127	managers	(x,xx) leadership
128	exercise	participation
129	association	(xx) emotions
		(x,xx) school
		(x,xx) education
130	coping	strategy/ies
		(x,xx) use
131	learning	efficiency
		effectiveness
		(x) experience
		(x,xx) themes
132	categories	process
133	implementation	death/s
134	suicide	(x) self-harm
		prevention
		mortality
135	workplace	spirituality
		bullying
		violence
		culture
		conditions
136	diabetes	mellitus
		(x) classes
		(x) education
		self-management
137	adolescents	(x) leukemia
		(x) cancer
		(xx) years

302

No.	Node	Collocates	No.	Node	Collocates
		(xx) points	138	adolescent	psychiatry
11	research	fellows			healthcare
		ethics			(x) domains
		(x) committee			(x,xx) adult
		assistant/s			mothers
		design/s			(x,xx) development
		question/s			(x) issues
		team			(x)
		articles			parents
		project	139	aspects	(x,xx) caregiving
12	children	(x) disabilities	140	majority	(x,xx) respondents
		(x,xx) adolescents			(x,xx) participants
		(x,xx) thalassemia	141	residents	(xx) adjustment
		(x,xx) leukemia	142	reliability	(x,xx) validity
		(x,xx) cancer			coefficient
13	nurse	anaesthetists			testing
		staffing	143	adults	(x) years
		educator/s	144	version	(xx) instrument
		consultant			(xx) questionnaire
		manager/s	145	condition	management
		leader/s	146	functioning	(xx) life
		prescribing	147	responsibility	(x,xx) accountability
		administrators	148	consent	form/s
		practitioner/s	149	sex	differentials
		mentors			ratio/s
		call			(x,xx) partners
		(x,xx) physician			(x,xx) age
14	parents	(x,xx) children	150	values	(x,xx) beliefs
		(x) siblings	151	relation	(xx) assessment
15	support	(x) assistant			(x,xx) risk
		(x,xx) friends	152	carers	(x,xx) district
		systems			(x,xx) knowledge
		(x,xx) siblings	153	caregiver	(xx) burden
16	risk	(x,xx) ulceration	154	units	(xx) hospitals
		(x) infertility	155	limitation/s	(xx) study
		(x) stunting	156	meaning	units
		assessment	157	meaning-making	themes
		taking	158	self-efficacy	expectations
		(xx) defects			(x,xx) exercise
		(x,xx) reduction			(x,xx) outcome
		(x) tools			(x,xx) community
		(x,xx) safety			(x,xx) scale
		factor/s	159	crisis	response/s
		(x,xx) harm			(xx) situations
		(x,xx) developing			(x,xx) services
		(x,xx) bias			intervention
		management			(x) team
		behaviors	160	expectations	(x) recovery
		(x,xx) suicide			(x) outcome
17	information	sheet			(xx) expectations
		(x) instructions	161	discharge	(xx) planning

No.	Word	Collocations
18	work	session, intensification, environment/s, (x,xx) shifts, engagement, demands, hours, processes
19	practice	environment/s, guidelines, standards
20	child	rearing, (x,xx) psychiatry, (x) thalassemia, (x,xx) disabilities, protection, (x) adolescent, (x,xx) comfort, (x,xx) diagnosis, (x,xx) cancer, (x,xx) condition
21	use	(x) lubricants, (x) fertility, (x) seclusion, (x,xx) restraint, (x,xx) technology, (x,xx) treatments, (x,xx) tools
22	analysis	(x) variance
23	illness	blogs, suffering, belief/s, expressions, narratives, severity, (x) model, (x) non-illness
24	experience	(x) loneliness
25	knowledge	utilization, (x) ads, base, sharing, (x,xx) district, (x) skill/s, (xx) innovation, translation, (x,xx) attitudes, (x) adherence
26	staff	turnover, (xx) commitment, member/s

No.	Word	Collocations
		education, (xx) home, (x,xx) hospital, (x,xx) items
162	instrument	
163	violence	prevention, (x) aggression, period
164	collection	(x,xx) analysis, (x,xx) process
165	users	(x,xx) families
166	feelings	(x,xx) guilt, (x,xx) fear
167	mortality	rate/s
168	fatigue	(x,xx) depression
169	validity	(x) reliability, index
170	barriers	(x) facilitators, (x,xx) implementation
171	caregiving	responsibilities, role
172	clinicians	(xx) lack
173	models	(x) caregiving
174	reporting	(x) bias, interval
175	confidence	(x) comfort, (x,xx) ability
176	gender	(x) ethnicity, (x) male, (x) age, (x,xx) status, (x,xx) education, (x,xx) level
177	outcome	expectations, measure/s, variables
178	emergency	department/s, room, (xx) service/s
179	aim	(xx) study
180	consumers	(x) carers
181	roles	(x) responsibilities, workers
182	shift	(x,xx) medicines, errors
183	administration	(x,xx) medication
184	inclusion	(xx) criteria, (x) exclusion, (x,xx) education
185	symptom	occurrence, (xx) severity, cluster

No.	Word	Collocates
27	education	classes, (x,xx) occupation, programs/me/mes, (x,xx) income, (x) training
28	using	(x,xx) package, (x,xx) software, (x) statistics, (x,xx) tests, (x,xx) methods, (x) tools, (x) content
29	group	discussion/s, interview/s
30	settings	(x,xx) samples
31	level	(x) consciousness, (x) significance, (x) empathy, (x) satisfaction, (x) anxiety
32	life	expectancy, course, events, situations
33	members	(xx) couple
34	age	(x,xx) gender, (x,xx) ethnicity, (x) sex, (x,xx) years, (x,xx) cohort, group/s, range, (x,xx) status, (x) diagnosis
35	need	(x) assistance
36	treatment	regimen, options, prognosis, plan, intensity, adherence, decisions
37	cancer	trajectory, survivors, diagnosis, mortality, treatment, experience
38	hospital	charges, stay/s, admission/s

No.	Word	Collocates
		(xx) distress, management, (x) strategies
185	ward	managers, (x) staff, level
187	survivors	(x) childhood, (xx) cancer
188	guidelines	(x) policies
189	oncology	patients, group, nurse
190	birth	defects, weight, cohort, control
191	aggression	(x) violence
192	burden	(xx) caregivers
193	clinic	visit/s
194	engagement	(xx) users
195	siblings	(x,xx) children
196	childhood	cancer
197	transition	(x) parenthood, (xx) services, experience
198	baseline	characteristics, (x) data, age
199	approaches	(x,xx) body
200	receiving	(x,xx) diagnosis, (x,xx) information, (x) education, (x,xx) support, terms
201	interaction	(x,xx) interventions
202	effectiveness	(x) using
203	physicians	(x,xx) nurses
204	prevalence	(x,xx) pain
205	perception	(xx) efficacy
206	consequences	(x,xx) violence
207	correlation	coefficient/s
208	discourse	analysis
209	inpatient	ward, unit/s, setting/s, care
210	adherence	(xx) lifestyle
211	interactions	(x) others
212	involvement	(x,xx) discharge, (xx) education
213	burnout	(x,xx) job

		administrators			(xx) satisfaction
		discharge	214	identity	(x) definition
		costs	215	alcohol	consumption
		oncology			intake
		setting/s			(xx) drinking
		(x,xx) clinic			(xx) suicide
		(x,xx) unit			use
39	quality	assurance	216	screening	tool/s
		(x) life	217	codes	(x,xx) categories
		(x) sleep	218	frequency	(x) drinking
		improvement	219	prevention	strategies
		(x,xx) care	220	subscale	scores
		(x) safety			items
40	relationship	quality	221	evaluation	phase
41	management	strategies	222	therapy	center
		efforts	223	self-harm	(x) suicide
		plan	224	behavior	change
		practices	225	disorders	(x,xx) depression
42	pain	relief	226	improving	(x,xx) quality
		(x,xx) nausea	227	severity	(x,xx) illness
		(xx) movement	228	subscales	(x,xx) scores
		management	229	managing	(x,xx) deterioration
		prevalence			(x,xx) symptoms
		(x) anxiety			(x,xx) risk
		control	230	regression	model/s
		medication			(xx) analysis/es
43	experiences	(x,xx) feelings	231	components	(x) dimensions
44	needs	(x) concerns	232	loneliness	(x,xx) isolation
45	sample	size/s	234	cohort	effect/s
		(x) characteristics			(x,xx) period
		(x,xx) adolescents			studies
		(x,xx) adults	235	restraint	(x) asking
46	scale	(x,xx) items	236	vaccine	intention
		(xx) reliability			(x,xx) attitudes
47	service	user/s			(x) beliefs
		delivery	237	reducing	(xx) medication
		provision			(x) use
		providers	238	delivery	room
		system	239	topics	(x,xx) inclusion
48	intervention	(xx) communities	240	comfort	(xx) levels
		(xx) skill	241	recruitment	(x) retention
		(xx) sites			methods
		period			process
		(xx) group/s	242	associations	(xx) distress
		models	243	conversations	(x) families
		(x) control	244	perspectives	(xx) transition
		(x) training	245	variable	(xx) points
		area	246	treatments	(x,xx) procedures
49	role	(x) identity	247	dementia	damage
50	levels	(x,xx) distress	248	documentation	(x) delirium
		(x) anxiety	249	motivation	(xx) activity

		(x) burnout
		(x,xx) satisfaction
		(x) job
		(x,xx) stress
		(x,xx) depression
51	caregivers	(x,xx) adults
52	symptoms	(x,xx) depression
		(x) anxiety
53	team	meetings
		leader
		members
		(x,xx) communication
54	person	centredness
		(x) diabetes
		(x,xx) illness
55	researchers	(x) clinicians
56	behaviours	(x) managers
57	behaviors	(x) perceptions
58	diagnosis	(x,xx) prognosis
		(x) schizophrenia
		(x,xx) treatment
		(x,xx) plan
		(x,xx) diabetes
		(x,xx) cancer
59	control	districts
		(x) prevention
		group/s
		variables
		(x,xx) quality
60	evidence	base
		(xx) effectiveness
61	focus	group/s
		(x) interview/s
62	population	(x,xx) cohort
63	satisfaction	survey
		(xx) levels
		(x,xx) scores
64	medication	administration
		(x) error/s
		adherence
		preparation
		side-effect/s
		(x) assessment
		(xx) safety
65	assessment	tool/s
66	risk-assessment	(x) safety planning
		(x) management
	family	
67	assessment	(x) intervention
68	problems	(x,xx) fatigue
		(xx) anxiety
69	interview	guide

250	visits	(xx) inhabitants
251	ethics	committee/s
		(x,xx) approval
		(x) board
252	parenting	role
253	disorder	(x,xx) cohesion
254	duration	(x) dysmenorrhea
255	networks	(xx) relationships
256	sharing	(x,xx) experiences
		(x,xx) information
257	competence	(x) encouraging
258	domains	(xx) responsibilities
259	neighbourhood	(x) inequality
		(x) disorder
		(x,xx) income
		characteristics
		(xx) safety
260	identifying	information
261	consistency	(x) reliability
262	wards	(xx) units
263	literacy	limitations
264	provider	recommendation
265	uncertainty	(xx) relationship
266	lifestyle	changes
267	admission	(x,xx) discharge
268	assessing	(x) managing
		(x,xx) risk
269	online	(x) appendix
		(x,xx) figure
		(x) survey
270	outpatient	(xx) clinic/s
		hemodialysis
271	vaccination	uptake
		rate
272	placement	process
273	massage	(x) child
274	empathy	(x,xx) attitudes
275	disability	handicap
276	predictors	(x,xx) exercise
		(x) depression
277	dimensions	(x,xx) community
278	clinics	(xx) characteristics
279	descriptions	(x,xx) issues
280	addressing	validity
281	construct	(x) mortality
282	morbidity	score
283	predictor	variable
284	coding	process
285	physician	(xx) nurse
286	rounding	chart
287	schizophrenia	(x,xx) disorder

		transcripts	288	impairment	(x,xx) functioning
		(xx) questions	289	indicators	(x,xx) ranges
		(x,xx) participant data	290	adjustment	(x,xx) coping
70	effects	(x) medication	291	focusing	(x) tasks
71	understanding	(x,xx) phenomenon	292	rounds	(x,xx) sites
72	review	board/s	293	shifts	(x,xx) night
		articles	294	respondent	(x) level
		approval	295	nutrition	support
		course	296	translation	process
		literature	297	trauma	center
73	training	courses	298	carer	burden
		(x) skill	299	delirium	(xx) symptoms
		programs/me/mes	300	insight	(x,xx) understanding
		(x,xx) development	301	grandparents	(x) children
74	interviews	(xx) min	302	segregation	barrier/s
75	beliefs	model	303	cardiovascular	(xx) disease
		(x,xx) beliefs			(x) risk
		(x) attitudes	304	center	(x) disease
		(x,xx) practices	305	workload	allocation
		(x,xx) illness	306	strengths	(x) weaknesses
76	model	fit	307	thalassemia	(x,xx) child
77	individuals	(xx) diagnosis	308	clinician	engagement
78	parent	caregiver	309	deterioration	event/s
79	parent-child	relationships	310	staffing	levels
		communication	311	simulation	(x,xx) learning
80	context	(x,xx) culture	312	narratives	(x,xx) experiences
81	community	(x,xx) centers	313	impacts	(x) individual
		(xx) teams	314	facilitating	(x) change
		setting/s			(x,xx) beliefs
		(x,xx) services	315	smokers	(xx) family
		(x,xx) resources	316	illnesses	(xx) cancer
82	lack	(x) continuity	317	stigma	(xx) suicide
		(x) energy	318	hospitalization	(xx) duration
		(x) awareness	319	dynamics	(x,xx) bullying
		(x) confidence	320	triage	scale/s
		(x) engagement			(x) crisis
		(x) knowledge	321	cluster	analysis
		(x) communication	322	induction	(xx) training
		(x) understanding	323	cessation	(xx) intervention
		(x) time	324	onset	(x,xx) dementia
83	characteristics	(x,xx) variables	325	empowerment	(x,xx) family
84	depression	(x,xx) anxiety	326	disabilities	(x,xx) mothers
		(x) fatigue	327	exhaustion	(x) depersonalisation
		(x,xx) sleep	328	teamwork	survey
		(x,xx) depression scale	329	de-escalation	skill
85	caring	(x,xx) empathy	330	leukemia	(x) lymphoma
		behaviors/our/ours	331	tumor	(xx) survivors
		(xx) child	332	cohesion	(x,xx) neighbourhood
86	disease	(x,xx) progression	333	centers	(x) disease
					(xx) control
			334	mentoring	relationships

No.	Nodes	Collocates	No.	Nodes	Collocates
		(xx) prevention	335	yoga	programme
		duration	336	stressors	(xx) life
		burden	337	chemotherapy	(x) radiation
		control	338	extraversion	(x,xx) neuroticism
		(x,xx) disease	339	utilization	process
		(x,xx) treatment	340	homebirth	governance
87	survey	instrument	341	initiating	structure
		respondents			(x,xx) style
		(x,xx) questions			(x) leadership
88	job	satisfaction	342	mentorship	(x,xx) leadership
		demands	343	prognosis	(xx) disease
		performance	344	seclusion	(x) restraint
		stress	345	dyads	(xx) relationship
		(x,xx) job	346	rationing	(x) nursing
		(x,xx) control			(x,xx) care
89	literature	review	347	subgroup	(xx) characteristics
		search	348	mentor	(xx) behaviours
					(x,xx)
90	safety	planning	349	facilitators	implementation
		culture	350	influenza	vaccination
		plan			vaccine
		(x,xx) practices	351	tumors	(xx) tumors
		(x,xx) quality	352	psychiatry	service
		(x) management	353	obesity	(x,xx) wo/men
		issues	354	dyspnea	(x,xx) loneliness
91	questionnaire	item/s	355	internet	access
92	variables	(x,xx) gender	356	accessing	(x,xx) services
		(x,xx) age	357	person-centeredness	(x,xx) care
93	differences	(x,xx) groups	358	hemodialysis	(x) caregiver
		(x,xx) scores	359	inhaler	use
94	importance	(x) having	360	test-retest	reliability
95	response	rate/s	361	telehealth	videoconferencing
		options	362	understandings	(x) recovery
		bias	363	healthcare	personnel
96	communication	networks			costs
		skills			organisations
97	unit	champion			professionals
		operation			system
98	mothers	(x) fathers			providers
		(x,xx) daughters			workers
		(x,xx) survivors			services
99	practices	(x,xx) policies			

L5   Adverb + Adjective   nodes=6, 11 pairs

No.	Nodes	Collocates	No.	Nodes	Collocates
1	significantly	(xx) higher	3	statistically	significant
		(xx) lower	4	potentially	relevant
		(x) likely different	5	culturally	(xx) appropriate specific
2	strongly	disagree/d agree	6	clinically	irrelevant

L6	**Verb + Adverb**	nodes=8, 10 pairs			
No.	**Nodes**	**Collocates**	**No.**	**Nodes**	**Collocates**
1	described	(x,xx) elsewhere	4	discussed	(xx) below
		below	5	viewed	(x) positively
		above	6	transcribed	verbatim
2	showed	(x) statistically	7	correlated	positively
3	noted	(xx) earlier	8	disagree	(x,xx) strongly
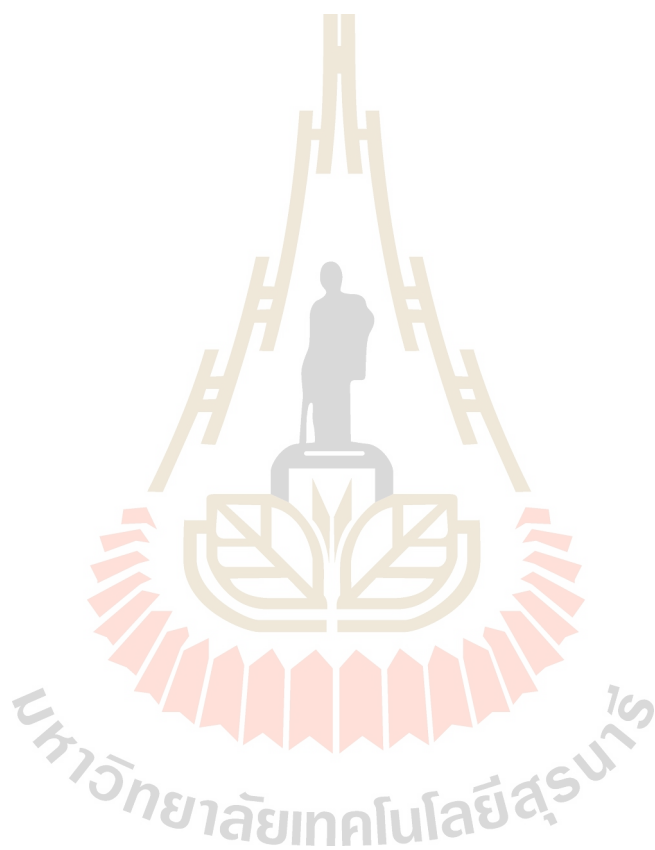
Appendix H

Lexical Collocations with Combination Types
Not in accordance with the Set Framework

N1: Noun + Adjective 68 nodes, 84 pairs

No.	Nodes	Collocates	No.	Nodes	Collocates
1	study	(x) voluntary	31	adolescents	aged
2	nurses	(x,xx) tactile	32	setting	specific
		(x) affective	33	adults	aged
3	patients	(x) stable			(x,xx) mental
4	children	(x,xx) special	34	functioning	(x) everyday
		(x,xx) aged	35	consent	(x,xx) prior
		(x) young	36	finding	(x) consistent
		(x,xx) chronic	37	gender	(x,xx) marital
5	risk	(x,xx) adverse			(xx) male
6	being	able			(x,xx) educational
7	child	(x) comfortable	38	perspective	(xx) interpersonal
		(x,xx) chronic	39	aim	(xx) present
8	use	(x) antipsychotic	40	consumers	(x,xx) medical
9	staff	(x) affective			(x,xx) mental
10	using	(x) thematic	41	transition	(x) tertiary
		(x) statistical	42	prevalence	(x,xx) chronic
		descriptive	43	medications	(x,xx) correct
11	age	(x) marital	44	outcomes	(xx) premature
12	findings	(x) consistent	45	groups	(x) audio-recorded
13	families	(x,xx) chronic	46	relationships	(x,xx) spiritual
14	need	(x,xx) better	47	practitioners	(x,xx) young
15	results	(x) consistent	48	adherence	(x) healthy
		(xx) present	49	awareness	(x) previous
16	levels	(x) perceived	50	improving	(x,xx) physical
17	score/s	(x,xx) greater	51	populations	(xx) rapid
		(x,xx) higher	52	vaccine	behavioral
18	review	(x,xx) qualitative	53	emotions	(x,xx) negative
19	status	(x,xx) educational	54	motivation	(x) physical
20	community	psychiatric	55	disorder	(x,xx) social
		mental	56	neighbourhood	(x,xx) social
21	impact	(x) parental	57	online	supplementary
22	students	(x) registered	58	promoting	adolescent
23	activities	(x) daily	59	bullying	(x,xx) public
24	skills	(x,xx) necessary	60	collaboration	(x) participatory
25	ability	(x,xx) primary	61	trajectory	(x,xx) critical
26	recovery	orientated	62	segregation	(x,xx) mental
		oriented	63	impacts	(x,xx) individual
		focused	64	implementing	(x) new
27	emotion-focused	coping	65	mentors	(x,xx) senior

No.	Nodes	Collocates	No.	Nodes	Collocates
28	participation	(x,xx) voluntary	66	subgroups	(x) significant
29	exercise	(x,xx) older	67	uptake	evidence-based
30	implementation	(x,xx) new	68	constructs	(x,xx) physical

N2: Noun + Adverb   4 nodes, 4 pairs

No.	Nodes	Collocates	No.	Nodes	Collocates
1	parents	(x,xx) newly	3	interviews	(xx) verbatim
2	education	(xx) newly	4	responses	(x,xx) strongly

N3: Verb + Adjective   23 nodes, 34 pairs

No.	Nodes	Collocates	No.	Nodes	Collocates
1	used	(x) cross-sectional	10	showed	(x,xx) significant
2	reported	(x,xx) moderate			(x,xx) higher
3	found	(x,xx) significant	11	identify	(x,xx) potential
4	need	(xx) aware	12	diagnosed	(x)oncology
		(x,xx) better			pediatric
5	included	(x) following	13	demonstrated	(x,xx) significant
6	associated	(x,xx) decreased	14	reflect	(xx) own
		(x,xx) lower	15	indicating	greater
		(x,xx) increased			(x) higher
		(x,xx) higher	16	experiencing	(x) mental
		(x,xx) severe	17	calculated	(x) each
		(x,xx) greater	18	addressed	prior
7	provide	(x) safe	19	resulted	(x,xx) increased
		(x,xx) appropriate	20	implemented	(xx) educational
8	considered	(x,xx) essential	21	perceive	greater
		(x,xx) important	22	focuses	(x,xx) individual
9	indicate/s/d	(x,xx) high/er	23	utilized	(x) post

N4: Adjective + Verb   10 nodes, 13 pairs

No.	Nodes	Collocates	No.	Nodes	Collocates
1	each	(x,xx) rated	6	descriptive	(xx) used
2	important	(x) note	7	ethical	(xx) obtained
		(x, xx) consider	8	registered	(x) working
3	significant	(xx) noted	9	statistical	(x,xx) performed
4	higher	(x) indicate/s/ing			(x,xx) set
5	previous	(x,xx) showing	10	eligible	(x) participate
		(xx) shown			

N5: Adjective + Adjective   51 nodes, 82 pairs

No.	Nodes	Collocates	No.	Nodes	Collocates
1	social	(x,xx) political	23	cognitive	affective
		(x,xx) economic			(x) somatic
		cognitive			(x,xx) behavio(u)ral
		(x) cultural			(x,xx) emotional
2	physical	(x,xx) psychosocial	24	multiple	logistic

No.	Nodes	Collocates
		(x,xx) psychological
		(x) emotional
		(x,xx) sexual
		(x,xx) cognitive
3	high	(x,xx) low
4	medical	(x) surgical
		(x) primary
5	higher	(xx) greater
		(xx) higher
6	individual	(x,xx) environmental
7	positive	(x) negative
		(x) therapeutic
8	professional	(x) familial
		(x) personal
9	perceived	(x,xx) parental
		(x,xx) negative
		(xx) social
10	negative	(xx) psychological
11	personal	(x,xx) professional
12	primary	(x) secondary
13	emotional	(x,xx) spiritual
14	low	(x) moderate
		(x,xx) high
15	psychological	(x,xx) social
		(x) emotional
16	qualitative	(x) quantitative
		(x) descriptive
17	limited	(x,xx) available
18	chronic	obstructive
		(x) pulmonary
19	educational	(xx) material
20	overall	mean
21	lower	socioeconomic
		(x,xx) educational
		(xx) perceived
22	acute	lymphoblastic
		(x) psychiatric
		(x,xx) chronic
		linear
25	pediatric	critical
26	adult	(x) spousal
		(xx) mental
27	palliative	(x) supportive
28	female	(x,xx) male
29	demographic	(x) clinical
30	organisational	(xx) organisational
31	severe	(xx) mental
32	consistent	(x,xx) previous
33	aged	(x,xx) older
34	internal	(x,xx) external
35	psychosocial	(x,xx) spiritual
		(x) emotional
36	descriptive	(x) correlational
		(xx) qualitative
37	registered	(x,xx) practical
38	moderate	(x) severe
		(x) high
39	cross-sectional	descriptive
40	affective	(x) somatic
41	paediatric	(x) intensive
42	engaging	(x) supportive
43	quantitative	(x) qualitative
44	cardiac	educational
45	marital	(x,xx) educational
46	physiological	(x,xx) psychological
47	tertiary	medical
		(x,xx) primary
48	sociodemographic	(x,xx) clinical
49	randomized	controlled
		(x) clinical
50	neonatal	intensive
51	dyadic	individual

N6: Verb + Verb   12 nodes, 23 pairs

No.	Nodes	Collocates	No.	Nodes	Collocates
1	used	(x) analyse/ze	4	needed	(x) determine
		(x) compare			(x) explore
		(x) collect			(x) help
		(x) assess	5	informed	(xx) obtained
		(x) measure	6	completed	(x,xx) returned
		(x) examine	7	diagnosed	(xx) living
		(x) describe	8	stated	(x,xx) know
		(x) evaluate	9	recruited	(x) participate
2	need	(x) consider	10	reviewed	(x) approved

		(x) develop	11	interviewed	expressed
		(x) understand	12	disagree	(xx) agree
3	provided	(x,xx) written			
		(x,xx) informed			

N7: Adverb + Verb   6 nodes, 12 pairs

No.	Nodes	Collocates	No.	Nodes	Collocates
1	significantly	correlated	4	positively	associated
		associated			related
		(x,xx) related	5	negatively	correlated
		increased			affect
2	strongly	associated			associated
3	specifically	designed	6	independently	coded

N8: Adverb + Noun   4 nodes, 7 pairs

No.	Nodes	Collocates	No.	Nodes	Collocates
1	significantly	(x,xx) scores	3	statistically	(x) difference/s
2	approximately	(x) min(utes)			(x) correlation
		half	4	positively	(x,xx) job
		(xx) people			

Appendix I

List of Keywords According to Parts of Speech

NOUNS

ability	constructs	identity	parent	sessions
access	consumers	illness	parent-child	setting
accessing	content	illnesses	parenting	settings
actions	context	impact	parents	severity
activities	control	impacts	participant	sex
activity	conversations	impairment	participants	sharing
addressing	coping	implementation	participation	shift
adherence	correlation	implementing	pathway	shifts
adjustment	correlations	importance	patient	siblings
administration	couples	improving	patients	simulation
admission	crisis	inclusion	perception	situations
adolescent	criteria	indicators	perceptions	skills
adolescents	culture	individuals	person	sleep
adults	data	induction	person-centeredness	smokers
age	decisions	influence	perspective	smoking
aggression	de-escalation	influenza	perspectives	spirituality
aim	delirium	information	physician	spouses
alcohol	delivery	inhaler	physicians	staff
analyses	dementia	initiating	placement	staffing
analysis	depression	inpatient	population	statistics
anxiety	descriptions	insight	populations	status
approach	deterioration	instrument	practice	stigma
approaches	diabetes	interaction	practices	strategies
articles	diagnoses	interactions	practitioners	strengths
aspects	diagnosis	internet	predictor	stress
assessing	differences	intervention	predictors	stressors
assessment	difficulties	interventions	prevalence	students
association	dimensions	interview	prevention	studies
associations	disabilities	interviewer	problems	study
attitudes	disability	interviews	procedures	subgroup
authors	discharge	investigator	process	subgroups
awareness	discourse	involvement	processes	subscale
barriers	discourses	issues	prognosis	subscales
baseline	disease	item	program	suicide
behavior	disorder	items	programs	support
behaviors	disorders	job	promoting	surgery
behaviours	distress	knowledge	provider	survey
being	documentation	lack	providers	survivors
beliefs	domains	leadership	providing	symptom

bias	duration	learning	psychiatry	symptoms
birth	dyads	leukemia	quality	tasks
bullying	dynamics	level	questionnaire	tattoos
burden	dyspnea	levels	questionnaires	team
burnout	education	life	questions	teamwork
cancer	educators	lifestyle	rationing	telehealth
cardiovascular	effectiveness	limitation/s	receiving	test-retest
care	effects	literacy	recommendations	thalassemia
caregiver	emergency	literature	recovery	theme
caregivers	emotion-focused	living	recruitment	themes
caregiving	emotions	loneliness	reducing	therapy
carer	empathy	majority	regression	tool
carers	empowerment	management	relation	tools
caring	engagement	managers	relationship	topics
categories	environment	managing	relationships	training
category	errors	massage	reliability	trajectory
center	ethics	meaning	reporting	transcripts
centers	evaluation	meaning-making	research	transition
cessation	evidence	measures	researcher	translation
challenges	exercise	medication	researchers	trauma
changes	exhaustion	medications	residents	treatment
characteristics	expectations	medicines	resources	treatments
chemotherapy	experience	members	respondent	triage
child	experiences	mentor	respondents	trials
childhood	extraversion	mentoring	response	trust
children	facilitating	mentors	responses	tumor
clinic	facilitators	mentorship	responsibility	tumors
clinician	factor	methods	restraint	uncertainty
clinicians	factors	midwives	results	understanding
clinics	families	model	review	understandings
cluster	family	models	risk	unit
codes	fatigue	morbidity	risk-assessment	units
coding	feedback	mortality	risks	uptake
coefficient/s	feelings	mothers	role	use
cohesion	finding	motivation	roles	users
cohort	findings	narratives	rounding	using
collaboration	focus	need	rounds	utilization
colleagues	focusing	needs	routines	vaccination
collection	follow-up	neighbourhood	safety	vaccine
comfort	framework	networks	sample	validity
community	frequency	neuroticism	sampling	values
communication	functioning	nurse	satisfaction	variable
community	gender	nurses	scale	variables
competence	grandparents	nursing	scales	version
competencies	group	nutrition	schizophrenia	violence

complications	groups	obesity	score/s	visits
components	guidelines	oncology	screening	ward
concerns	health	online	seclusion	wards
condition	hemodialysis	online	segregation	work
confidence	homebirth	onset	self-care	workload
consent	hospital	outcome	self-efficacy	workplace
consequences	hospitalization	outcomes	self-harm	yoga
consistency	hospitals	outpatient	service	
construct	identifying	pain	services	

VERBS

access	decrease	explored	living	recruited
address	demonstrated	expressed	manage	reduce
addressed	describe	facilitate	measure	reflect
affect	described	facilitated	measured	reported
affected	determine	focuses	need	resulted
analys/zed	develop	found	needed	reviewed
assess	diagnosed	highlighted	noted	selected
assessed	disagree	identified	observed	showed
associated	discussed	identify	obtained	stated
bereaved	engage	impacted	participate	suggest
calculated	enhance	implemented	participated	transcribed
coded	enrolled	improve	perceive	translated
collected	evaluate	improved	performed	undergoing
compared	evaluated	included	promote	understand
completed	examine	indicate/s/d	provide	use
conducted	examined	indicating	provided	used
considered	excluded	influenced	ranged	utilized
consisted	experienced	inform	rated	viewed
contribute	experiencing	informed	received	
correlated	explore	interviewed	receiving	

ADJECTIVES

acute	developmental	increased	palliative	self-administered
adult	different	individual	parental	self-reported
adverse	discursive	influencing	participating	semi/structured
affective	diverse	initial	pediatric	sensory
aged	documented	institutional	perceived	severe
antipsychotic	dyadic	intensive	personal	sexual
appropriate	each	internal	person-centred	shared
additional	educational	interpersonal	physical	significant
behavio(u)ral	effective	interprofessional	physiological	similar
bereaved	eligible	limited	positive	social
biomedical	emotional	logistic	postoperative	sociodemographic
cardiac	engaging	longitudinal	potential	socioeconomic
challenging	ethical	low	preoperative	somatic

chronic	everyday	lower	present	specific
clinical	evidence-based	marital	previous	standardized
cognitive	familial	mean	primary	statistical
collaborative	family-centered	meaningful	prior	stressful
completed	female	medical	problematic	suicidal
consistent	functional	mental	professional	supportive
contextual	geriatric	moderate	psychiatric	surgical
core	greater	multidisciplinary	psychological	systematic
critical	grounded	multiple	psychometric	tactile
cross-sectional	healthy	narrative	psychosocial	tertiary
cultural	helpful	negative	qualitative	thematic
current	high	neonatal	quantitative	theoretical
customer-oriented	higher	nurse-led	quantitative	therapeutic
daily	high-risk	older	randomized	total
decreased	holistic	ongoing	registered	validated
demographic	hospital-based	oral	relational	verbatim
depressive	hospitalized	organis(z)ational	relevant	
descriptive	important	overall	respiratory	
developed	included	paediatric	selected	

ADVERBS

approximately	culturally	negatively	potentially	specifically
clinically	independently	positively	significantly	statistically
				strongly

Appendix J

List of 200 Most Frequent Lexical Collocations in the SCNRA

No.	Nodes	Collocates	Freq.	MI scores
1	mental	(*ill, and physical*) health	1699	7.18099
2	health	care	1516	5.46083
3	family	member/s	1075	7.70444
4	health	service/s	514	5.39321
5	physical	(*ill, and mental*) health	385	5.60188
6	nursing	home/s	366	7.285255
7	mental	illness/es	352	6.83562
8	care	provider/s	328	6.655335
9	service	user/s	326	10.0054
10	present	study	306	6.62802
11	palliative	(*and supportive*) care	304	7.43407
12	family	caregiver/s	303	6.139965
13	social	support	300	7.04352
14	physical	activity	281	8.9678
15	health	(*care*) provider/s	280	6.430495
16	data	collection	264	8.88004
17	quality	(*of, of nursing*) care	261	5.574
18	mental	(*health*) service/s	254	5.823345
19	nursing	practice	226	5.87812
20	job	satisfaction	225	9.40835
21	focus	group/s	222	7.47001
22	quality	(*of*) life	221	7.68685
23	previous	studies	217	7.63892
24	significant	difference/s	214	8.255645
25	patient	education	201	5.91467
26	chronic	(*physical*, *conditions for*)	190	9.096985
27	higher	(*mean*, *and lower*) score/s	187	7.241805
28	risk	assessment	187	7.50134
29	registered	nurse/s	186	6.832025
30	nurse	manager/s	185	7.89047
31	older	adult/s	182	8.220515
32	high	level/s	176	6.55749
33	older	people	176	7.44611
34	mean	score/s	175	7.702515
35	sample	size/s	172	9.547435
36	data	(*were*) collected	168	8.1993
37	nursing	student/s	168	6.240085

No.	Nodes	Collocates	Freq.	MI scores
38	clinical	(*nursing*) practice	166	6.75781
39	children	(*with*, *diagnosed with*) cancer	161	6.47166
40	age	(*of -*, *ranged between -*) years	160	6.82058
41	higher	level/s	160	6.569775
42	primary	care	158	5.7012
43	risk	factor/s	157	5.936195
44	pediatric	oncology	156	11.07548
45	nursing	staff	155	5.55006
46	aged	(*under -*) years	154	8.92289
47	statistically	significant	154	10.01163
48	patients	(*and*, *and their*) families	152	5.01531
49	parents	(*of*, *and their*) children	150	5.69411
50	critical	care	149	6.06109
51	previous	research	147	6.76126
52	total	(*mean*, *health literacy*) score/s	147	7.81536
53	health	problems	145	5.30555
54	health	status	142	5.31281
55	patient	safety	142	6.67556
56	current	study	141	5.83692
57	study	(*was*) conducted	138	5.38618
58	research	team	136	6.5534
59	data	analysis	135	5.63066
60	health	outcomes	134	5.20824
61	health	literacy	133	7.3999
62	acute	(*psychiatric*, *and primary*) care	130	5.9923
63	informed	consent	128	10.34874
64	providing	(*quality*, *efficient health*) care	126	5.76249
65	intensive	(*follow-up*, *support and*) care	125	6.90627
66	illness	belief/s	123	6.569745
67	score/s	indicate/ed/ing	122	6.9964388
68	childhood	cancer	121	9.10938
69	chronic	(*disease*) condition/s	121	8.63409
70	psychological	distress	118	9.37247
71	primary	(*family*) caregiver/s	117	7.798315
72	systematic	review/s	117	10.390245
73	nurses	working	116	5.32592
74	medication	administration	115	9.55464
75	inclusion	(*and exclusion*) criteria	114	10.57658
76	participate	(*in the*) study	113	5.8202
77	assessment	tool/s	112	8.01942
78	mean	age	110	7.14542
79	qualitative	(*research*) study/ies	110	5.430305
80	care	settings	109	5.3268
81	health	system/s	109	5.26906

No.	Nodes	Collocates	Freq.	MI scores
82	health	issues	109	5.05659
83	depressive	symptoms	107	10.22404
84	control	group/s	105	6.03599
85	patient	satisfaction	104	5.92498
86	research	question/s	102	5.721915
87	emergency	department/s	100	10.544365
88	review	board/s	100	9.470885
89	internal	consistency	99	11.56271
90	strongly	agree	98	10.24256
91	age	group/s	97	5.559825
92	response	rate/s	97	7.53264
93	anxiety	(*and*) depression	95	8.64065
94	age	(*and*) gender	94	8.36083
95	coping	strategy/ies	94	8.743655
96	family	functioning	94	6.4883
97	workplace	spirituality	94	11.3839
98	daily	life/ves	93	7.44662
99	risk	management	93	6.00847
100	adolescent	(*mental*) health	92	5.60674
101	chronic	(*physical*, *obstructive*	91	8.803585
102	demographic	(*and socioeconomic*)	91	9.23526
103	everyday	life/ves	91	8.39174
104	primary	family	90	5.64974
105	each	(*questionnaire*) item	88	7.42557
106	work	environment/s	88	6.66943
107	aim	(*of this*) study	87	6.09552
108	descriptive	statistics	87	11.24674
109	health	behaviours	86	5.46625
110	knowledge	(*and*) skill/s	86	6.861275
111	marital	status	86	10.60619
112	sleep	(*and appetite*) problem/s	86	7.32009
113	team	members	86	6.73214
114	medication	(*administration*) error/s	85	8.74906
115	pain	management	85	6.80611
116	staff	member/s	85	5.477155
117	caring	behaviors/our/ours	84	7.1793525
118	psychiatric	nurses	83	5.38188
119	institutional	review	81	9.35581
120	children	(*and*, *and young*) adolescents	80	7.14593
121	clinical	setting/s	80	6.542585
122	institutional	(*review*) board/s	80	10.79775
123	acute	(*care*, *care hospital*) setting/s	79	8.137035
124	content	analysis	79	7.44504
125	limitation/s	(*of the*) study	79	5.355845

No.	Nodes	Collocates	Freq.	MI scores
126	participants	(*were*) asked	79	6.01484
127	child	(*with*, *diagnosed with*) cancer	78	5.72764
128	ethics	committee/s	78	11.690585
129	leadership	style/s	78	10.39333
130	score/s	range/s/d	78	7.492545
131	alcohol	consumption	77	12.07942
132	surgical	patients	77	6.09286
133	tactile	massage	77	13.13364
134	regression	(*and meditation*) analysis/es	76	8.4296
135	risk	taking	76	7.46499
136	social	capital	76	8.68454
137	increased	risk	75	6.38798
138	pediatric	(*oncology*) patients	75	5.77861
139	high	school	74	7.54612
140	mental	(*health*) problems	74	5.87767
141	analys/zed	using	73	7.874495
142	evidence-based	practice/s	73	5.06217
143	oral	care	73	6.17471
144	screening	tool/s	73	9.56171
145	care	units	72	5.459
146	higher	(*scores*) indicate/s/ing	72	7.3265225
147	intensive	(*care*) unit/s	72	9.15044
148	lower	(*education*, *baseline energy*)	72	6.636615
149	qualitative	(*exploratory*) research	72	6.36078
150	relationship	quality	71	6.41975
151	used	(*to*) measure	71	6.6832
152	paediatric	nurses	70	6.96883
153	safety	planning	70	9.4534
154	adult	(*mental*) health	69	5.29067
155	diagnosed	(*with*, *with cervical*) cancer	69	7.68772
156	medication	adherence	69	9.2047
157	outpatient	(*heart failure*) clinic/s	69	10.688345
158	sensory	room/s	69	10.753215
159	severe	(*level of*) pain	69	8.19037
160	strongly	disagree/d	69	11.29134
161	being	able	68	6.68901
162	smoking	cessation	68	10.94261
163	emotional	exhaustion	67	10.85763
164	positive	(*and*) negative	67	7.27918
165	score/s	(*indicated*, *indicates a*) higher	67	5.905485
166	significant	correlation/s	67	7.937
167	diabetes	(*management*) education	66	7.3443
168	ethical	approval	66	11.2317
169	health	crisis/es	66	5.60456

No.	Nodes	Collocates	Freq.	MI scores
170	intervention	(*and control*) group/s	66	5.26845
171	low	level/s	66	5.99803
172	nurse	leader/s	66	6.89521
173	research	ethics	66	7.78759
174	risk	(*and*, *assessment and*) safety	66	6.28379
175	semi/structured	interview/s	66	8.44483
176	study	aims/ed	66	5.780475
177	results	show/ed	65	6.98478
178	significantly	(*associated with*) higher	65	7.71748
179	educational	(*programs and*) intervention/s	64	6.680835
180	patient	outcomes	64	5.25737
181	studies	(*were*, *have been*) conducted	64	5.92268
182	data	(*were*) analysed/zed	63	6.91889
183	each	participant	63	6.58316
184	educational	(*and income*) level/s	62	6.261525
185	family	(*and*, *members and*) friends	62	6.66246
186	group	interview/s	62	5.390845
187	nursing	interventions	62	5.09223
188	parental	presence	62	9.14564
189	statistically	(*significant*) difference/s	62	8.982385
190	community	mental	61	5.75027
191	educational	program/s/me/mes	61	7.53087
192	oncology	patients	61	5.91493
193	consent	form/s	60	8.815495
194	interviews	(*were*) conducted	60	7.25117
195	literature	review	60	7.55423
196	social	network/s	60	7.413565
197	symptoms	(*of*, *such as*) depression	60	7.23027
198	training	programs/me/mes	60	7.446715
199	affective	commitment	59	11.65039
200	community	setting/s	59	6.899865

Appendix K

List of 100 Most Frequent Adjective + Noun Collocations found from the SCNRA

No.	Nodes	Collocates	No.	Nodes	Collocates
1	mental	(*ill*, *and physical*) health	51	tactile	massage
2	physical	(*ill*, *and physical*) health	52	social	capital
3	mental	illness/es	53	increased	risk
4	present	study	54	pediatric	(*oncology*) patients
5	palliative	(*and supportive*) care	55	mental	(*health*) problems
6	social	support	56	high	school
7	physical	activity	57	evidence-based	practice/s
8	mental	(*health*) service/s	58	oral	care
9	previous	studies	59	qualitative	(*exploratory*) research
10	significant	difference/s (*physical*, *conditions for*)	60	lower	(*education*, *baseline energy*) level/s
11	chronic	illness/es	61	intensive	(*care*) unit/s
12	higher	(*mean*, *and lower*) score/s	62	paediatric	nurses
13	registered	nurse/s	63	adult	(*mental*) health
14	older	adult/s	64	severe	(*level of*) pain
15	high	level/s	65	sensory	room/s
16	older	people	66	significant	correlation/s
17	mean	score/s	67	emotional	exhaustion
18	clinical	(*nursing*) practice	68	low	level/s
19	higher	level/s	69	ethical	approval
20	primary	care	70	semi/structured	interview/s (*programs and*)
21	pediatric	oncology	71	educational	intervention/s
22	aged	(*under -*) years	72	each	participant
23	critical	care	73	parental	presence
24	previous	research (*mean*, *health literacy*)	74	educational	(*and income*) level/s
25	total	score/s	75	educational	program/s/me/mes
26	current	study (*psychiatric*, *and primary*)	76	social	network/s
27	acute	care (*follow-up*, *support and*)	77	mean	(*age -*) years
28	intensive	care	78	affective	commitment
29	chronic	(*disease*) condition/s	79	positive	outcomes
30	psychological	distress	80	emotional	support
31	primary	(*family*) caregiver/s	81	daily	living
32	systematic	review/s	82	cognitive	impairment
33	mean	age	83	thematic	(*content*) analysis
34	qualitative	(*research*) study/ies	84	social	cohesion
35	depressive	symptoms	85	negative	(*health*) effects

No.	Nodes	Collocates	No.	Nodes	Collocates
36	internal	consistency	86	mental	(*health*) crisis/es
37	daily	life/ves	87	mental	(*health*) issues
38	chronic	(*physical*, *obstructive pulmonary*) disease/s (*and socioeconomic*)	88	different	types
39	demographic	characteristics	89	high	(*response*) rate/s
40	everyday	life/ves	90	medical	records
41	primary	family	91	individual	(*in-depth*) interviews
42	each	(*questionnaire*) item	92	positive	effect/s
43	descriptive	statistics	93	significant	relationship/s
44	marital	status	94	socioeconomic	status
45	psychiatric	nurses	95	emotional	distress
46	institutional	review	96	organisational	culture
47	clinical	setting/s	97	social	worker/s
48	institutional	(*review*) board/s (*care*, *care hospital*)	98	negative	emotions
49	acute	setting/s	99	educational	attainment
50	surgical	patients	100	mental	(*health*) triage
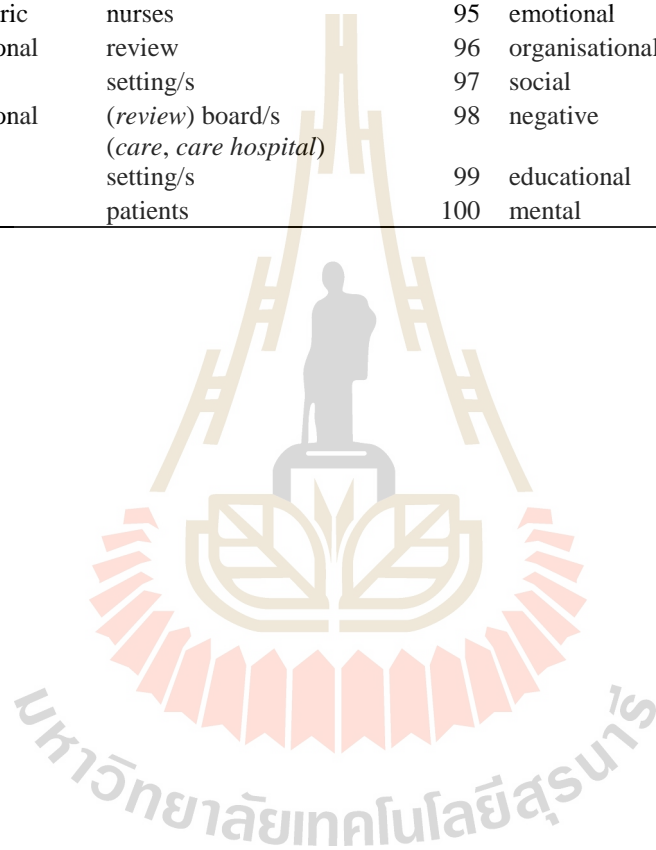
Appendix L

List of 100 Most Frequent Noun + Noun Collocations found from the

SCNRA

No.	Nodes	Collocates	No.	Nodes	Collocates
1	health	care	51	age	(*and*) gender
2	family	member/s	52	coping	strategy/ies
3	health	service/s	53	workplace	spirituality
4	nursing	home/s	54	risk	management
5	care	provider/s	55	adolescent	(*mental*) health
6	service	user/s	56	work	environment/s
7	family	caregiver/s	57	aim	(*of this*) study
8	health	(*care*) provider/s	58	health	behaviours
9	data	collection	59	knowledge	(*and*) skill/s
10	quality	(*of*, *of nursing*) care	60	team	members
11	nursing	practice	61	sleep	(*and appetite*) problem/s
12	job	satisfaction	62	staff	member/s
13	focus	group/s	63	pain	management
14	quality	(*of*) life	64	medication	(*administration*) error/s
15	patient	education	65	caring	behaviors/our/ours
16	risk	assessment	66	children	(*and*, *and young*) adolescents
17	nurse	manager/s	67	content	analysis
18	sample	size/s	68	limitation/s	(*of the*) study
19	nursing	student/s	69	child	(*with*, *diagnosed with*) cancer
20	children	(*with*, *diagnosed with*) cancer	70	leadership	style/s
21	age	(*of -*, *ranged between -*) years	71	ethics	committee/s
22	risk	factor/s	72	alcohol	consumption
23	nursing	staff	73	risk	taking
24	patients	(*and*, *and their*) families	74	regression	(*and meditation*) analysis/es
25	parents	(*of*, *and their*) children	75	screening	tool/s
26	health	problems	76	care	units
27	health	status	77	relationship	quality
28	patient	safety	78	safety	planning
29	research	team	79	medication	adherence
30	data	analysis	80	outpatient	(*heart failure*) clinic/s
31	health	outcomes	81	smoking	cessation
32	health	literacy	82	health	crisis/es
33	providing	(*quality*, *efficient health*) care	83	research	ethics
34	illness	belief/s	84	nurse	leader/s
35	childhood	cancer	85	risk	(*and*, *assessment and*) safety
36	medication	administration	86	intervention	(*and control*) group/s
37	inclusion	(*and exclusion*) criteria	87	diabetes	(*management*) education
38	assessment	tool/s	88	patient	outcomes
39	care	settings	89	family	(*and*, *members and*) friends
40	health	system/s	90	nursing	interventions
41	health	issues	91	group	interview/s
42	control	group/s	92	oncology	patients
43	patient	satisfaction	93	symptoms	(*of*, *such as*) depression
44	research	question/s	94	training	programs/me/mes

No.	Nodes	Collocates	No.	Nodes	Collocates
45	review	board/s	95	literature	review
46	emergency	department/s	96	consent	form/s
47	age	group/s	97	diagnosis	(*and*) treatment
48	response	rate/s	98	community	setting/s
49	anxiety	(*and*) depression	99	discharge	education
50	family	functioning	100	nurse	staffing
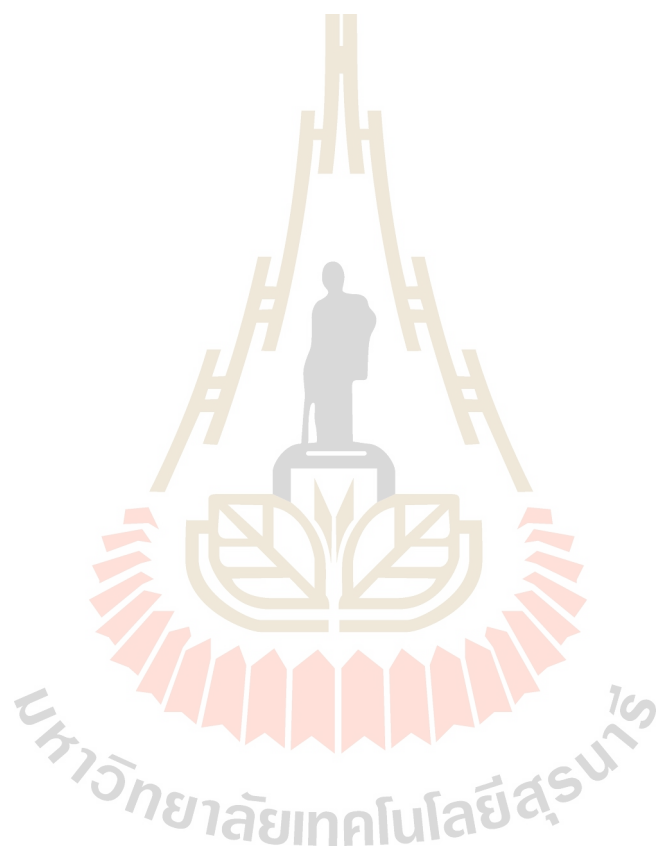
Appendix M

List of 200 most frequent collocations in SCNRA based on each node

No.	Nodes	Collocates	Freq.	MI scores
1	mental	(*ill, and physical*) health	1699	7.18099
2	health	care	1516	5.46083
3	family	member/s	1075	7.70444
4	physical	(*ill, and mental*) health	385	5.60188
5	nursing	home/s	366	7.285255
6	care	provider/s	328	6.655335
7	service	user/s	326	10.0054
8	present	study	306	6.62802
9	palliative	(*and supportive*) care	304	7.43407
10	social	support	300	7.04352
11	data	collection	264	8.88004
12	quality	(*of, of nursing*) care	261	5.574
13	job	satisfaction	225	9.40835
14	focus	group/s	222	7.47001
15	previous	studies	217	7.63892
16	significant	difference/s	214	8.255645
17	patient	education	201	5.91467
18	chronic	(*physical, conditions for*) illness/es	190	9.096985
19	higher	(*mean, and lower*) score/s	187	7.241805
20	risk	assessment	187	7.50134
21	registered	nurse/s	186	6.832025
22	nurse	manager/s	185	7.89047
23	older	adult/s	182	8.220515
24	high	level/s	176	6.55749
25	mean	score/s	175	7.702515
26	sample	size/s	172	9.547435
27	clinical	(*nursing*) practice	166	6.75781
28	children	(*with, diagnosed with*) cancer	161	6.47166
29	age	(*of -, ranged between -*) years	160	6.82058
30	primary	care	158	5.7012
31	pediatric	oncology	156	11.07548
32	aged	(*under -*) years	154	8.92289
33	statistically	significant	154	10.01163
34	patients	(*and ,and their*) families	152	5.01531
35	parents	(*of ,and their*) children	150	5.69411
36	critical	care	149	6.06109
37	total	(*mean, health literacy*) score/s	147	7.81536

No.	Nodes	Collocates	Freq.	MI scores
38	current	study	141	5.83692
39	study	(*was*) conducted	138	5.38618
40	research	team	136	6.5534
41	acute	(*psychiatric, and primary*) care	130	5.9923
42	informed	consent	128	10.34874
43	providing	(*quality, efficient health*) care	126	5.76249
44	intensive	(*follow-up, support and*) care	125	6.90627
45	illness	belief/s	123	6.569745
46	score/s	indicate/ed/ing	122	6.9964388
47	childhood	cancer	121	9.10938
48	psychological	distress	118	9.37247
49	systematic	review/s	117	10.390245
50	nurses	working	116	5.32592
51	medication	administration	115	9.55464
52	inclusion	(*and exclusion*) criteria	114	10.57658
53	participate	(*in the*) study	113	5.8202
54	assessment	tool/s	112	8.01942
55	qualitative	(*research*) study/ies	110	5.430305
56	depressive	symptoms	107	10.22404
57	control	group/s	105	6.03599
58	emergency	department/s	100	10.544365
59	review	board/s	100	9.470885
60	internal	consistency	99	11.56271
61	strongly	agree	98	10.24256
62	response	rate/s	97	7.53264
63	anxiety	depression	95	8.64065
64	coping	strategy/ies	94	8.743655
65	workplace	spirituality	94	11.3839
66	daily	life/ves	93	7.44662
67	adolescent	(*mental*) health	92	5.60674
68	demographic	(*and socioeconomic*) characteristics	91	9.23526
69	each	(*questionnaire*) item	88	7.42557
70	work	environment/s	88	6.66943
71	aim	(*of this*) study	87	6.09552
72	descriptive	statistics	87	11.24674
73	knowledge	(*and*) skill/s	86	6.861275
74	marital	status	86	10.60619
75	sleep	(*and appetite*) problem/s	86	7.32009
76	team	members	86	6.73214
77	pain	management	85	6.80611
78	staff	member/s	85	5.477155
79	caring	behaviors/our/ours	84	7.1793525
80	psychiatric	nurses	83	5.38188
81	institutional	review	81	9.35581

No.	Nodes	Collocates	Freq.	MI scores
82	content	analysis	79	7.44504
83	limitation/s	(*of the*) study	79	5.355845
84	participants	(*were*) asked	79	6.01484
85	child	(*with, diagnosed with*) cancer	78	5.72764
86	leadership	style/s	78	10.39333
87	alcohol	consumption	77	12.07942
88	surgical	patients	77	6.09286
89	tactile	massage	77	13.13364
90	regression	(*and meditation*) analysis/es	76	8.4296
91	increased	risk	75	6.38798
92	analys/zed	using	73	7.874495
93	evidence-based	practice/s	73	5.06217
94	oral	care	73	6.17471
95	screening	tool/s	73	9.56171
96	lower	(*education, baseline energy*) level/s	72	6.636615
97	relationship	quality	71	6.41975
98	used	(*to*) measure	71	6.6832
99	paediatric	nurses	70	6.96883
100	safety	planning	70	9.4534
101	adult	(*mental*) health	69	5.29067
102	diagnosed	(*with, with cervical*) cancer	69	7.68772
103	outpatient	(*heart failure*) clinic/s	69	10.688345
104	sensory	room/s	69	10.753215
105	severe	(*level of*) pain	69	8.19037
106	being	able	68	6.68901
107	smoking	cessation	68	10.94261
108	emotional	exhaustion	67	10.85763
109	positive	(*and*) negative	67	7.27918
110	diabetes	(*management*) education	66	7.3443
111	ethical	approval	66	11.2317
112	intervention	(*and control*) group/s	66	5.26845
113	low	level/s	66	5.99803
114	semi/structured	interview/s	66	8.44483
115	results	show/ed	65	6.98478
116	significantly	(*associated with*) higher	65	7.71748
117	educational	(*programs and*) intervention/s	64	6.680835
118	studies	(*were, have been*) conducted	64	5.92268
119	group	interview/s	62	5.390845
120	parental	presence	62	9.14564
121	oncology	patients	61	5.91493
122	consent	form/s	60	8.815495
123	interviews	(*were*) conducted	60	7.25117
124	literature	review	60	7.55423
125	symptoms	(of, *such as*) depression	60	7.23027

No.	Nodes	Collocates	Freq.	MI scores
126	training	programs/me/mes	60	7.446715
127	affective	commitment	59	11.65039
128	community	setting/s	59	6.899865
129	diagnosis	(*and, and specific*) treatment	59	6.34207
130	disagree	(*to strongly*) agree	59	12.34105
131	discharge	education	59	7.45818
132	manage	(*mental, their physical*) health	59	5.12137
133	cognitive	impairment	57	10.71249
134	thematic	(*content*) analysis	57	8.81699
135	bereaved	(*by*) suicide	56	10.75413
136	findings	suggest	56	7.35298
137	measured	using	56	7.29466
138	negative	(*health*) effects	56	6.37994
139	practice	environment/s	56	6.372525
140	different	types	55	7.95275
141	individual	(*qualitative*) interviews	55	6.75376
142	medical	records	55	8.263185
143	hospital	stay/s	54	8.338785
144	scale	ranged/ing	54	7.746915
145	socioeconomic	status	54	9.08523
146	healthcare	professionals	53	7.63016
147	organisational	culture	53	8.92321
148	transcribed	verbatim	53	12.99078
149	condition	management	52	7.60663
150	interview	data	52	5.30888
151	cancer	treatment	51	5.51064
152	cross-sectional	(*research, survey research*) design	51	10.30939
153	education	programs/me/mes	51	6.021365
154	antipsychotic	medication	50	10.06543
155	grounded	theory	50	11.19794
156	personal	(*and professional*) experience/s	50	5.47017
157	person-centred	care	50	6.28034
158	sex	ratio/s	50	10.27748
159	statistical	analysis/es	50	7.79329
160	factors	(*that*) influence/d	49	6.32096
161	mortality	rate/s	49	7.814405
162	information	(*was, could be*) provided	48	5.44169
163	person	centredness	48	10.50962
164	reliability	(*and, and construct*) validity	47	9.14541
165	self-efficacy	expectations	47	9.28383
166	associated	(*with, with a*) higher	46	5.93423
167	correlation	coefficient/s	46	10.469305
168	deterioration	event/s	46	10.18993
169	geriatric	team	46	9.97692

| No.	Nodes	Collocates	Freq.	MI scores
170	majority	(*of, of the*) participants	46	5.71545
171	professional	identity	46	8.5053
172	provide	(*appropriate, the needed*) support	46	5.34203
173	activities	(*of*) daily	45	7.64633
174	depression	(*and*) anxiety	45	7.56265
175	factor	structure	45	8.82845
176	logistic	(*and linear*) regression	45	12.51813
177	longitudinal	(*aging*) study/ies	45	6.14983
178	multiple	(*linear*) regression	45	9.73073
179	risk-assessment	(*and safety*) planning	45	8.53102
180	unit	operation	45	9.93697
181	questionnaire	item/s	44	6.06187
182	self-care	agency	44	5.82961
183	therapeutic	conversation/s	44	9.699
184	suicide	death/s	43	8.547
185	important	role	42	5.56348
186	lack	(*of*) knowledge	42	5.98375
187	theoretical	framework	42	10.742795
188	consistent	(*with, with the*) previous	41	7.83332
189	improve	(*the, access and*) quality	41	6.94804
190	inpatient	care	41	5.19948
191	tertiary	medical	41	9.33307
192	caregiver	(*stress and*) burden	40	9.32308
193	randomized	(*controlled, controlled clinical*)	40	11.65146
194	access	(*the, mental health*) services	39	6.6829
195	crisis	response/s	39	7.287955
196	meaning	units	39	8.9866
197	online	supplementary	39	12.56247
198	perceived	(*social, higher family*) support	39	5.32784
199	psychometric	properties	39	13.64112
200	validity	(*and*) reliability	39	8.87622

Appendix N

The Raw scores of the Pre-test and the Post-test

No.	Id	Pre Part1	Post Part1	Pre Part2	Post Part2	Pre part3	Post Part3	Pre Total	Post Total	Pre Total(%)	Post Total(%)
1	B5790012	21	23	13	16	3	8	37	47	61.67	78.33
2	B5790029	23	24	12	17	8	7	43	48	71.67	80.00
3	B5790036	22	27	12	16	5	5	39	48	65.00	80.00
4	B5790043	16	26	9	15	0	1	25	42	41.67	70.00
5	B5790050	12	14	4	8	4.5	6	20.5	28	34.17	46.67
6	B5790067	23	22	6	15	1	7.5	30	44.5	50.00	74.17
7	B5790081	9	12	6	8	3.5	5.5	18.5	25.5	30.83	42.50
8	B5790098	22	26	14	18	6.5	7.5	42.5	51.5	70.83	85.83
9	B5790111	12	22	11	18	2.5	5	25.5	45	42.50	75.00
10	B5790128	17	21	11	15	5	5.5	33	41.5	55.00	69.17
11	B5790142	13	19	9	12	4.5	4.5	26.5	35.5	44.17	59.17
12	B5790166	24	27	14	17	8.5	9	46.5	53	77.50	88.33
13	B5790173	19	21	12	15	1.5	3	32.5	39	54.17	65.00
14	B5790197	12	16	10	12	1.5	5.5	23.5	33.5	39.17	55.83
15	B5790210	19	29	11	20	5	4.5	35	53.5	58.33	89.17
16	B5790227	18	23	9	9	2	5	29	37	48.33	61.67
17	B5790234	18	25	13	14	2.5	7	33.5	46	55.83	76.67
18	B5790241	25	23	14	17	7	9	46	49	76.67	81.67
19	B5790258	15	18	12	14	0.5	1.5	27.5	33.5	45.83	55.83
20	B5790296	20	21	14	16	6	5	40	42	66.67	70.00
21	B5790319	11	17	8	10	4	4	23	31	38.33	51.67
22	B5790333	20	24	15	15	6.5	6.5	41.5	45.5	69.17	75.83
23	B5790340	18	23	8	14	2	5.5	28	42.5	46.67	70.83
24	B5790357	20	25	10	16	1.5	8.5	31.5	49.5	52.50	82.50
25	B5790364	19	19	12	15	6.5	7	37.5	41	62.50	68.33
26	B5790371	16	19	9	13	1	4	26	36	43.33	60.00
27	B5790395	20	20	11	11	3.5	6	34.5	37	57.50	61.67
28	B5790401	16	18	6	13	3	4.5	25	35.5	41.67	59.17
29	B5790418	22	27	15	14	8	9	45	50	75.00	83.33
30	B5790425	23	28	11	15	4.5	7	38.5	50	64.17	83.33
31	B5790432	19	19	11	16	7.5	8.5	37.5	43.5	62.50	72.50
32	B5790449	17	19	14	14	4	8.5	35	41.5	58.33	69.17
33	B5790463	18	21	8	13	1	4.5	27	38.5	45.00	64.17
34	B5790487	16	20	8	11	2	4.5	26	35.5	43.33	59.17
35	B5790494	15	17	7	13	3	6.5	25	36.5	41.67	60.83
36	B5790517	20	16	5	9	6	5	31	30	51.67	50.00
37	B5790548	16	18	13	15	4	6	33	39	55.00	65.00
38	B5790555	10	13	10	12	2.5	3.5	22.5	28.5	37.50	47.50
39	B5790562	15	17	9	11	0.5	3	24.5	31	40.83	51.67
40	B5790593	11	16	4	10	1	3.5	16	29.5	26.67	49.17
41	B5790609	9	17	9	10	2	7	20	34	33.33	56.67

42	B5790623	16	17	6	12	2	4.5	24	33.5	40.00	55.83
43	B5790630	14	16	11	11	3.5	8.5	28.5	35.5	47.50	59.17
44	B5790647	12	13	8	8	6	6.5	26	27.5	43.33	45.83
45	B5790722	15	28	5	20	4.5	5	24.5	53	40.83	88.33
46	B5790746	17	27	14	11	1.5	6	32.5	44	54.17	73.33
47	B5790760	17	20	7	10	2.5	3	26.5	33	44.17	55.00
48	B5790777	18	19	14	16	4.5	7	36.5	42	60.83	70.00
49	B5790784	19	20	9	10	3	3	31	33	51.67	55.00
50	B5790791	15	15	8	10	3.5	4.5	26.5	29.5	44.17	49.17
51	B5790807	15	17	8	11	2.5	4	25.5	32	42.50	53.33
	MIN	9	12	4	8	0	1	16	25.5	26.67	42.50
	MAX	25	29	15	20	8.5	9	46.5	53.5	77.50	89.17
	MEAN	17.04	20.47	9.98	13.35	3.64	5.62	30.66	39.44	51.09	65.74
	SD	3.96	4.37	3.02	3.06	2.19	1.94	7.41	7.65	12.36	12.76
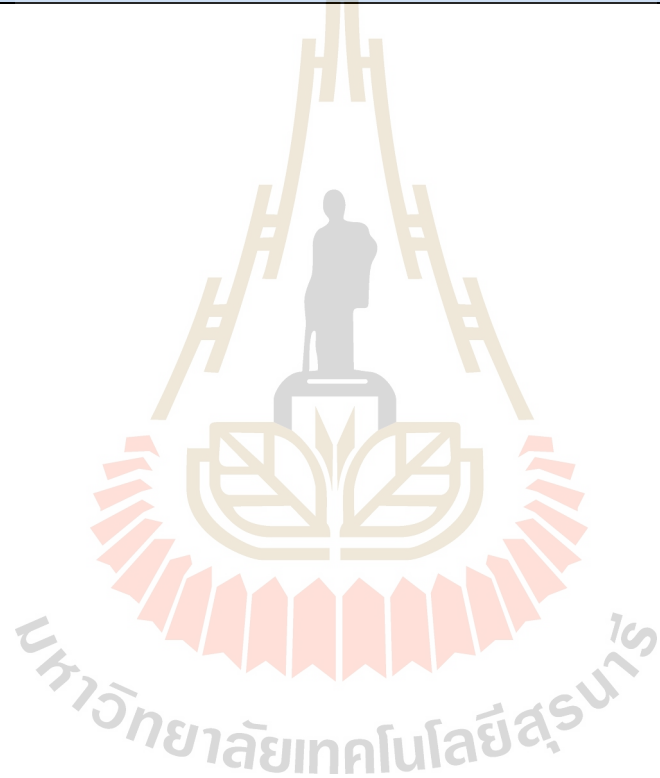
CURRICULUM VITAE

Kantapat Trinant was born in Ubon Ratchathani, Thailand. He received his Bachelor of Arts in Politics in 1997 from La Trobe University, Australia and Master of Arts in Language and Communication in 2004 from National Institute of Development Administration, Thailand. He is a full-time lecturer of English at the Faculty of Liberal Arts, Ubon Ratchathani University where he teaches fundamental English and academic English courses. His research interests include EFL/ESL and ESP learning and teaching as well as corpus-based language studies.