NOUN-NOUN COMBINATIONS IN TECHNICAL ENGLISH

Nguyen Mai Linh

A Thesis Submitted in Partial Fulfillment of the Requirements for

the Degree of Master of Arts in English Language Studies

Suranaree University of Technology

Academic Year 2010

คำนามผสมในภาษาอังกฤษเชิงเทคนิค

นางสาวเหงียน มาย ลิงค์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาศิลปศาสตรมหาบัณฑิต
สาขาภาษาอังกฤษศึกษา
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2553

เหงียน มาย ลิงค์ : คำนามผสมในภาษาอังกฤษเชิงเทคนิค (NOUN-NOUN COMBINATIONS IN TECHNICAL ENGLISH) อาจารย์ที่ปรึกษา : รองศาสตราจารย์ ดร. เจเรมี วอด, 111 หน้า

การสร้างนามวลีแบบนามประสมเป็นวิธีการหนึ่งในการย่อข้อมูลในภาษาอังกฤษเทคนิค โดยที่ความสัมพันธ์เชิงอรรถศาสตร์ระหว่างนามที่ใช้ในการประสมอาจสูญหายไปในกระบวนการดังกล่าว เพื่อที่จะตีความหมายของนามวลีแบบนามประสมนี้ จำเป็นต้องมีการนำความสัมพันธ์เชิงอรรถศาสตร์นี้กลับคืนมา การศึกษานี้มีวัตถุประสงค์หลักเพื่อศึกษารูปแบบพื้นฐานของความสัมพันธ์ทางอรรถศาสตร์ระหว่างนามในนามวลีแบบนามประสมในตำราภาษาอังกฤษด้านวิศวกรรมไฟฟ้า โดยอาศัยความรู้ด้านความสัมพันธ์เชิงอรรถศาสตร์นี้ ผู้เรียนสามารถฝึกให้เกิดความตระหนักและสามารถตีความหมาของนามวลีแบบนามประสมในตำราเรียนได้ การศึกษานี้ประกอบไปด้วยหลายขั้นตอน กล่าวคือ การรวบรวมนามวลีแบบนามประสมจากคลังข้อมูลตำราด้านวิศวกรรมไฟฟ้า และนำตัวอย่างของนามวลีแบบนามประสมจำนวน 370 ตัวอย่างมาวิเคราะห์หาความสัมพันธ์เชิงอรรถศาสตร์ เพื่อสร้างความสมดุลระหว่างการนำมาใช้ในการสอนและการศึกษาโดยครอบคลุม ผู้วิจัยได้จำแนกความสัมพันธ์ออกเป็น 8 ประเภท ได้แก่ ความสัมพันธ์ด้านตำแหน่ง วัตถุประสงค์ การวัด การแทน แหล่งที่มา การปฏิบัติการ เป้าหมาย และโครงสร้าง (เรียงลำดับความถี่ของการปรากฏจากมากไปหาน้อย) ความสัมพันธ์เหล่านี้ครอบคลุมมากกว่า 80% ของนามวลีแบบนามประสมในด้านจำนวนนับและมากกว่า 75% ในด้านประเภทของความสัมพันธ์ ทั้งนี้ ผู้วิจัยได้เสนอแบบฝึกจำนวน 14 แบบสำหรับใช้เป็นกิจกรรมและแบบฝึกหัดในการสอน แบบฝึกหัดเหล่านี้แบ่งออกเป็นระดับต่าง ๆ ตั้งแต่ง่ายถึงยาก และมีรูปแบบต่าง ๆ ตั้งแต่ แบบฝึกหักจับคู่ คำถามแบบปรนัย และคำถามสำหรับตอบแบบสั้น ๆ

สาขาวิชาภาษาอังกฤษ ลายมือชื่อนักศึกษา_______________________

ปีการศึกษา 2553 ลายมือชื่ออาจารย์ที่ปรึกษา________________

NGUYEN MAI LINH : NOUN-NOUN COMBINATIONS IN TECHNICAL ENGLISH. THESIS ADVISOR : ASSOC. PROF. JEREMY WARD, Ph.D., 111 PP.

NOUN-NOUN COMBINATIONS/ TECHNICAL ENGLISH

The formation of noun-noun combinations (NNCs) is a means to compress information in technical English and the semantic relation among nouns in the combination may be lost in this compression process. To interpret the NNC, the relation lost needs to be retrieved. The initial objective of the study is to investigate the semantic relationship of NNCs in English textbooks in the field of electrical engineering in order to find out a list of common relations. With the knowledge of such relations, the student could be trained to recognize and interpret NNCs in their textbooks. Several steps are carried out to reach that goal. A list of NNCs was extracted from an electrical engineering textbook corpus, and a sample of 370 combinations was taken from the list for the analysis of semantic relations. Due to the effort to balance between teachability and the coverage of the relation classification, 8 relation categories were drawn, including location, purpose, measure, representation, source, operation, objective, and structure (in order of commonness from the most to the least). These relations covered over 80% of the NNCs in terms of tokens and over 75% in terms of types. Fourteen tasks which could be used as teaching activities and exercises are introduced. These tasks are classified into different levels from the least to the most difficult and presented in different formats from matching, multiple choices to short answers.

School of English Student's Signature__________________

Academic Year 2010 Advisor's Signature__________________

ACKNOWLEDGEMENTS

This thesis would not have been made possible without the help of many kind people around me. The space is not enough to mention all of them here, and there is no word which could express their kind help and support for me. First of all, I owe my deepest gratitude to my supervisor, Associate Professor Dr. Jeremy Ward, without whom this thesis would not have been done. His encouragement, guidelines, suggestions and patience have been a great help for me to overcome difficult times to conduct and complete the study. I would also like to thank my committee, Dr. Jitpanat Suwanthep, Dr. Suksan Suppasetseree, and Dr. Butsakorn Yodkamlue, for their valuable suggestions.

I am also indebted to my many teachers and staffs from the School of English, Associate Professor Dr. Anchalee Wannaruk - the current chair, Assistant Professor Dr. Siriluck Usaha-the former chair, Dr. Sirinthorn Seepho – the MA coordinator, the school secretaries and other teachers who cannot mention all here, for their support in both personal and academic issues.

A great support came from the Faculty of Information Technology, Thai Nguyen University. I would like to thank the faculty for making it possible for me to have a chance to study here at the School of English, Suranaree University of Technology.

As an international student, it was hard for me to start my study in a new environment. I am very grateful to all of my classmates and my friends for their emotional support.

Finally, I would like to show my gratitude to my family, my parents and my brother, for staying beside me, for always being someone I could lean on in any situation.

Nguyen Mai Linh

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

LIST OF TABLES

Table **Page**

CHAPTER 1

INTRODUCTION

The Faculty of Information Technology, Thai Nguyen University, where the researcher works, offers two undergraduate programs for students majoring in Automatic Control and Electronic Telecommunications, two sub-fields of electrical and electronic engineering. Engineering students from both sub-disciplines are required to study fundamental electrical courses. Reading English textbooks and other English supplementary materials are recommended. Technical texts are characterized by high frequency of technical vocabulary which is considered difficult to ESL/EFL learners (e.g. Anwar, 1994; Chung & Nation, 2003; Farrell, 1990). To help engineering students in terms of technical English, ESP courses (mainly focus on reading) are offered by the Faculty as core courses. However, as the common situation in ESP teaching and learning in Vietnam (e.g. Nguyen, 2006), these courses are inefficient. Based on the difficulty that technical vocabulary in general and noun-noun combinations in particular may bring about for engineering students and due to the failure of ESP courses, the current study is expected to make a small contribution to improve the reading performance of the two mentioned above majors at the Faculty of Information Technology by raising their awareness of such class of vocabulary and providing them with a tool to be trained to read more efficiently.

1.1 Rationale

1.1.1 Importance of vocabulary in reading comprehension

″Without grammar very little can be conveyed; without vocabulary nothing can be conveyed″ (Wilkins 1972:111). This quotation has shown how important vocabulary is. The beauty of a language lies in its vocabulary; the information one utterance conveys is mostly in the words. However, the history of vocabulary teaching and researching is full of ups and downs. Vocabulary had a priority in Roman time, but its position was taken by grammar for centuries since medieval period (Schmitt, 2000). Recently, vocabulary has gained back its favor with plenty of publications such as *The lexical approach: the state of ELT and a way forward* (Lewis, 2002), *The lexical syllabus: a new approach to language teaching* (Willis, 1990), *Implementing the lexical approach: putting theory into practice* (Lewis & Gough, 2002). Vocabulary, especially technical vocabulary, plays even a more crucial role in specialized texts, because this class of vocabulary is the key to a specific area where mastery of its special vocabulary shows that one has knowledge in that subject matter.

The academic world is developing faster and faster; the world powers both in economy and technology are English speaking countries. Most of world knowledge is daily accessed via written English, which makes reading comprehension become a very important skill in every academic field, especially in technology which keeps changing day by day. Reading comprehension and vocabulary have a high correlation (Anderson & Freebody, 1981; Carroll, 1993; Koda, 1989; Laufer, 1992; Laufer & Sim, 1985; Mezynski, 1983; Qian, 1999; Zhang & Anual, 2008,). Vocabulary knowledge is a good predictor whether one is a good reader or not, which means if one gets a high score in a vocabulary test, he or she has a high possibility of being a good reader. In L1 reading, three kinds of relationship have been discussed: Instrumentalist hypothesis, aptitude hypothesis, and knowledge hypothesis (Anderson & Freebody, 1981; Hiebert & Kamil, 2005). Instrumentalist view can be explained in a very straightforward way: if learners has a good knowledge of vocabulary (i.e. getting a high score in a vocabulary test), they can have a good understanding of the text. In aptitude point of view, good vocabulary knowledge means good verbal ability which results in good comprehension. Looking at vocabulary – comprehension relationship from the angle of knowledge hypothesis, good comprehension is the results of good knowledge of culture which is obtained from good vocabulary knowledge.

In L2 reading, building vocabulary knowledge is a prerequisite. If in L1 reading, when learners start to read, they already build a vocabulary of some size which can facilitate them to comprehend the text at an appropriate difficulty level, in L2 reading, learners start to read at the same time or even earlier than other skills as well as vocabulary building. No comprehension can be achieved without knowing any word; no reading strategies can be applied; no word can be inferred from the context. Sufficient vocabulary knowledge not only results in an acceptable comprehension, but also brings about incidental vocabulary learning (Koda, 2005). Lack of vocabulary could interfere the process of reading. Imagine we are reading a text in a language we just begin to learn, the text is full of unknown words. We stop frequently to look up the dictionary, mostly a bilingual one, which firstly demotivates our reading by making us feel frustrated. Secondly, even if we get some kind of comprehension, it is an incomplete picture assembled from L1 translation, which loses L2 nuances.

Studies in L2 reading also show that vocabulary knowledge is a good predictor of comprehension (Koda, 1989; Laufer, 1992; Qian, 1999). Vocabulary knowledge even surpasses syntax (Laufer & Sim, 1985; Ulijn & Strother, 1990) and background knowledge (Laufer & Sim, 1985) in terms of the importance to reading comprehension.

Different aspects of this relationship have been explored but no researchers can refute the conclusion by Anderson and Freebody (1981: 111): "It is also clear that word knowledge is a requisite for reading comprehension: people who do not know the meaning of very many words are most probably poor readers."

1.1.2 Vocabulary learning and multi-word items

As widely mentioned in lexical studies, lexical items are the building block of language, which was the grammar or structure in traditional view (Lewis, 2002; Willis, 1990). However, lexical items do not just mean individual words, single words, which are recognized by spaces, but means chunks, patterns, and collocations. This results in the fact that multi-word items have gradually played a more and more important role in comparison with single-word items in vocabulary learning in particular, and in second language learning in general. A multi-word item in Moon's (1997) definition is "a vocabulary item which consists of a sequence of two or more words (a word being simply an orthographical unit). This sequence of words semantically and/or syntactically forms a meaningful and inseparable unit" (pp. 43). These multi-word items can be compounds (e.g. *bandpass filter, Prime Minister*), phrasal verbs (e.g. *give up, put out, break in*), idioms (e.g. *rain cats and dogs, the early bird catches the worm*), fixed phrases (e.g. *of course, in fact*), and prefabs or lexical bundles (e.g. Biber & Barbieri, 2007; Hyland, 2008) (e.g. *...what we are going* *to do…-in spoken university register, it can be seen that – in academic writing,*). Among them, “compounds typically denote and have high information content – often because they are technical terms or having specific reference” (Moon, 1997:56). Compounds are good linguistic devices to convey highly compact information, so they become technical terms. Whatever the cause is, it is widely agreed that compounds, especially noun compounds, reflect the genre of technical writing and overwhelmingly occur in specialized texts (Cohen et al., 1988; Beardon and Turner 1993, Ter Stal 1994, cited in Lauer, 1995; Norman, 2003; Pueyo & Val, 1996; Salager-Meyer, 1984; Ward, 2007; Wasuntarasophit, 2008). One point which should be noted is that in different studies, different researchers may use different terms such as noun phrases, noun compounds, nominal groups, complex nominal, compound nominal but they all refer with slight differences to one thing - noun + noun combinations.

1.1.3 Noun-noun combinations in technical writing

Noun-noun combinations occur with high frequency in technical writing. A large number of noun phrases occurs in textbooks of natural and social science subjects, including genetics, biology, political science, and history (Cohen et al. 1988). In addition to science textbooks, engineering textbooks also show up numerous nominal groups (Ward, 2007; Wasuntarasophit, 2008). A similar phenomenon occurs in the text in the plastic field (Pueyo & Val, 1996). Moreover, it is pointed out that a large number of these noun compounds are the product of the nominalization process, one type of grammatical metaphor. Grammatical metaphor is explained below.

In the point of view of Systemic-Functional Grammar, language is a semiotic system including three components: discourse-semantics, lexicogrammar, and phonology, in which phonology realizes lexicogrammar that, in turns, realizes discourse-semantics (Halliday, 1985/1994, cited in Saenz, 2000). Some grammatical form is typically used to realize a semantic choice; for example, processes are expressed by verbs, participants/entities by nouns, attributes by adjectives. In other words, expressing processes, participants, attributes are unmarked function of verbs, nouns, adjectives, respectively (Banks, 2005). The following is Halliday's (2004: 87) summary of typical functions of word classes:

1. Processes (actions, events, mental processes, relations) are expressed by verbs;
2. Participants (people, animals, concrete and abstract objects that take part in processes) are expressed by nouns;
3. Circumstances (time, place, manner, cause, condition) are expressed by adverbs, and by prepositional phrases;
4. Relations between one process and another are expressed by conjunctions.

When this congruence does not take place, there occurs grammatical metaphor, or grammatical metaphor is a variation of the typically grammatical realization of semantic choice. The following example is provided by Downing (1991, cited in Sasenz, 2000):

We walked in the evening along the river to Henley → *Our evening walk along the river took us to Henley.*

The first sentence is typical grammatical wording. The congruence of word classes and their functions cannot be seen in the second sentence. The process *walk* is now expressed by the noun; the two circumstances *in the evening* and *along* *the river* now become classifier and qualifier expressed in the form of a noun and a prepositional phrase.

Types of grammatical metaphor are differently classified by different researchers. Nineteen types of which 5 are the cases of nominalization are listed by Ravelli (1988); thirteen types, of which nominalization accounts for 3, are classified in Halliday's (1998) list (cited in Banks, 2001). It can be seen that in both classifications, nominalization plays a central role in grammatical metaphor. Nominalization has also been exhaustively discussed in studies of scientific writing (e.g. Halliday 1993, 2004)

Nominalization is the formation of noun phrases to express processes and quality instead of verbs and adjectives (Bank, 2005), for example: *to modulate signal* → *signal modulation*, m*etal is conductive* → *metal conductivity.* A clause can also be nominalized to become a noun phrase like:

The temperature increases sharply → *a sharp increase in temperature* (example in Pueyo & Val, 1996)

The reason for the wide use of nominalizations in scientific writing is two features of scientific subject-matter, the structure of scientific argument and structure of scientific knowledge (Halliday, 1993). Firstly, scientific reasoning often occurs in a sequence; nominalization allows packaging the process in the previous sentence into a nominal group functioning as a theme in the next sentence to provide given information which will lead to the new information as can be seen in the following example:

Recombinant falcipain rapidly hydrolyzed both denatured and native hemoglobin. Hemoglobin hydrolysis was blocked by cysteine protease inhibitors... (Norman, 2003)

The information, the process in the first sentence is nominalized as *hemoglobin hydrolysis* which functions as theme in the second sentence and what is new here is encoded in *blocked by cysteine protease inhibitors.*

The second lies in the structure of scientific knowledge. The world is changing, but it is easier to begin to feature it if we imagine it staying still. Nominalization allows scientists to give the image of a static world so that they can classify, analyze, and create relations among things. In the words of Norman (2003), he sees this function of nominalization from the semantic point of view; it has the "reification" effect, which "makes nominalization particularly suitable for scientific writing, by giving it an appearance of solidity, stability, and fixed factuality" (pp. 350). Reviewing previous studies about nominalizations, Pueyo and Val (1996) adds that nominalization allows packaging of a complex phenomenon in a few words. For example, *phase modulation* is the compression of the process of sending information by modifying (modulating) the difference in phase (fraction of a wave length) between a signal and a reference.

For these reasons, nominalizations are widely preferred by technical writers. Nominalizations could be a linguistic device to form a good deal of noun-noun combinations (NNCs) which are preferable in technical texts. However, these combinations are a source of difficulty in reading comprehension (Cohen et al., 1988; Pueyo & Val, 1996).

1.1.4 Difficulties in noun-noun combination interpretation

There are several reasons for such difficulty in interpretation of noun-noun combinations or noun compounds. Firstly, “the problem of noun compounds is that as compounding is a highly productive process, any individual compound may not figure in any particular lexicon, so its meaning has to be constructed by readers/hearers” (Jones, 1983:1), for example *bus voltage* could be interpreted as voltage of batteries used for buses if the readers/hearers are reading/listening a paper about bus manufacture, or it could be interpreted as voltage in buses in a computer if one is reading about computer architecture. Compounding is highly productive; at any time a new compound can be formed to fit in the context (Bauer, 1983). One example in Weiskopf (2007) is *watermelon place*, which in one context could be interpreted as the place is in front of a watermelon as opposed to other places in front of other fruits at a dinner table or could be interpreted as a place which a water melon is put in another context.

The second lies in the degree of lexicalization or opacity of noun-noun combinations. The following opacity scale taken from Levi (1978):

Productive forms

Transparent	*mountain village, family reunion, lemon peel…*
Partly opaque	*grammar school, brief case, polar bear…*
Excocentric	*birdbrain, razorback, ladybird, blockhead…*
Partly idiomatic	*polka dot, monkey wrench, flea market…*
Completely idiomatic	*honeymoon, fiddlesticks, duck soup …*

Idiomatic forms

Depending on the degree of opacity of the noun-noun combination, we can retrieve the meaning of the compound from the meanings of its constituents. Take "random signal" as an example.

Random

1. lacking any definite plan or prearranged order; haphazard

2. (Mathematics & Measurements / Statistics) *Statistics*

a. having a value which cannot be determined but only described probabilistically

b. chosen without regard to any characteristics of the individual members of the population so that each has an equal chance of being selected

Signal

1. any sign, gesture, token, etc., that serves to communicate information

2. anything that acts as an incitement to action

3. (Electronics & Computer Science / Telecommunications)

a. a variable parameter, such as a current or electromagnetic wave, by which information is conveyed through an electronic circuit, communications system, etc.

b. the information so conveyed

c. (*as modifier*) signal strength a signal generator

Based on the meanings of *random* and *signal* from the dictionary, *random signal* can be interpreted as an electromagnetic wave selected by chance, but the fact is that "*random signal*" means "waveforms having at least one parameter (usually amplitude) that is a random function of time" (Graf, 1999). The relationship of *random, signal,* and *random signal* in the example can be seen as partially opaque; the meaning of *signal* in the combination can be referred from the meaning of *signal* as an individual word, but *random* cannot be. The other case of opacity is totally opaque. Take another example of *control center.* Suppose we know *control* – regulation or operation, *center* – a place, then *control center*- a place for regulation. If we guess so, we are wrong. *Control center* is actually a preamplifier; the full definition can be found in Graf (1999): "a switching, amplification, and equalization component designed to select input signals, amplify them by amounts from 0 to 60 dB, and deliver an output voltage compatible with input requirements of a power amplifier" (p. 150), which gives us a much more specific referent of *control center*, a typical feature of technical vocabulary.

However, the biggest difficulty of noun-noun combination interpretation does not lie in these idiomatic compounds, because with each such combination we can learn one at a time (Weiskofp, 2007). Moreover, this type of combination does not play a large part in the language, especially in technical language.

The third source of difficulty concerns the differences in the structure of different languages which could affect L2 readers. Cross-linguistic research provides one more difficulty in interpreting noun-noun combination for EFL/ESL learners. Different languages are strictly or more freely based on some word order and these word order patterns may more or less affect the reading/listening process of EFL/ESL learners (e.g. Flynn, 1989; Rutherford, 1989; Gass, 1989), in this case word order in phrases. At word-level of reading skills, after words are recognized (decoding the lexicographical form of words), the reader integrates them into larger units like phrases, clauses (Fender, 2001). The difference in word order patterns could influence this process. For example, an English noun-noun combination is structured by modifier + head, then the later word determines the category and the former distinguishes it from others in the same category. However, the process is reversed in such languages as Thai and Vietnamese where the structure is head + modifier. For example, *input signal* become *tín hiệu đầu vào* in Vietnamese, in which *input* is equivalent to *đầu vào,* and *signal* to *tín hiệu.* The cross-linguistic studies also show that the difference in the degree of explicitness of the noun-noun relation could also hinder the full comprehension of the combination. Pastor (2008) points out that in Spanish, complex noun phrases are formed and linked together by prepositions to show the relationship, but it is often not the case in English complex noun phrases. This fact causes difficulty for Spanish-speaking learners in interpreting complex noun phrases.

The last source of difficulty is implicitness of semantic relations between two constituent nouns. Such nominal groups as *GM car, woman doctor, diesel engine, voltage source,* products of nominalization process, often cause difficulty in interpretation. To interpret *GM car*, a general knowledge that GM is an automobile manufacturer is necessary; similarly, a knowledge that diesel is one kind of fuel is essential to correctly interpret *diesel engine.* It can be seen that there are two different semantic relations existing between the two nouns in the nominals. The underlying relation in the former nominal is "*manufactured by*", while the semantic relation in the latter one is "*powered by.*" These semantic relations are lost during the packaging process, a common phenomenon in technical English. This relation needs to be recovered for a full interpretation during reading. Even such cases where this kind of knowledge is not required, ambiguity still exists as in the cases of *woman* *doctor, voltage source. Woman doctor* is a doctor for woman or a doctor who is a woman. *Voltage source* is a voltage for the source or voltage from the source.

Among all the reasons, implicit semantic relations may lie at the center. The productivity of compounding leads to the fact that hearers/readers have to interpret compounds without reference to any lexicon. However, the semantic relation is not explicit, which causes difficulty in their interpretation.

1.1.5 Taxonomies of semantic relations

The first issue is whether or not semantic relations can virtually be classified into taxonomy. There are arguments about these attempts to establish such a taxonomy. The example taken is a phrase from British newspapers *canoe wife* which cannot be interpreted by the linguistic surface of the compound. To be able to understand it, we need to know the underlying story that there is a woman telling a lie that her husband died in a canoe accident to get the insurance. The case seems to be common in newspaper headlines for the sake of saving space, and these types of compound as defined by Bauer (1983) as nonce formations, "a nonce formation can be defined as a new complex word coined by a speaker/writer on the spur of moment to cover some immediate need" (pp. 45). He also adds "one point that is very characteristic of some kinds of nonce formations considered in isolation is their potential ambiguity" (pp. 46), which can be seen in *canoe wife.* It can be argued that as many nonce-formations, *canoe wife* merely occurs in that moment and its meaning can be specified by the particular situation. Moreover, *canoe wife* has a very particular referent, the wife in that story, not anyone else; it is neither productive nor lexicalized. However, this type of combining is not the case of noun-noun combination in neither technical texts nor the one discussed by cognitive scientists in conceptual combination. It is necessary to classify different semantic relations in nominal groups to facilitate the process and the correctness of their interpretation when readers are aware of them, especially L2 readers who are not familiar with this kind of concept modification.

The fact is that linguists have been attempting to classify such relations on different bases from transformational to functional points of views based on syntactic and semantic nature of noun compounds. In the earlier stage, from the transformational grammar point of view, a noun compound is the surface structure of an underlying sentential form (Lees, 1960). For example, the relation of *subject-predicate* can be found in the noun compound *girl friend (the friend is a girl)*, or *subject-object* in the compound *car thief (the thief [stole] the car)*. Under the frame of generative semantic theory, noun compounds are the products of the nominalization (e.g. *signal detection – […] detect the signal*) or deletion process of relative clauses or prepositional phrases (*drug death – the death which is caused by drug*) (Levi, 1978). Even though this analysis is closer to functional approach, it is still primarily syntactic. This is because of the fact that "generative semantics, the name notwithstanding, is heavily concerned with the form and structure of language and the structural transformations which operate on sentences" (Finin, 1980:33). From a semantic view, the underlying relationships could be classified into such categories as *location* (e.g. *anode region – the region located at the anode*), *whole - part (*e.g. *table leg – the leg is a part of the table)*, *time* (e.g *Sunday newspaper – the newspaper comes out on Sundays*) (e.g. Downing, 1977; Warren, 1978; Halliday, 1985; Girju et al., 2005). Based on the assumption that noun compounds can be paraphrased into clauses or phrases, lists of prepositions are proposed for compound noun interpretation such as *at* (e.g. *airport food – food at the airport*), *in* (e.g. *morning talk – talk in the morning*), *from* (e.g. *reactor waste – waste from the reactor*). Different taxonomies are suggested on the basis of different aspects of noun compounds, and there is also no agreement on the definiteness of these taxonomies. Some researchers postulate that the list of such relations is limited (Finin, 1980; Lauer, 1995; Levi, 1978; Warren, 1978); some argue for its unconstrained nature (Downing, 1977). For the standpoint of a complete list of relations, there is no consensus on the length of such list, from 9 verbs and prepositions (Levi, 1978), 8 prepositions (Lauer, 1995), a list of 12 relations (Warren, 1978), 8 categories of noun compound (Master, 2003) to potentially unlimited list from the nominal role viewpoint (Finin, 1980). Whatever the length of the list is, the shorter taxonomy provided to cover a large number of noun compounds, the more abstract it is. Take Levi's (1978) taxonomy as an example. The taxonomy includes high abstract verbs such as *have, make*. *Mountain top* is classified in *have* category (*the top that a mountain has*); similarly, *vegetable soup* is interpreted as *the soup that has vegetable in it*. However, the former one refers to location, and the latter one indicates composition. The dilemma is that details are sacrificed in a short taxonomy, and a detailed taxonomy results in an unlimited list of relations. With a pedagogical purpose in mind, a helpful taxonomy is characterized by its specificity and based on semantic relations between two constituents.

In short, noun-noun combinations are pervasive in technical texts. The process of forming these combinations is productive; some combinations are idiomatic. Noun-noun combinations have different word order patterns regarding to modifiers and heads in different languages. The underlying semantic relationship of the modifier and the head noun is implicit. All these facts make it difficult to engineering students in reading technical texts. However, playing the central role is implicitness of semantic relationship. Idiomatic combinations could be learned at one at one time and they do not account for a large number in specialized texts. Differences in word order patterns could be taught to students for them to be able to recognize the structure of English noun-noun combinations. The difficulty caused by productivity lies in the implicit relation. Because the relation is implicit, novel combinations formed due to productivity are more difficult to interpret. Although there have been studies about the underlying relationship, these studies are mostly in general English and not for application in teaching. There raises a need for a taxonomy of semantic relations applicable in teaching technical English.

1.2 Purpose of the study

In this study the semantic relations of noun-noun combinations in the corpus of electrical engineering textbooks built by Wasuntarasophit (2008) were investigated. The purpose of the study was to identify the common semantic relations of technical noun-noun combinations, from which exercises to train engineering students to interpret these combinations were suggested.

1.3 Research question

To realize the purpose of the study, the research question was raised:

- What are the most common semantic relations in technical noun-noun combination?

1.4 Significance of the study

Filling the gap in the literature

This study is conducted based on two assumptions which are (1) noun-noun combinations are ubiquitous in specialized texts and (2) underlying semantic relations of two constituent nouns are implicit, which causes difficulty for L2 learner in interpreting such combinations. Although such relations are intensively studied in general texts, little research has been done on technical texts, and on electrical engineering texts in particular as the case of this study. Moreover, the previous studies on this field are mostly pure linguistics or computational linguistics without implications for teaching. This research attempts to fill that gap.

Implication for teaching

This study is expected to benefit electrical background-needed engineering students' L2 reading comprehension as well as other technical fields. As has been noted, the ubiquity of nominal compounds and ambiguity of their interpretation could bring about difficulties for L2 learners. This study will raise students' awareness of these semantic relations. Especially, exercises will be suggested to help teachers to train students to recognize the relation. The number of nominal compounds is higher and higher in more and more specialized texts. Being able to recognize the implicit semantic relation is more important. The significance of this study is not limited in the field of electrical engineering and related areas; using exercises to train students' recognition can be applied in other engineering areas.

1.5 Definitions of terms

This section concerns two issues. The first issue is the distinction between noun compounds and noun phrases of which the argument is that there is no well-established distinction between these two groups. The second issue is related terms used in the literature.

First, regarding the criteria to distinguish noun compounds and noun phrases: it should be clearly stated that there will be no distinction between noun compounds and noun phrases applied in this study for the criteria to distinguish between compounds and noun phrases are still controversial. Some view a compound as "a lexical unit" which meets some criteria regarding to five aspects: orthographic, phonological, morphological, semantic, and syntactical (Bauer, 2006:719).

In terms of *orthographic* criteria, a compound is written as a single word, but this rule seems to be superficial. Bauer (2006) takes the examples of *rainforest, rain-forest,* and *rain forest* which could be found in standard dictionaries as an evidence to show that this criterion is not sufficient. In a more serious case as in "*a New York–Los Angeles flight,*" *York – Los,* if this criterion is applied, is a compound.

Stress patterns are considered as *phonological* criteria. As a compound, the preceding constituent or the premodifier is often assigned the primary stress (Lees, 1960). *Church-warden* is, however, not fixed to any stress pattern.

In *morphological* aspects, compounds do not allow inflection of the modifying word such as *trouser braces, scissor lift,* even though *trousers* and *scissors* are always plural as individual words. Although this criterion seems to be true to many instances, there are still exceptions like *arms race, suggestions box.*

When *syntactic* criteria are applied, compounds are syntactically treated as a single word, which could be tested by anaphora. Examples are provided by Bauer (2006) as follows*: I thought this house had aluminum windows, not wooden ones*, in which *ones* are replaced for *windows* in *aluminum windows*, a syntactic construction. Differently, in the example of *I installed a combination lock and now I can't remember it*, the object pronoun *it* is used to replace *combination lock*, not only lock. However, the rule could still be broken on occasions as in *I want to give myself a headache by banging it on the floor!*

Regarding to *semantic* criteria, compounds are at some degree of lexicalization. For example, *push-chair* and *wheel-chair* are denoted as two different entities even though *push-chair* has wheels, and *wheel-chair* can be pushed. Bauer (2006) argues against this point of view on two bases. Firstly, such sentence as *how do you do?* is lexicalized or idiomatized, but it is still a sentence. Therefore, *push-chair* and *wheel-chair* could still be syntactic phrases as they were newly constructed. Secondly, specialization in the meaning such as in the cases of *push-chair* and *wheel-chair* is not only the product of lexicalization, but it could be the meaning selected when it was first formed. Downing (1977) takes the example of *apple-juice seat* with the meaning of a seat of which a glass of apple juice is placed, but not several other possible meanings such as a seat which is splashed by apple juice.

Bauer's (2006) arguments about the distinction between compounds and phrases are in the same line with Downing (1977). She casts a doubt on the well-established cut-off between compounds and nominal phrases. The study will adopt Bauer's and Downing's point of view about the division between the two groups; "[…] the criteria to which reference is generally made do not allow us to distinguish between a class of noun + noun compounds and a class of noun + noun syntactic constructions" (Bauer, 1998:65).

The second point relates to related terms in the literature. One more point needs to be noted is that there are different terms used through the literature of noun compound research like complex nominals (Levi, 1978), noun compounds (Lauer, 1995), nominal compounds (Finin, 1980), compound nouns or noun + noun compounds (Downing, 1977), noun-noun combinations (Gagné, 2001, 2002; Gagné & Shoben 1997; Maguire et al., 2008). The different terms used reflect the slightly diversion of the working definitions. Downing (1977) adopts Li's (1971) definition (cited in Downing, 1977) which defines an N+N compound as a concatenation of two or more nouns with nominal functions. The same definition is used in Finin (1980); he further lists some related terms like complex nominals, nominal phrases, and noun-noun modification. Levi (1978) includes in her study of complex nominals non-predicate adjectives such as *electrical* in *electrical engineering.* Lauer (1995) defines a noun compound as "any consecutive sequence of nouns at least two words in length that functions as a noun, but which contains no genitive markers and is not a name" (pp. 31). It seems more restricted than the previous definitions. He excludes from his study such compounds such as *dog's breakfast, firetruck,* but includes gerund forms like *laughing children, horse riding*, and there is no consideration of conversion as in the case of *plastic encapsulation* (*plastic* can either be a noun (*encapsulation* made of *plastic*) or an adjective (*encapsulation* is elastic)). In this study, noun-noun compounds are a string of two or more nouns which has the function of a noun. Despite the fact that the string of nouns can be very long 4, 5 or more nouns, the underlying semantic relations are similar to the case of two noun combination. This study will analyze only compounds of two nouns; the results could be applicable to longer compounds. Throughout the study, all terms such as noun compounds, compound nouns, noun-noun compound, and nominal compounds may be encountered, but they all refer to the concept of noun-noun combinations as explained above.

This chapter provides a rationale for the study which is the basis to draw up the research question. The next chapter will review related studies in the literature regarding to the ubiquity of noun-noun combination in technical English, properties of noun-noun combination which cause difficult to L2 learners, taxonomies of semantic relations and exercises designed for noun-noun combination interpretation.

CHAPTER 2

LITERATURE REVIEW

2.1 Ubiquity of noun-noun combinations

The frequent occurrence of noun-noun combinations has been proved through lots of studies in literature of English for specific purposes in different fields. Noun-noun combinations have been proved to be ubiquitous and played different roles in technical and scientific texts.

Salager-Meyer (1984) investigated the complex nominal phrases (CNPs) in English scientific-technical texts in comparison with general English in terms of frequency of occurrence, average length and proportion. Complex nominal phrases in ten 2,000 - word extracts of medical text (ME) and ten 2,000 - word extracts of general English (GE) were counted and recorded. The results of this study were then compared with her previous studies of technical English. Salager-Meyer (1984) reported the following results:

Table 2.1 Complex nominal phrases in Technical English, Medical English, and General English

	Number of compounds	Number of words involved in the compounds	Average length	Percentage of compound words in the text
TE (technical English)	1,179	3,073	2.61	15.37
ME	751	1,953	2.55	9.76
GE	69	173	2.51	0.87

A T-test was carried out and the result showed that the frequency of occurrence of CNPs in ME was significantly higher than that in GE ($p < 0.05$). This frequency in TE was even higher than in ME. The CNP average length was similar among texts but their distribution was different. The number of long CNPs (more than 2 words) was much higher in ME than in GE. Looking in more detail in each extracts, the researcher found out that the more specialized texts were, the longer CNPs were.

Cohen et al. (1988) studied the difficulties the student encountered in reading for specialized purposes. They used the texts from 4 areas including genetics, biology, political science, and history, and information about difficulties of the texts was provided by the students' reading comprehension. The study found out that heavy noun phrases (noun phrases are not necessarily long and complex but are difficult to process) functioning as subjects of main clauses, and subjects of subordinate clauses, or objects, were problematic for students.

In ter Stal's (1994, cited in Lauer, 1995) study of 293 technical abstracts from *Engineered Materials Abstracts,* he reported 3514 different noun compounds, which resulted in 12 compounds per abstract in average.

Beardon and Turner (1993, cited in Lauer, 1995) took a different approach and reported even a higher number of noun compounds in Computer Graphics abstracts. 27% of all words in six abstracts of their sample were found to occur in complex nominals.

Pueyo and Val (1996) pointed out that nominalization is a common process of forming technicality in the field of plastics, which resulted in a large number of such nominal groups as *reaction injection moulding, increasing blowing pressure, high resistance polystyrene.* These nominal groups compress information into a part of a clause, and are the sources of reading difficulties so that the authors suggested that teachers help students to practice unpacking the information in the nominal groups for an easier understanding.

Norman (2003) analyzed anaphoric references in research article abstracts in the biomedical field. The results showed that nominalizations (packaging devices in Norman's wording) functioning as an anaphoric device occur very frequently. Complex nominal groups (multi – lexeme nominal groups) as the products of the nominalization process account for 11.86% of anaphors. Norman (2003) provided examples of nominal groups functioning as anaphoric devices in his study of biomedical research abstract samples:

> *Recombinant falcipain rapidly hydrolyzed both denatured and native hemoglobin. Hemoglobin hydrolysis was blocked by cysteine protease inhibitors...*
>
> *To further evaluate the role of falcipain, we expressed the enzyme in bacterial and viral expression systems. [...] Recombinant falcipain rapidly hydrolyzed…*

Ward (2007) analysed corpora of chemical, industrial, civil, mechanical, and electrical engineering textbooks. One of his findings was that the occurrence of complex noun phrases is ubiquitous and highly discipline – specific. Take *system* in his corpora as an example. *System* is widely distributed throughout all 5 subcopora with the PEAKRATIO (PR- the ratio of the mean occurrence and the maximum occurrence of one word throughout all disciplines) of 2.11 (lower PR value, wider range of distribution and vice versa); however, such nominal groups as *real systems, system equation* have PR values of 4.36 and 3.48, which means a narrow distribution of these nominal groups. His analysis also showed that more than often the technical terms combine with other words to form a combination. For example, *reaction* occurrs 2311 times in the corpus, and there are 1538 times it involves in a combination, which accounts 67% of *reaction* occurrences.

Wasuntarasophit (2008) compiled a corpus of 122, 209 running words from five electrical engineering textbooks in four sub-fields of electrical engineering including Control Systems, Power Systems, Electronics, and Communications. All noun phrases (which are termed nominal compounds in this study) were extracted from the corpus, which resulted in a number of 6,043 different noun phrases (56.31% of types – different word forms) with the occurrence of 10,707 (9.96% of tokens – number of occurrences), and the number of words involving in the noun phrases accounting for 25,429 running words (20.80%) of the corpus. The type-token ratio of noun phrases is very high, which means that lots of these noun phrases occur only once or twice throughout the corpus and lead to a high learning load. Most of these noun phrases are technical noun phrases (5,500 types accounting for 51.98% types, 10,069 tokens accounting for 19.02% tokens, and involving 24,601 (35.56%) running words of the whole corpus); half of them are noun-noun combination.

Wansuntarasophit (2008) manually analysed and extracts all the noun phrases from the corpus. He adapted Yang's (1986) criteria to identify complex noun phrases (NP) and used 5 following criteria in his study:

(1) Multi-word terms are mainly nominals;
(2) Multi-word terms cannot go across punctuation marks;
(3) Verbs may be terms by themselves but no part of a multi-word term because of 1.;
(4) Adverbs may be part of multi-word term (e.g. 'VERY high frequency', 'POSITIVELY charged ions', but adverbs for text cohesion (e.g. 'subsequently', 'naturally', 'usually', etc.) should be excluded
(5) No multi-word terms can end up with an adjective or adverb. (pp. 64)

All technical NPs were then identified from the list of all NPs based on the following criteria:

> Nominal groups as noun phrases which provide their meaning related to the fields of general science and engineering are classified as technical noun phrases. These noun phrases refer to scientific principles, items, materials, units, properties, functions, processes, and concepts. (pp. 66)

The others were classified as non-technical NPs and further divided into academic and general NPs. This procedure resulted in a list of mixed types of noun compounds (or NPs in Wasuntarasophit's study). From the working criteria of complex NPs, it could be seen that the list of NPs could include such types of structures like *adv + adj + noun, adj + noun, adj + noun + noun, noun + noun, noun + noun + noun*, and so on. However, adj + noun compounds are often not difficult to interpret. The meaning could simply be the intersection of the meanings of the adjective and the noun. It is not the case of noun + noun compounds which are concerned in this study. Therefore, a list of technical noun + noun compounds need to be extracted from the raw list of technical NPs. Only technical compounds are under consideration, because they characterize the electrical engineering language and could be defined in the context of electrical engineering or in another word, they are discipline-dependent.

Being aware of the high number of nominal compounds in technical texts, Salager (1983) listed word-compound as one of three broad linguistic features of fundamental medical English which should be noted. She added that "they acquire a terminological function and belong to a medical subfield, because words become technical by the compounding of several sub-technical terms." Therefore, they should take a room in teaching of English for specific purposes. Pritchard and Nasr (2004) included recognizing and understanding nominal compounds as one of reading comprehension skill in a study which aimed at improving reading performance of Egyptian engineering students.

Lots of studies in the literature have shown that noun-noun combinations are ubiquitous in technical and scientific English. These combinations may function as anaphora or a specific vocabulary which is distinct from other disciplines. Many of these combinations occur only once or twice which results in a high information load for the student. These studies have also indicated that these combinations are the source of reading comprehension difficulty to engineering students.

2.2 Properties of noun-noun combinations and implications for teaching

2.2.1 Productivity

Compounding is a highly productive process in English word formation (See Chapter 1). According to Downing (1977:53), "compounding thus serves as a back door into the lexicon." Noun compounds account for the largest number of English compounds (Bauer, 1983). She distinguishes 4 types of noun compounds: endocentric, exocentric, appositional, and dvandva on the basis of semantic criteria. An endocentric compound has the referent as a member of the category defined by the head noun such as *honey bee, circuit diagram*. Exocentric compounds do not refer to a concept/thing belonging to the head noun category like *bird-brain (idiot), egg-head (intellectual)*. In appositional compounds, entities denoted by compounds are subcategories of both categories denoted by heads and modifiers such as *maidservant*. For the last type of dvandva compounds (*e.g. singer-actor, parents-teacher*), heads and modifiers are not always clear, and entities referred by compounds are not members of any categories denoted by either heads or modifiers. Most noun compounds are endocentric. The exocentric group is very limited in productivity. The most productive group is the combination of two common nouns, one sub-group of endocentric noun compounds. "This is by far the most productive type of compound, and hundreds of examples can be found in any newspaper, magazine or dictionary" (pp. 204) such as *acid rock, adventure playground, aversion therapy, bang zone, battered baby syndrome, body jewel, bullet train, cable television, credibility gap, domino theory, family planning.* The reason for this productivity could be explained as "the possibility to form compounds from two nouns are unlimited whether they are actually formed, however obviously in the need" in Paul (1995, cited in Fernández-Domínguez, 2009:22). For its highly productivity, naming function, and condensation, noun-noun combinations are intensively made use of in scientific and technical texts. The productivity meets the dynamism of a technology world; naming function and condensation are for the requirements of precision and space-saving in technical texts.

2.2.2 Idiomaticity

From Levi's scale of opacity (Chapter 1), it can be seen that at the other end of the lexicalization scale is idiomaticity. A novel noun-noun combination is formed as productive compounding. This combination could be accepted by a community and gradually enter the lexicon. Under the process of broadening or narrowing the meaning of one or both constituents of a combination or as a result of the lexicalization process, the combination becomes opaque or idiomatic. The meaning of an idiomatic compound cannot be interpreted basing on the meaning of each constituent and the underlying relationship. The historical course of forming *horsepower* is an example of this type. Originally, *horsepower* was the power of a horse, but nowadays it is used as a measurement unit of engines, motors. The meaning of *horsepower* cannot be understood from the combination of *horse* and *power* anymore, but it is understood as a unit with values from 735.5 and 750 watts. Despite the lexical process it undergoes, the typical feature of the idiomatic compound is its meaning deviating from the combination of its constituent parts.

2.2.3 L2 reading and noun-noun combinations

The reading process distinguishes word-recognition and word-integration (Fender, 2001). In word-recognition, words are decoded in terms of their graphical forms and mapped to the sound forms. In cross-language reading, this process may be influenced by the difference in orthography of different languages. For example, Chinese is a logographic writing system, so the word 天 (*sky*) is mapped to the sound *tian1*. This is different from English, in which *sky* is mapped */s/* +*/k/* +*/aɪ/* which are then combined to become */skaɪ/*. Similarly to English, Korean is an alphabetic language. Wang, Koda, and Perfetti (2003) studied the effects of nonalphabetic and alphabetic L1 on English word identification in the case of Chinese learners and Korean learners of English and found out that Chinese learners of English may depends more on orthographic information than on phonological information in comparison to Korean counterparts.

Following word-recognition is word-integration in which readers integrate words into phrases, clauses. This process may be influenced by different syntactic structures of different languages. In a study of Arabic literacy, Fender (2008) postulated

> "Arabic word-integration or sentence-parsing procedures utilize a set of right-branching word-integration or sentence-parsing processes in which the head of a phrase (e.g. the noun in a noun phrase or the verb in a verb phrase) guides and constrains how subsequent words (such as modifiers or complements) are attached or integrated into the preposition, noun, and verb phrase structures." (pp.112)

This process of word-integration is different from such languages as Japanese and Korean but similar to English, which makes English learners of these languages perform differently at the word-integration level (Fender, 2001). In a comparative study of native Arabic- and Japanese-speaking learners of English as a second language in terms of word-integration, Fender (2003) found out Arabic-speaking learners perform better than Japanese-speaking learners. In the word integration experiment of the study, each chunk of the sentence appeared, and then disappeared when the next chunk appeared in the next position. The sentences in the experiment have the structure as the following sentence: *The waiter in the kitchen will bring the food to the table.* This structure is similar to Arabic but different from Japanese. A true/false question was asked to check the learner's comprehension. Arabic learners had better results in terms of word interpretation accuracy which was shown by better comprehension scores. Fender suggested that the overlapping in terms of word integration at phrase and clause level of word integration between Arabic and English languages may explain such difference in the performance of Arabic and Japanese L2 learners.

The similarity and difference in phrase structures between L1 and L2 do influence the quality of L2 reading. Concerning complex nominal compounds, Pastor (2008) conducted a study on Spanish speaking English learners' interpretation of complex noun phrases derived from the differences in English and Spanish that

"English can simply juxtapose different words by their semantic relationship, as in *blood urea nitrogen concentrations.* However, Spanish is not a synthetic language and needs connectors to join the different elements of a phrase […]" (pp. 39). Due to such difference, Spanish-speaking learners find it difficult to identify the headword of a complex noun phrase. Similar to Spanish, Salager-Meyer (1984) stated that French uses prepositions or relative clauses instead of long complex noun phrases.

Table 2.2 Complex noun phrases in English and French

English	French
Water vapor	Vapeur *d'*eau
Sample pressure	Pressure *de* l'échantillon
Relative intensity measurement	Mesure *d'*intensité relative
Polyphrase series commutator motor	Moteur polyphasé série *à* collecteur

Premodifiers of noun compounds in English are also moved to postmodification positions in Vietnamese. For example, *water vapor* is equivalent to *hơi nước* in Vietnamese, in which *hơi* means *vapor,* and *nước* means *water.* This syntactic difference may slow down or interfere Vietnamese-speaking EFL learners' reading. Especially with long complex nominal compounds, the head word must be first identified to be able to integrate with other modifiers. While Vietnamese speakers have formed the routine of recognizing the first word of a noun string as the headword which guides the subsequent words, this routine needs to be adjusted when they read in English, the language of which the headword of a noun phrase is often not at the initial position.

2.2.4 Implicitness of underlying semantic relations

According to Halliday (1985), a noun can be premodified by adjectives, participles, nouns, genitives, adverbs or other phrases, and sentences. Adjectives in the position of premodification qualify and describe the head noun in the same way as they are in the predicative position; for example, *the signal is periodic* and *the periodic signal*. Both present and past participles can be premodifiers. Premodification by participles is possible when there is a tendency of a permanence relation between the quality described by the participle and the thing such as *transmitting signals, transmitted signals*. Past participles often have passive meanings when they are in the position of premodification. A noun can be premodified by genitives such as *Moore's Law*. In some places like *far-away places, round-the–clock services*, and *do-it-yourself jobs*, the nouns *places, services,* and *jobs* are modified by an adverb, a phrase, and a sentence, in respectively. Similar to premodification by participles, a noun can be premodified by another noun when there is a possibility of a relative permanence relationship of the modification. Different from other types of premodification items, "noun premodifiers are so closely associated with the head as to be regarded as compounded with it. In many cases, it appears to reduced – explicitness relation with prepositional postmodifiers" (Greenbaum & Quirk, 1990: 387). An example can be taken such as *voltage source* whether it can be interpreted as *voltage for source* or *voltage from source*. Many other examples like such a noun – noun combination could be found. Raffray, Pickering, & Branigan (2007) took examples of *tourist castle* (castle for tourist) and *mountain castle* (castle located in mountain). Both of them have the same N + N structure but the underlying semantic relation between the two nouns are different, which causes the difference in interpreting the two combinations.

Downing (1977: 820) writes "words are used to denote real, multifaceted entities. The compound picks out but one aspect of the entity on which to base a classification – its shape, its location, or its purpose.” Which basis a combination is on is not made clear by its linguistic surface. For example, *pumpkin bus* may be formed by the shape similarity between a *pumpkin* and a *bus*, but *pumpkin soup* is not inherent such relation. Ryder (1994) discusses the unpredictable output of a noun-noun compound as the product of productivity by taking such examples as:

Water hole: hole containing water

Mouse hole: hole a mouse goes through

Bullet hole: hole created by a bullet

It can be clearly seen that there is no signal to inform which aspects of *hole* will be used; thus there is no way to identify the underlying semantic relationship merely based on their linguistic surface. She comes to the conclusion that “noun-noun compounds exhibit an apparently unlimited range of semantic interpretations.” (pp.5)

Premodification by noun or noun + noun compounding is said to be intensively used in scientific and technical texts for naming, packaging and space-saving purposes, which results in ambiguity in such texts, especially for one who does not share the same background or lack of background knowledge in the field. Pastor (2008) discusses the difference in the structure of English and Spanish as the source of Spanish speaking ESL learners’ difficulty in identifying the key words in a long complex noun phrases. This difficulty is doubled by the implicitness of the relationship between them.

All these previous properties of noun-noun combinations more than often go together, which makes its interpretation multiple difficulties. Productivity results in a number of combinations which L2 learners may encounter for the first time. The meaning of the combination is often based on one aspect of the two entities denoted by its modifier and head noun. This aspect is not explicit, which makes it difficult to interpret. Semantic changes may make some compounds become opaque, which means that they need learning as a whole. Different languages may have different order patterns of modifiers and head nouns, which also may cause obstacles for students in terms of identifying the correct syntactic structure. This may result in an incorrect interpretation.

2.3 Taxonomies of underlying relationship between constituents of noun-noun combinations

Attempts to identify the underlying relationship in noun-noun combinations have been made by lots of researchers in different fields such as linguistics, psycholinguistics, and computational linguistics. Researchers from different approaches have come up with different lists of relationship.

2.3.1 A taxonomy based on underlying syntactic relationship

From transformational view of the relationship, Lees (1960) listed nine underlying grammatical relations:

1. Subject – Predicate: *girlfriend (the friend is a girl), fighter plane (the plane is a fighter)*
2. Subject – Middle Object: *artist's model (the artist has a model), arrow head (the arrow has a head)*
3. Subject – Verb: *talking machine (the machine that talks), population growth (the population grows*)
4. Subject – Object: *steamboat (the steam powers the boat), knife wound (the knife causes a wound)*

5. Verb – Object: *eating apple (eating the apple) farm land (farm the land)*
6. Subject – Prepositional Object: *farm boy (the boy is from the farm), body fluid (the fluid is in the body)*
7. Verb – Prepositional Object: *grindstone (grind knives on the stone), boiling point (boil at the point)*
8. Object – Prepositional Object: *school grammar (grammar taught in school), bedtime story (story told at bedtime)*
9. Proper nouns and naming: *Hemingway book, Keynes approach*

Lees' taxonomy was criticized because of its arbitrariness (Downing, 1977). He gave no explanation for his classification, so one compound can be interpreted with more than one underlying relations. For example, *grindstone* can be verb – prepositional object or subject – prepositional object (*stone is for grinding*)

From the generative semantic view of the relationship, while the main principle was still based on syntactic structure, Levi (1978) proposed two processes forming noun compounds including nominalization and deletion. As for nominalization, the verb of the predicate is converted from the head noun and the modifier is either the subject or object of the nominalized verb (subject as in *population growth*, object as in *sound synthesizer*). In terms of deletion, she proposed that compounds are derived from the process of deleting an underlying relative clause or a complement of a noun phrase. She listed nine predicates:

Cause	*excitation energy*
Have	*insulation layer*
Make	*paper capacitor*
Use	*steam engine*

Be	*oscillator circuit*
In	*field mouse*
For	*bird sanctuary*
From	*peanut butter*
About	*abortion problem*

For the purpose of including all compounds, Levi provided a short and highly abstract list of deleteable predicates. She claimed that all noun-noun combinations could be interpreted by one of predicates in the list. Downing (1977) criticized Levi's analysis in three aspects. Firstly, there was no distinction between nominalization and deletion; Levi gave no explanation why she classified *feminine intuition* is in *have* category but not nominalization. Secondly, more than one predicate can be used to interpret one compound; for example, *peanut butter* could be classified in *make* category or in *from* category. Finally, even if only one predicate is identified to be appropriate, it is still ambiguous by itself; for example, *headache pill* and *fertility pill* are both in *for* category, but one is for reducing headache and another is for enhancing fertility. This ambiguity makes the use of her predicate list not help much in figuring out the meaning of the compound.

2.3.2 A taxonomy based on underlying semantic relationship

Different from Levi (1978), Downing argued that the semantic relations in compounds are unlimited, but that the following are the most common relations:

Whole – part	*duck foot*
Half – half	*giraffe – cow*
Part – whole	*pendulum clock*
Composition	*stone furniture*

Comparison	*pumpkin bus*
Time	*summer dust*
Place	*Eastern Oregon meal*
Source	*vulture shit*
Product	*honey glands*
User	*flea wheelbarrow*
Purpose	*hedge hatchet*
Occupation	*coffee man*

It can be seen some overlapping between Downing's relation list and Warren's (1978) taxonomy which was established from the study of 4557 non-nominalized noun compounds extracted from Brown Corpus. Warren (1978) classified all the relations of noun compounds in her study into 6 groups including: constitute, possession, location, purpose, activity-actor, resemblance, which can be further divided as follows:

Source – result	*clay bird*
Copular	*oak tree*
Resemblance	*club foot*
Whole – part	*spoon handle*
Part – whole	*armchair*
Size – whole	*22-inch board*
Goal-obj	*moon rocket*
Place – obj	*coast road*
Time – obj	*Sunday paper*
Origin – obj	*engine noise*

Purpose	*coffee cup*
Activity – actor	*cowboy*

However, the distribution of these relations are not the same, there is preference for some relations over others. Among them, whole-part relations which are the most preferred account for 23% of all compounds in her study; the least common is resemblance accounting for 1.8%.

Despite the difference in point of view about the limitation number of relations, Downing's most frequent relations and Warren's taxonomy share a lot in common. The relations *part-whole, whole-part, time, place, purpose* are in both lists. Some other categories have different names, but share similar nature, for example *composition – source result*, *comparison – resemblance, source – origin, occupation – activity actor* (the former in Downing's work, the latter in Warren's taxonomy)

Four most common relations which could be found in both lists are presented in Halliday's (1985):

Part – Whole	*clay soil (soil with clay)*
Place	*garden fence (fence around garden)*
Time	*morning train (train in the morning)*
Whole – Part	*board member (member of board)*

From a different angle, Finnin (1980) based his interpretation on the theory of nominal role, in which all nouns are implicitly nominalizations, even in the case when the noun and verb are not morphologically related. e.g. *recipe book* is the case of the implicit nominalization of the verb *write, the book writes about recipe*. As a result, his semantic interpretation went into very particular details. For instance, *dog house* is a house which a dog dwells in, or *cat food* is the food that a cat eats.

2.3.3 A taxonomy based on prepositional paraphrases

With the assumption that all the relation classifications finally boil down paraphrasing noun compounds, Lauer (1995) suggested a list of prepositions to paraphrase noun-noun combinations, including: *of, for, with, in, on, at, about,* and *from*:

Of	*state laws means laws of the state*
For	*a baby chair means a chair for babies*
In	*morning prayers means prayers in the morning*
At	*airport food means food at the airport*
On	*Sunday television means television on Sunday*
From	*reactor waste means waste from a reactor*
With	*gun men means men with guns*
About	*war story means story about war*

One thing which could be easily recognized is the high abstraction which results in ambiguity. For example, *in* could indicate location or time. *City bus* (*a bus in the city*) indicates a location relation, but *morning pray (pray in the morning)* implies a time relation. A similar situation happens to *at* and *on.* On the other hand, the same semantic relation is paraphrased with two different prepositions such as *night flight* (*flight at night*) and *Sunday paper* (*paper on Sundays*). In some other cases, the paraphrase preposition is identified, but the semantic relation is vague, especially with preposition *of*. For instance, *circuit diagram* could be paraphrases *diagram of circuit,* but this paraphrasing has no meaning in terms of the semantic relation indication. For these reasons, a taxonomy of prepositions may have low pedagogical values.

2.3.4 A mixed taxonomy

Girju et al. (2005) approach the issue in a combined manner. In their study, they found out that there were nominal compounds which could not be paraphrased by any of Lauer's (1995) list such as *bus service, daisy flower.* Therefore, in addition to 8 prepositions, Girju et al. (2005) adds 35 more semantic relations. Among them, the most frequent relations are part-whole (*girl mouth*), attribute-holder (*quality sound*), purpose (*migraine drug*), location (*Taxas city*), topic (*art museum*), and theme (*car salesman*).

2.3.5 A taxonomy in technical writing

All the previous studies of the underlying semantic or syntactic relation classifications are based on nominal compounds in general English. A taxonomy which is mainly established for technical writing is proposed by Master (2003). With the same assumption of the possibility to paraphrase nominal compounds into sentential forms or prepositional phrases, he suggests a classification of noun compounds based on the underlying relations:

Properties	*requires adjective+noun, e.g., strong wire, not a noun compound*
Material	*copper wire (wire that is made of copper)*
Operation	*friction brake (a brake that works by means of friction)*
Purpose	*air filter (a filter for cleaning air)*
Location	*field mouse (a mouse that lives in fields)*
Time	*night hawk (a hawk that hunts at night)*

Shape/form	*worm gear (a gear that is shaped like a worm)*
Inventor/Professional user	*Bunsen burner (a burner that was invented by Robert Bunsen)*

Except for the first category because it requires *adjective + noun*, all the others are applied to noun compounds in the consideration of this study or the combination of two nouns. Apart from common relations which are similar to the taxonomies of general noun compounds such as time, purpose, location, these others like material, operation, shape/form, and inventor/professional users are very typical in technical language.

Looking throughout the literature, lots of efforts have been made to classify the underlying semantic relations. Different taxonomies have been suggested from five most common relations by Halliday (1985) to a long list of 8 prepositions and 35 other relations by Girju et al (2005), and even to potentially unlimited numbers by Finin (1980) with nominal roles. However, most of them work on general English. Is there any difference in specialized English in general, and in electrical engineering English in particular? Is there any preferred relation?

2.4 Exercises of training L2 learners' interpretation of nominal compounds in the literature

Different types of exercises have been recorded in the literature to teach noun compounds in technical reading. Ward (2005) suggested several types of exercises to teach collocations which in his research refer to nominal compounds. Exercises could be very simple like underlining all the encountered noun compounds in the reading text to get students' attention to such compounds in the text. A little more difficult exercises could be matching the correct definition of a compound or rearranging the order of words to have a correct order noun compound. The two most difficult types are interpretation and production. The format of interpretation exercise may look like the below example (reading texts are previously provided):

Answer each of the following yes, no, or maybe

Is hydrogen sulphide:

a chemical species?	*yes*
a product species?	*maybe*
a reagent species?	*maybe*
a gas phase?	*yes* (Example from Ward (2005))

Short – answer exercises could be one simple format of production type. Students could be asked about the information they read in the text and they need to look for a noun compound to answer.

Master (2003) suggests 4 steps to train students' interpretation of noun compounds going along with 4 types of exercises. The first step is to ask students to classify a given list of compounds into its categories (his list of 9 categories of noun compounds including properties, material, operation, location, time, shape/form, inventor/professional user). The second step is to paraphrase nominal compounds into relative clauses or change them into postmodification structures. Matching exercises could realize this step like the following examples:

Instructions

Choose the correct definition for the noun compound on the left.

Example: worm gear *a. a worm that lives in gears*

b. a gear shaped like a worm

Answer: b. a gear shaped like a worm

The third step is to practice students to produce a correct nominal compound. Students are asked to form nominal compounds from the given original forms. For example, “the *vent for air* should be open” is rewritten “the *air vent* should be open.” The final step deals with production of complex nominal compounds (more than two words) in a long paragraph (instead of one sentence like the preceding example) to get the student attention to the use of such nominal compounds in the actual language.

In conclusion, this chapter reviews research on noun-noun combinations in the literature. These studies show that noun-noun combinations are ubiquitous in technical language. However some of their properties such as productivity, idiomaticity, different structures in different languages, and implicitness of semantic relationship makes them difficult to interpret for L2 learners. Among these properties, the implicitness of semantic relationship plays the center role. A large number of research has been done on creating a taxonomy of these relations, but little has been done on technical English as well as designing exercises to help L2 learners overcome the difficulties caused by these combinations.

CHAPTER 3

METHODOLOGY

In this chapter, the procedures to conduct the study are presented. Generally speaking, the steps were as follows: creating a list of noun-noun combination from the corpus, sampling, examining the taxonomy, classifying the agreed taxonomy, and calculating the frequency of relation. The details of each step will be discussed in the following sections.

3.1 Criteria for a noun-noun combination (NNC) to be included

To conduct the study, a list of technical noun-noun combinations needed to be established from Wasuntarasophit's (2008) corpus of electrical engineering textbooks, so-called EEC. The following criteria were applied for the inclusion of one noun-noun combination in the study:

Any combination of two nouns, in which a noun "refers semantically to those aspects of our experience which we perceive as 'things' or 'entities'. The term 'thing' refers here not only to concrete entities such as persons, objects, places, institutions and other 'collectives', but also to the names of actions (reading, laughter), abstractions (thought, experience), relationships (friendship, obedience), qualities (beauty, speed), emotions (anger, excitement), phenomena (thunder, success), and many other classes of entities." (Downing and Locke, 2002: 406)

Based on this definition, the following combinations were included:

- Noun - noun combinations of which one constituent is gerund, *e.g. bandpass signaling, addressing scheme*. Gerunds and present participles are distinguished by the assumption that present participles have characteristics of a genuine adjective, and gerunds have characteristics of a noun. Therefore, a present participle can be in predicate position and can be modified by adverbs, while a gerund function as a noun which cannot be predicative and cannot be modified by adverbs. A gerund can have an adjective in the position of premodification. For example, a *swimming pool* is a *pool for swimming*, not *a pool is swimming*. We have a *beautiful swimming pool,* but not a *beautifully swimming pool.* In contrast, a *sleeping child* is a *child who is sleeping*. The interpretation of *a child for sleeping* is implausible. We can also say *a deeply sleeping child.*

Noun – noun combination of which one constituent could be an adjective in one case and be a noun in another case. For example, *plastic wire* could be interpreted as *a wire made of plastic* or *an elastic wire*. In the first interpretation, *plastic* functions a noun which denotes the material of the entity *wire*; in the second interpretation, *plastic* functions as an adjective with the meaning of *having the quality of plastic*.

Also based on these criteria, all combinations with head nouns modified by more than one noun like *clock recovery circuit* are not be included for the simplification purpose. By bracketing them into nominal groups, the taxonomy of underlying relations could be applied between the new nominal groups because of recursion of noun compounding. For example, *clock recovery circuit* could be bracketed as (*clock recovery*) *circuit,* and *clock recovery* is regarded as one nominal group with which *circuit* has an underlying semantic relationship. Moreover, noun-noun combinations of more than 2 words do not frequently occur throughout the corpus.

In addition, all the noun phrases with only one noun functioning as head noun modified by one or more than one adjective and adverb such as *positively charged ion* were excluded from the study. For the noun phrases which consist of more than 2 words such as the structure *adverb* + *adjective* + *noun* + *noun*, the combination noun-noun were extracted to be included. However, there was one exception; noun – noun combinations in which one of the constituents is proper noun, e.g. *Newton law, Microsoft products* were not included in the study.

One more point that should be noted is that some researchers include in their study of nominal compounds which in Levi's analysis is named "non-predicate adjectives." This class of premodifiers was not included. Non-predicate adjectives are adjectives which have the surface of an adjective but are derived from nouns. Different from real adjectives, these adjectives have different meanings when they are in attribute positions and when they are in predicate positions. In some cases, the sentential paraphrase of the compound by changing the non-predicate adjective in the attribute positions to predicate positions as we can do with genuine adjectives is nonsense, e.g. *electrical engineer* cannot be interpreted as *engineer is electrical*[*]. The correct interpretation should be *engineer who specializes in electrical engineering.* In some other causes, when one appears as a predicate, the meaning is not the same as it functions as an attribute. For example, compare *criminal lawyer* and *the lawyer who is criminal*. One more criterion to distinguish those adjectives is that they cannot be modified by intensifier such as *very, extremely, fairly*. There is no a *very electrical engineer*.

In conclusion, this study covered only combinations of two nouns. Adjectives and adverbs were omitted for the combination of *noun and noun* in *adverb* + *adjective* + *noun* + *noun* to be included. In the case the modifier having –ing form, that combination would be included if the –ing form was identified as a gerund. A combination of which the modifier is a case of conversion from an adjective to noun was counted.

3.2 The corpus

This study worked on the specialized language, or language used electrical engineering in particular. The study was expected to help students majoring in electrical engineering and related disciplines in reading technical texts, especially in interpreting technical noun-noun combinations which are pervasive in technical texts. Therefore, a corpus of electrical engineering textbooks is needed. Fortunately, Wasuntarasophit (2008) has developed such a corpus for the purpose of studying the proportion of different types of lexical units such as technical, academic, and general lexical units in electrical engineering textbooks. He kindly allowed the writer to use his corpus for the purpose of this study.

This corpus of electrical engineering textbooks developed by Wasuntarasophit (2008) was compiled from 5 textbooks in 4 sub-fields in the field of electrical engineering including Control Systems, Power Systems, Electronics, and Communications, one textbook from each field and one textbook for background knowledge needed for electrical engineering major. One hundred pages were randomly selected from each book, changed into electronic form, and stored in the form of text files. Because of limitations imposed by copyright and manual analysis, the corpus was limited to the size of approximately 120,000 running words. All complex noun phrases were extracted from the corpus (see Chapter 2 for detailed criteria and procedures). For the purpose of his study, he listed a list of complex technical noun phrases. The problem is that he included in his list all 2 word noun phrases, 3 word noun phrases, and more. All types of noun phrase structures were included such as *adjective* + *noun, adverb* + *adjective* + *noun,* and *noun* + *noun*. With the aim of this study, a new list needed to be established. The criteria in section 3.1 were applied to establish this list.

3.3 Creation and analysis of NNC list from NP list extracted from the corpus

A list of NNCs was created based on the criteria in section 3.1 and the available list of NPs extracted from the corpus. The creation and analysis of the list were carried out as follows:

1. Eliminating all *adj+n* combinations in the list based on the above criteria such as *distortionless amplifier, logarithmic scale,* and *logical sum.*

2. Eliminating combinations of more than two nouns such as *amplifier bias circuit, channel enhancement transistor,* and *circuit output impedance.*

3. Omitting adjectives in combinations of *adj* + *n* + *n* structures such as *applied gate pulses (gate pulses), balanced phase capacitors (phase capacitors),* and *constant terminal voltage.* If the *n+n* combination remained after the adjective is omitted is the same type or token with any NNC in the list, its frequency will be added up to the frequency of that NNC. For example, the frequency of *terminal voltage* is 30, and the

frequency of *constant terminal voltage* is 1; thus, after omitting *constant* from *constant terminal voltage,* the frequency of *terminal voltage* is 31.

4. Eliminating combinations with hyphens of such structures as *voltage-regulation characteristics, shorted-emitter design, signal-to-noise ratio,* and *common-mode signal.*

5. Recalculating different types of NNCs with the same meaning. The original NP list was built based on types, so such NNCs as *minority carriers* and *minority carrier* are listed in two lexical items. The target of the analysis was the semantic relationship of the two nouns of the combination, so it involved the meaning of the NNC as a whole and the meanings of each constituent. There is no difference in terms of meaning between *minority carrier* and *minority carriers,* thus *minority carriers* was omitted from the list and its frequency was added to the frequency of *minority carrier.*

6. Identifying the relations of the modifier and the head noun of the combination by linking the meaning of the modifier and the head noun to the meaning of the combination. High frequency relation types were coded by a number for frequency calculation.

3.4 Random sampling

The purpose of the study was to find out high frequency semantic relations. A complete list of available relations was not the aim. Thus, analysis of all noun-noun combinations was not expected; it is very time-consuming but the results cannot be said to be much more precise than the results from sampling which is more effective. Stratified random sampling was an appropriate method to be carried out for the two reasons. Firstly, "a stratified random sample of given size will yield a more precise estimate than a simple random sample of the same size, since much smaller variability is encountered […]" (Neter & Whitmore, 1988:457) Secondly, it provided more meaningful information about the subpopulation (Neter & Whimore, 1988); in this case, it was the different frequency noun-noun combinations.

For the particular case of the study, the stratified random sampling and analysis were conducted as follows. All noun-noun combinations were divided into three groups in terms of frequency, including high frequency group (of NNCs with frequency of over 10), medium frequency group (of NNCs with frequency from 3 to 9), and low frequency group (of NNCs with frequency of 1 and 2). The purpose of this study was to find out the high frequency relations. High frequency relation types could result from one or several high frequency types or many low frequency types having the same relation. For an explication about tokens and type, each occurrence of one word form is counted as one token; thus in the previous sentence, we have 22 tokens. The same word form is counted as one type, so we have 15 types in that sentence. In this study, two-word combinations were considered as a lexical unit and the concept of tokens and types of one-word lexical units were applied to two-word lexical units. For example, in this sentence "*one application of filter circuits is in the "conditioning" of non-sinusoidal voltage waveforms in power circuit,"* three underlined combinations are, for our purposes, three types. *Filter circuits* occur only once in the sentence, so the number of tokens of *filter circuits* is one.

High frequency relation types could result from one or several high frequency types; for example, if 30 occurrences are a high frequency, and *output voltage* occurs 30 times in the corpus, the relation *location* in this combination will have high frequency of 30. High frequency relation types could result from many low frequency types having the same relation. For example, if 30 types occur once through the corpus (30 low frequency types) are identified as having the same *location* relation, *location* is a high frequency relation resulting from many low frequency types.

One thing we may be sure of is that high frequency types, certainly, provide a list of high frequency relations. Therefore, all high frequency types were analyzed. High frequency types, however, often do not account for a large number of all noun-noun types. There is a possibility that the long list of low frequency types could make up a list of high frequency relations. A sample from the medium frequency group and one from low frequency group are also taken to be analyzed. This method allows an insight into low frequency and medium frequency types which could not be done if a random sampling of the whole list of noun-noun combinations is taken. The resulting relation list was compared with the list of relations of the high frequency group to find out high frequency relations.

3.5 Interpretation of NNCs

The following sections present how the NNCs were interpreted or paraphrased to be able to identify the semantic relations exhibited by the combinations. Basically, there are two ways of paraphrasing: clausal paraphrasing and prepositional paraphrasing. But for the purpose of precision for the analysis, clausal paraphrasing was used, but for the purpose of teaching, prepositional paraphrasing could be useful. Both ways of paraphrasing will be presented in detail in the two following parts.

3.5.1 Clausal paraphrasing

According to Levi (1978), an NNC is shortened form of a relative clause or a complement structure. For example, *induction motor* is derived from the relative clause *a motor that works by means of induction*; *block diagram* is shortened from the clause *a generator that is structured in block manner*. Complement structures are derived from nominalized verb phrases; for example, *data compression* is derived from *[we should] compress the data* or *carrier generation* from *[the device] generates the carrier*. In Fabre' (1996:364) words, "interpreting nominal compounds consists in retrieving the predicative relations between the constituents." *Motor works by means of induction* is retrieved from *motor induction; data are compressed* from *data compression*. In other words, a NNC has a clausal paraphrase, which is also pointed out by Downing (1977).

3.5.2 Prepositional paraphrasing

In addition to clausal paraphrases, an NNC could be paraphrased by postmodifying with prepositional phrases. "In most cases, premodifying nouns correspond to postmodification with prepositional phrases" (Quirk et al., 1985:1330). For example:

Input current	*current at the input*
Voltage drop	*drop in voltage*
Transmission line	*line for transmission*
Commutation problem	*problem concerning commutation*
Series combination	*combination in series*
Flux density	*density of flux*

Weaknesses of prepositional paraphrasing

For purposes of brevity as well as naming purposes, premodification by noun or noun-noun combination is chosen instead of a long description with postmodification. To interpret the NNC is to recover its postmodification form like in the examples. However, there are some three shortcomings with regard to prepositional paraphrases (see Chapter 2)

1. Not all NNCs could be paraphrased with a preposition (*e.g. carrier signal (the signal is a carrier), compound generator (the generator is a compound type))*. The relation of this combination class is often classified as *be* (Girju et tal., 2005; Levi, 1978) or *copular* (Warren, 1978). Examples taken from their studies are *solder ant, girl friend, Dallas city.*

2. There is more than one relation which could be conveyed by one preposition. For instance, *in* in *loss in the core*, paraphrase of *core loss*, expresses a location relation, but it is not the case of *in* in *drop in voltage*, paraphrase of *voltage drop*.

3. There are some prepositions which are very vague in terms of relation indication because of their metaphoric feature, especially in the case of *of*. For example, does *of* in *density of flux (flux density)* mean the same thing as *of* in *regulation of voltage (voltage regulation), or position of armature (armature position)*? The prepositional paraphrase in those examples seems not to help much in terms of interpretation the NNC.

Usefulness of prepositional paraphrasing

Despite those shortcomings, prepositional paraphrases still play a role in NNC interpretation. First, it is the most intuitively easy paraphrase because most NNCs as a premodification are derived from their postmodification, and it is the first step to retrieve its clausal structure. For example:

Drift current:

Current from the drift	*the current results from the drift*

Communication system:

System for communication	*the system is used for communication*

In some cases when the preposition exhibits some clear relation, it seems that prepositional paraphrases are easier than clausal paraphrases for the pedagogical purpose under the criterion of teachability. As in the case of *in* indicating the location relation:

Line current:

Current in the line and *current located in the line*

Core loss:

Loss in the core and *loss located in the core*

Secondly, prepositional paraphrasing may be useful for English learners whose first languages have different structure of NNCs, such as Vietnamese students. For example, *core loss* is translated into *tổn hao lõi,* in which *tổn hao* is equivalent to *loss* and *lõi* to *core,* or into *tổn hao trong lõi* which is actually equivalent to *loss in the core.* By converting premodification structure into postmodification one could help them to identify the head noun and modifier of NNC and by this way, identify the different functions of each noun in the combination. Converting premodification into postmodification by using prepositions is especially helpful when the combination consists of more than two words. For example, *armature voltage drop* will be bracketed as *armature (voltage drop)*, then paraphrased as *voltage drop in the armature*. *Voltage drop* is then paraphrased as *drop in voltage.* Then, the postmodification structure of *armature voltage drop* is *drop in voltage in the armature.* However, in accordance with the conditions of this study, combinations of more than 2 nouns will not be discussed in detail.

In conclusion, a NNC could have a clausal paraphrase and most of the time a prepositional paraphrase. Clausal paraphrases are more explicit, and for analysis purposes, clausal paraphrase were used (the underlying semantic relation determined the clausal paraphrase of the NNC). For teaching purposes, prepositional paraphrases could be useful for their intuitive ease and simplicity.

3.5.3 Questions and relations

Each NNC could be paraphrased by a clausal or prepositional phrase; another way to express this would be, say that a question could be posed based on the semantic relation of the NNC, and the answer to the question is the clausal or prepositional paraphrase of the NNC. Each relation would go with one question which helps to understand what the relation really means. For example, *gate signal* could be paraphrased in a sentence as *the signal is located at the gate* or in a prepositional paraphrase as *signal at the gate.* The question could be asked to have that answer is *where is the signal?* The relation between *gate* and *signal* could be classified as *location. Gate* specifies the location where the *signal* occurs. And the question *where is the signal?* could be generalized as *where is Y?* with Y as the head noun, X as the modifier. The question could be a means of explaining the way of naming the relation and could be used as a means to determine whether a NNC belongs to that relation or not. If the NNC gives an appropriate answer to the question, that NNC belongs to that group of semantic relation. For example, among *terminal voltage, induction motor, and compound generator,* only *terminal voltage* could give the appropriate answer to the question *Where is Y?* or in this case *Where is the voltage? At the terminal.* In the case of *induction motor, induction* could not provide a suitable answer for the question *where is the motor?*; thus, *induction motor* could not be classified under the *location* semantic relation; similarly to the case of *compound generator.*

Questions help to make the relations explain themselves and act as a tool to examine whether or not a NNC is under a relation category. Moreover, the questions are also helpful in terms of teaching. Questions could be asked by the teacher to guide students to interpret NNCs. Therefore, in addition to its analysis purpose, questions could also be used in teaching practice.

3.6 Semantic relation taxonomies

3.6.1 A semantic relation taxonomy proposed to be used in the study

This study concerns technical noun compounds and semantic relations of their constituents. The NNCs in the study were examined against the taxonomy of noun compound relations by Master's (2003):

Table 3.1 Master's Taxonomy of Noun Compound Relations

Properties	**requires adjective+noun, *e.g., strong wire, not a noun compound***
Material	copper wire (*wire that is made of copper*)
Operation	friction brake (*a brake that works by means of friction*)
Purpose	air filter (*a filter for cleaning air*)
Location	field mouse (*a mouse that lives in fields*)
Time	night hawk (*a hawk that hunts at night*)
Shape/form	worm gear (*a gear that is shaped like a worm*)
Inventor/Professional user	**Bunsen burner (*a burner that was invented by Robert Bunsen*)**

However, only noun-noun combinations were concerned, the first category *properties* was excluded. Similarly, proper nouns were excluded so that inventor category would not be included.

It should be noted that most of these relations could be expressed in terms of Levi's taxonomy or Downing's taxonomy.

Table 3.2 Comparison of three taxonomies

Master's	Levi's	Downing's
Material (*copper wire*)	Made of (*wire made of copper*)	Composition (*wire composited of copper*)
Operation (*friction brake*)	Use (*brake use friction*)	x
Purpose (*air filter*)	For (*filter for air*)	Purpose
Location (*field mouse*)	In/From (*a mouse in the field*)	Place (*a mouse that lives in the field*)
Time (*night hawk*)	In	Time
Shape/form (*worm gear*)	Be (gear is like a worm)	Comparison (*gear looks like a worm*)

It could be seen that one of Master's categories could be corresponded by one of Levi' relations. For *in* category in Levi's taxonomy, even though it could be used to expresses *time* relation (e.g. *morning paper: paper in the morning), in* cannot be used to paraphrase Master's example of time category *night hawk (a hawk at night)*. Similarly, *Sunday picnic (a picnic on Sunday)* could not be categorized in any of Levi's taxonomy. Despite attempting to cover all nominal compounds in her taxonomy, she fails to cover some exceptions like the previous examples. However, the biggest shortcoming of her taxonomy is the ambiguity of relations expressed by any of her items in the taxonomy. Is *a wire made of copper* similar to *a program makes errors* even though they are both in the category of *make*?

In the case of Downing's taxonomy, there is one blank, which means there is no equivalence found for *operation*. It could be explained by Downing's stance concerning the availability of unlimited relations existing between noun and noun in the combination (see pages 15, 37). Her list merely consists of the common relations she found in her study. *Operation* is not in the list because this category may be common in technical English, and it is not common in general English which is the domain of Downing's study. Moreover, the relations in her taxonomy do not render any technical aspect as in Master's taxonomy which is useful for engineering students in paraphrasing the combinations. For example, a noun compound is classified in *operation* category when it answers the question *how does it work?* This is a very common question in the technical context for any kind of devices, equipment, machines and so on. *Friction brake* is paraphrased as *brake works by means of friction.* Therefore, *operation*, itself, renders some technicalness.

In conclusion, the following taxonomy which was modified from Master's (2003) one were used to check against the high frequency combination types:

Table 3.3 Relation taxonomy used to examine high frequency combination types

Relations	Examples
Material	Iron hammer *(hammer made of iron*)
Operation	Induction motor (*motor works by means of induction*)
Purpose	Coffee grinder (*grinder for crushing coffee*)
Location	Source data (*data in the source*)
Time	Morning shift (*shift in the morning*)
Shape/form	Worm gear (*gear shaped like a worm*)

3.6.2 Examining existing taxonomies

3.6.2.1 Examining Master's taxonomy

All the high frequency combination types (NNCs with frequency of over 10) were classified into Master's taxonomy (X: the combination could not be classified into any category of the taxonomy)

Table 3.4 Classification of high frequency types into Master's taxonomy

No.	NNCs	Freq.*	Relations
1	Power factor	66	X
2	Phasor diagram	60	X
3	Output voltage	41	Location (*The voltage located at the output*)
4	Terminal voltage	33	Location (*The voltage located at the terminal*)
5	Voltage drop	33	X
6	Induction motor	32	Operation (*The motor works by means of induction*)
7	Circuit breaker	29	Purpose (*The breaker for breaking the circuit*)
8	Input impedance	26	Location (*The impedance located at the input*)
9	Voltage gain	26	X
10	Line current	25	Location (*The current located in the line*)
11	Base current	23	Location (*The current located at the base*)
12	Input signal	23	Location (*The signal located at the input*)
13	Transition region	23	X
14	Flux density	22	X
15	Power system	21	Purpose (*The system for distributing power*)
16	Voltage source	21	Purpose (*The source for supplying voltage*)
17	Input voltage	20	Location (*The voltage located at the input*)
18	Voltage regulation	20	X
19	Core loss	19	Location (*The loss located in the core*)
20	Output impedance	19	Location (*The impedance located at the output*)
21	Bit error	19	Location (*The error located in the bit*)
22	Phase angle	19	X
23	Copper loss	18	Location (*The loss located in the copper winding*)
24	Depletion region	18	X
25	Bit period	18	X
26	Line voltage	17	Location (*The voltage located in the line*)
27	Load current	17	Location (*The current located in the load*)
28	Current source	16	Purpose (*The source for providing current*)
28	Transmission line	16	Purpose (*The line for transmitting*)
30	Carrier signal	16	Purpose (*The signal for carrying other signals*)
31	Transfer function	16	X
32	Frequency component	16	X
33	Peak value	15	Location (*The value located at the peak*)
34	Valence band	15	X
35	Power supply	15	Purpose (*The supply for providing power, supply in this context means a device*)
36	Leakage flux	15	X
37	Armature current	14	Location (*The current located in the armature*)
38	Bias current	14	Purpose (*The current for biasing*)

No.	NNCs	Freq.*	Relations
39	Communication system	14	Purpose (*The system for communicating*)
40	Sine wave	14	X
41	Field current	13	Location (*The current located in the field*)
42	Gate oxide	13	Location (*The oxide located at the gate*)
43	Source voltage	13	Location (*The voltage located in the source*)
44	Frequency band	13	X
45	Power output	13	Purpose (*The output for power*)
46	Truth table	13	X
47	Operating point	13	X
48	Cutoff frequency	12	Location (*The frequency located at the cutoff point*)
49	Gate bias	12	Location (*Bias located at the gate*)
50	Input resistance	12	Location (*The resistance located at the input*)
51	Temperature rise	12	X
52	Lowpass filter	12	Purpose (*The filter for filtering lowpass signals*)
53	Correlation receiver	12	Operation (*The receiver works by means of correlation*)
54	Load line	12	X
55	Base speed	12	X
56	Minority carriers	12	X
57	Drain current	11	Location (*The current located in the drain*)
58	Input power	11	Location (*The power located at the input*)
59	Output signal	11	Location (*The signal located at the output*)
60	Voltage level	11	X
61	Wave function	11	X
62	Leakage reactance	11	X
63	Carrier frequency	11	X
64	Turns ratio	10	X
65	Bandpass filter	10	Purpose (*The filter for filtering bandpass signals*)
66	Logic gate	10	Operation (*The gate works by means of logic operations*)
67	Saturation current	10	X
68	Logic function	10	X
69	Receiver design	10	X
70	Block code	10	X
71	Harmonic currents	10	X

Note: * Frequency

It can be seen that only three relations in Master's taxonomy, *location, purpose*, and *operation* could be identified. Of 71 combination types, 32 could not be classified in any category in Master's taxonomy, which accounts for 45.07%. These 32 unlisted types account for 45.39% of tokens compiled by those types. The relation between the two nouns in the combination based on Master's taxonomy is identified by the semantics of the premodifier; if the modifier identifies location, material, time, shape, operation (manner), or purpose, then that will be the semantic relation of the combination. In other words, the modifier determines the relation. For example, in the case of *material* relation, the common interpretation of the NNC with this relation is the entity expressed by the head noun made of the material expressed in the modifier such as *copper wire* (*wire made of copper*). Thus the modifier should convey the meanings of some material. Similarly, in the case of *shape/form* relation, the modifier should convey the sense of some shape or form like *worm gear (the gear shaped like a worm).* However, many of the modifiers in the noun-noun combination list extracted from electrical engineering textbooks do not identify such relations. Most of them are physical quantities such as current, voltage, frequency, power, which may partially explain the low coverage of Master's taxonomy.

3.6.2.2 Examining Levi's taxonomy

Because of the failure of Master's taxonomy, Levi's taxonomy was checked out against all the high frequency combination types. It should be noted that Levi's taxonomy is one of the most frequently used (Gagné, 2001; Macguire et al., 2008) and also the vaguest one because of Levi's attempt to cover all nominal compounds. Levi used 4 prepositions (*in, for, from, about*) and 5 verbs (*cause, have, make, use, and be*) to describe the semantic relation in a NNC (the overlapping among relation taxonomies were discussed in section 3, chapter 2). The following result showed the failure of the taxonomy:

Table 3.5 Classification of high frequency combination types into Levi's taxonomy

No.	NNCs	Freq.	Relations
1	Power factor	66	X
2	Phasor diagram	60	X
3	Output voltage	41	X
4	Terminal voltage	33	X
5	Voltage drop	33	In (*drop in voltage*)
6	Induction motor	32	Use (*motor uses induction*)
7	Circuit breaker	29	For (*breaker for circuit*)
8	Input impedance	26	X
9	Voltage gain	26	In (*gain in voltage*)
10	Line current	25	In (*current in the line*)
11	Base current	23	X
12	Input signal	23	X
13	Transition region	23	Has (*region has transition*)
14	Flux density	22	X
15	Power system	21	For (*system of power*)
16	Voltage source	21	For (*source for voltage*)
17	Input voltage	20	X
18	Voltage regulation	20	X
19	Core loss	19	In (*loss in the core*)
20	Output impedance	19	X
21	Bit error	19	Be (*bit is the error*)
22	Phase angle	19	X
23	Copper loss	18	In (*loss in the copper winding*)
24	Depletion region	18	Has (*region has depletion*)
25	Bit period	18	X
26	Line voltage	17	In (*voltage in the line*)
27	Load current	17	In (*current in the load*)
28	Current source	16	For (*source of current*)
29	Transmission line	16	For (*line for transmission*)
30	Carrier signal	16	For (*signal for carrying*)
31	Transfer function	16	X
32	Frequency component	16	Make (*frequency made of components*)
33	Peak value	15	X
34	Valence band	15	X
35	Power supply	15	For (*supply for power*)
36	Leakage flux	15	X
37	Armature current	14	X
38	Bias current	14	For (*current for biasing*)

No.	NNCs	Freq.	Relations
39	Communication system	14	For (*system for communication*)
40	Sine wave	14	X
41	Field current	13	In (*current in the field*)
42	Gate oxide	13	X
43	Source voltage	13	In (*voltage in the source*)
44	Frequency band	13	X
45	Power output	13	For (*output for power*)
46	Truth table	13	X
47	Operating point	13	X
48	Cutoff frequency	12	X
49	Gate bias	12	X
50	Input resistance	12	X
51	Temperature rise	12	In (*rise in temperature*)
52	Lowpass filter	12	For (*filter for filtering lowpass signal*)
53	Correlation receiver	12	Use (*receiver use correlation*)
54	Load line	12	X
55	Base speed	12	X
56	Minority carriers	12	In (*carriers in the minority*)
57	Drain current	11	In (*current in the drain*)
58	Input power	11	X
59	Output signal	11	X
60	Voltage level	11	X
61	Wave function	11	X
62	Leakage reactance	11	X
63	Carrier frequency	11	Has (*carrier has frequency*)
64	Turns ratio	10	X
65	Bandpass filter	10	For (*filter for filtering bandpass signal*)
66	Logic gate	10	Use (*gate use logic operations*)
67	Saturation current	10	X
68	Logic function	10	X
69	Receiver design	10	X
70	Block code	10	In (*code in block*)
72	Harmonic currents	10	Has (*currents have harmonic*)

With the purpose of covering as much as possible all NNCs, Levi established a taxonomy of words which conveys very general meanings, but still, as the table shows 37 high frequency types are still unaccounted for. These 15 types account for 52.11% combination types and 54.44 % tokens. It can be seen that even if a very vague taxonomy which could be said that there was little pedagogical meaning such as Levi’s taxonomy was used, it could not cover all the NNCs although the coverage was higher than that of Master’s taxonomy in terms of types. The question raised is to establish a taxonomy which could have a good coverage and be specific enough to be teachable.

Master’s taxonomy was applied to classify the noun-noun relationship in this study. The noun compounds extracted were then classified under one of these categories. Because of the failure of Master’s taxonomy in covering the relation use, Levi’s taxonomy, the vaguest one, was also examined to see whether such ambiguous taxonomy could cover a large number of NNCs in the list. Despite its ambiguity for a better coverage, Levi’s taxonomy also failed. A new taxonomy was created on the basis of pedagogical purposes.

In conclusion, the study were conducted step by step as follows: (1) creating list of technical noun-noun combinations from the corpus; (2) sampling; (3) examining Master’s taxonomy by classifying all the high frequency types into different categories of the taxonomy; (3) creating a new taxonomy on the basis of keeping appropriate Master’s relation categories; (4) classifying the rest of noun-noun combinations sampled into the taxonomy used in the study. Clausal paraphrasing was used in the process of analyzing semantic relations. Exercises were then designed to familiarize engineering students with the idea of lost semantic relations.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Semantic relations available in the EEC

4.1.1 The new taxonomy and its coverage.

The new classification of semantic relations was built up by classifying all the NNCs which could not be categorized in Master's taxonomy. The relations of the high frequency NNC types which did not exhibit any relation in Master's classification were identified and the revised list of relations was checked against the samples from medium frequency types and low frequency types. The NNCs which did not exhibit the relations in the revised list were analyzed to identify the relations. If the occurrences of NNCs of a relation added up to 10 times (i.e. a relation with the number of tokens up to 10), the relation would be added to the revised list. The final taxonomy of relations was thus created.

The analysis resulted in a list of 8 common notional semantic relations, including: location, purpose, measure, representation, source, operation, objective, and structure. These 8 semantic relations answer 7 questions (objective does not go with any question). Before the results of each relation's frequency calculation is reported, the following section is a description of 8 notional semantic relations.

4.1.2 Description of 8 notional semantic relations

1. Location: Where is Y? (Y: head noun, X: modifier)

Location was the most common relation among the two nouns in the combination in this study. It was the highest frequency relation types among all the high frequency types. Location is among also the most common relations of different taxonomies.

If a combination XY has the location relation, it could be interpreted as Y located in or at X. In this way, X often conveys a location sense such as *output voltage, input signal, terminal voltage, base current*. In some other cases, X could be a device, in which a component (Y) is located: *Relay coil: The coil is located in the relay or the coil in the relay*

Electrocardiograph amplifier: The amplifier is located in the electrocardiograph or the amplifier in the electrocardiograph

Machine core: The core is located in the machine or the core in the machine

The question this relation answers is "*where is Y?":*

Output voltage:

Where is the voltage?

→ *(The voltage is located) at the output.*

Line current:

Where is the current?

→ *(The current is located) in the line.*

Core loss:

Where is the loss?

→ *(The loss is located) in the core.*

From the examples, it can be seen that two prepositions *at* and *in* could be used in prepositional paraphrase of NNCs with the location relation such as *voltage at the output, current in the line,* and *loss in the core.*

2. Purpose: What is the purpose of Y?

A purpose relation was exhibited by a combination when it could answer the question *what is the purpose of Y?* and very often in this field, the head noun is a device and the question became *what is the purpose of the device?* as in the case of:

Bandpass filter:

What is the purpose of the filter?

→ *(The purpose of the filter is) to filter the bandpass signal*

Circuit breaker:

What is the purpose of the breaker?

→ *(The purpose of the breaker is) to break the circuit.*

Signal generator:

What is the purpose of the generator?

→*(The purpose of the generator is) to generate the signal.*

But there are also cases which the head noun is not exactly a device such as in:

Transmission line:

What is the purpose of the line?

→ *(The purpose of line is) to transmit power.*

Power system:

What is the purpose of the system?

→ *(The purpose of the system is) to distribute power.*

In terms of prepositional paraphrases, the preposition *for* could be used in paraphrasing this relation. The previous examples could be paraphrased as follows with preposition for:

Bandpass filter: filter for filtering bandpass signals

Circuit breaker: breaker for breaking circuits

Signal generator: generator for generating signals

Transmission line: line for transmission

Power system: system of power

3. Measure: What Y is measured?

All of electrical quantities like voltage, current, resistance, reactance, and impedance could be measured in some way, which may explain why *measure* appeared in a semantic relation taxonomy for electrical engineering language. If a combination XY exhibited this relation, the interpretation could be that Y in X is measured. X was often some quantities like current, voltage, temperature, and Y was the result of the measurement of those quantities. Results could be some changes, some values of that quantity, some ratios.

The question this relation may answer was *What Y is measured?* For example,

Voltage gain:

What gain is measured?

→ *The gain in the voltage is measured.*

Voltage level

What level is measured?

→ *The level of the voltage is measured.*

Recombination rate

What rate is measured?

→ *The rate of the recombination is measured.*

Preposition *in* could be used to paraphrase some NNCs of this group such as *voltage gain (gain in the voltage), current drop (drop in the current)*. Previously, *in* could be used to paraphrase *location* relation, but *in* also could be used to paraphrase some NNCs of *measure* relation. These cases show the inadequacy of using preposition only in paraphrasing NNCs. For other cases such as *voltage level* and *recombination rate, of* could be used, *level of voltage, rate of recombination*.

4. Representation: What does Y represent?

Phasor diagram could be interpreted as *a diagram representing a phasor*; a *receiver design* could be paraphrased as a design represents a receiver. Thus, a *representation* relation exhibited when Y represents X or in other words, when the NNC answered the question *What does Y represent?*

Examples:

Regulation curve:

What does the curve represents?

→ *The curve represents the regulation.*

Wave function

What does the function represent?

→*The function represents how a wave transmits.*

Load line

What does the line represent?

→ *The line represents the operation of the load.*

This case, once again, proves the first shortcoming mentioned in the previous section. Prepositions could not paraphrase all NNCs. There is no preposition which could be used in paraphrasing these NNCs.

5. Source: How is Y created?

This relation answered the question: *How is Y created?* In this case, Y was the product of X, and X was the source of Y. The NNC could be interpreted as *Y is created by X.*

Examples:

Diffusion current

How is the current created?

→ *The current is created by diffusion.*

Clock pulse

How is the pulse created?

→ *The pulse is created by the clock.*

Quantization error

How is the error created?

→ *The error is created by quantization.*

In terms of prepositional paraphrasing, this relation could be loosely paraphrased with preposition *from*:

Diffusion current: Current from diffusion

Clock pulse: Pulse from the clock

Quantization error: Error from quantization

6. Operation: How does Y work?

Operation is also one of typical relation which could be found in technical language. The modifier X often expressed how the object Y, often a device, worked as in the case of the following examples:

Induction motor:

How does the motor work?

→ *(The motor works) by means of induction.*

Steam turbine:

How does the turbine work?

→ *(The turbine works) by means of steam.*

Correlation receiver:

How does the receiver work?

→ *(The receiver works) by correlating*

It can be seen that the paraphrase forms of these NNCs often go with preposition *by* to show the manner of the operation of the device, but *by* alone could be adequate to paraphrase the NNC. The verb *work* may need to be added for paraphrasing: *motor working by induction,* turbine *working by steam, receiver working by correlating.*

7. Objective

Objective in this taxonomy is grammatical objective. The relation is the case of nominalization or complement structure of paraphrase as Levi (1978) states. When the head noun is denominalized to be back to its verb form, the modifier is the object of that verb. For examples:

Efficiency calculation	*[…]calculate the efficiency*
Rotor construction	*[…]construct the rotor*
Matrix multiplication	*[…]multiply the matrix*

With preposition *of,* these NNCs could be converted to their postmodification forms as follows: *calculation of efficiency, construction of rotor, multiplication of matrix.* As discussed in the shortcomings of preposition paraphrases, however, *of* is an ambiguous preposition which cannot convey clearly the relation of these NNCs.

8. Structure: How is Y structured?

How something is configured, arranged, or organized is the sense the combination carrying this relation express. Y is structured in X manner. The question which could go with this relation is *How is Y structured?* For examples:

Compound generator

How is the generator structured?

→ *(The generator is structured) in a compound manner.*

Block code

How is the code structured?

→ *(The code is structured) in block manner.*

Series combination

How is the combination structured?

→ *(The combination is structured) in series manner.*

In and *of* could be used to change from the premodification form of these NNCs into postmodification form, for example, *generator of compound type, code in block, combination in series.* Among these 8 relations, there are more than one relations which could be paraphrased with preposition *in* and *of,* which is an evidence of the inadequacy of prepositional paraphrases. However, these prepositional paraphrase could be made clear when it is classified under a particular relation, and for teaching purpose, prepositions could still be applied for a simple way of paraphrase which is teachable.

Also from the data under analysis of the study, there is a group of NNCs with a head noun classified as shell nouns (Aktas & Cortes, 2008) such as *design requirements, current components, barrel types, welding process.* Mostly the head noun of these NNCs functions as a cohesion device which refers to the previous or the following part of the text. These NNCs of this group account for only 2.24% in terms of tokens and 4.56% regarding to types.

4.1.3 List of relations resulting from the analysis and their coverage

The following table displays the list of relations resulting from the analysis and how much each relation covers as well the total coverage of the taxonomy:

Overall coverage

Table 4.1 Semantic relation taxonomy and its token coverage

No.	Relations	Tokens	Proportion (%)
1	Location	591	30.06
2	Purpose	306	15.56
3	Measure	274	13.94
4	Representation	182	9.26
5	Source	89	4.53
6	Operation	70	3.56
7	Objective	63	3.20
8	Structure	59	3.00
Total		1615	82.15

Table 4.2 Semantic relation taxonomy and its type coverage

No.	Relations	Number of types	Proportion (%)
1	Location	87	23.32
2	Purpose	65	17.43
3	Measure	37	9.92
4	Representation	24	6.43
5	Objective	23	6.17
6	Source	19	5.09
7	Structure	17	4.56
8	Operation	11	2.95
Total		283	75.87

As pointed out previously, there is no taxonomy which could cover all NNCs. The higher coverage the taxonomy gives, the vaguer it is. With the attempt of balancing between the coverage of the taxonomy and the teachability of each relation, 8 relations which go with 7 questions were suggested. The questions helped to identify the relation or by answering the question, the relation could be identified and the NNC could be correctly paraphrased. For the *objective* relation, there was no question accompanying to clarify the relation. It was the case of nominalization where the head noun was nominalized and the modifier was the object of the process expressed in the head noun such as *data compression (compress the data), signal processing (process the signal), and voltage regulation (regulate the voltage).*

Location ranked number one both in terms of types and tokens. One possible explanation was because in many cases it was important to know the location of some electrical quantities such as voltage, current, power in an electric device such as *output voltage, input power, terminal voltage, node voltage, line current.* The electrical quantities at different location had or needed to have different characteristics which were necessary for the operation of some machine; thus, specifying the location of the quantity was important.

Measure is also a common relation which ranks the third and accounts for 13.9% of all NNCs. It may be brought about by the fact that electrical engineering is a field which is very close to physics and mathematics. Measurement and calculation are very common and related to nearly all activities of electrical engineering. It might be the reason for such high frequency of the relation *measure*.

In comparison with *measure, purpose* had much higher coverage in terms of types (*purpose*: 17.43% and *measure*: 9.92%) but not very higher coverage regarding to tokens (*purpose*: 15.56% and *measure:* 13.94%). The reason could be that there were a large number of combinations with *purpose* relation in the low or medium frequency group. This made the type-token ratio of combinations with *purpose* relation higher than that of combinations with *measure*.

Representation is also a common relation with high coverage in electrical engineering texts which could result from the fact that graphs and diagrams are widely used in the field, as diagram of a circuit, or as a device or graph to describe the relations among different electrical quantities. Because of this fact, such NNCs as *regulation curve, load line, drain characteristics, circuit diagram* and so on occurs in large numbers, which makes *representation* a common relation.

In conclusion, the taxonomy including 8 categories resulted from the balancing between coverage and teachability (one must sacrifice for the other's sake). *Location* was the most common relation. Some relations may have a high coverage in terms of types but low token coverage because of their high type-token ratio. The following part will provide the detail results from different groups of frequency.

Results from analysis of different frequency combination types

High frequency combination types

Table 4.3 Semantic relation taxonomy and its coverage of high frequency group

No.	Relations	Number of tokens/Proportion	Number of types/Proportion
1	Location	437/34.68	24/33.80
2	Measure	208/16.51	9/12.68
3	Purpose	197/15.63	12/16.90
4	Representation	145/11.51	8/11.27
5	Operation	54/4.29	3/4.23
6	Source	41/3.25	2/2.82
7	Objective	20/1.59	1/1.41
8	Structure	10/0.79	1/1.41
Total		1112/88.25	60/84.51

Take relation *location* as an example. The table could be interpreted as follows. The total number tokens of high frequency NNC types which was classified as having *location* relation is 437; this number accounts for 34.68% of the total tokens of high frequency NNC types. There were 24 high frequency combination types which were classified under *location* relation category. These 24 combination types cover 22.80% of the total number of high frequency combination types.

As shown in the table, in general, the order pattern of relations of high frequency group is similar to the overall sample. There are only some minor differences; for example, *measure* ranks higher than *purpose* in terms of tokens which could be explained that the *purpose* relation cover a higher number of tokens in medium and low frequency groups which add up and make it more common in overall.

Medium frequency combination types

Table 4.4 Semantic relation taxonomy and its coverage of medium frequency group

No.	Relations	Number of tokens/proportion	Number of types/proportion
1	Location	102/22.87	23/24.47
2	Purpose	56/12.56	12/12.77
4	Measure	45/10.09	10/10.64
4	Source	39/8.74	9/9.57
5	Structure	36/8.07	6/6.38
6	Objective	24/5.38	6/6.38
7	Representation	23/5.16	6/6.38
8	Operation	9/2.02	2/2.13
Total		334/74.89	74/78.72

Location, purpose and *measure* are still the three most common relations among medium frequency combination types. The biggest difference lies in the frequency of *representation* relation in medium group in comparison with the overall. *Representation* ranks the seventh in this group but it ranks the fourth overall. The reason could lie in the fact that many NNC types of *representation* relation might be among high frequency types.

Low frequency combination types

Table 4.5 Semantic relation taxonomy and its coverage of low frequency group

No.	Relations	Number of tokens/proportion	Number of types/proportion
1	Purpose	53/20.38	41/19.71
2	Location	52/20.00	40/19.23
3	Measure	51/8.08	18/8.65
4	Objective	19/7.31	16/6/69
5	Representation	14/5.38	10/4.81
6	Structure	13/5.00	10/4.81
7	Source	9/3.46	8/3.85
8	Operation	7/2.69	8/2.88
Total		188/72.31	149/71.63

Purpose is the most common relation among low frequency combination types. The number of types of NNCs with *purpose* relation in low frequency group was more than three times comparing to the sum of both high and medium frequency combination types of this relation. Although the order of other relation changed, the three most common relations were still *location, purpose,* and *measure,* which shows the consistency with high frequency, medium frequency groups and with the overall result.

All relations in the overall result were indentified among the high frequency combination types and these high frequency combination types make a large contribution to the ranking of different relations in the whole sample in terms of tokens.

Through the analysis results of three frequency groups, the low frequency group has the lowest coverage both in terms of tokens and types. This is because the type-token ratio of combinations in this group is high. More type could bring about more relations; thus, the relations in this group are very diverse. To reach a higher coverage of this group, more categories need to be created. However, one relation category would not add up a high coverage, and a large number is needed, which results in a large number of relations in the taxonomy. This goes against the teaching purpose.

4.2 Difficulties in technical NNC relation identification encountered in the study

As mentioned in Chapter 1 and 2, there are several factors which lead to the difficulty of interpreting NNCs; one of the most prominent one is the semantic relation which is lost in the process of packaging in technical language. To be able to interpret the NNC, the semantic relation needs to be recovered through unpacking the combination, a process discussed in Ward (2007:26). Unpacking could be by paraphrasing the NNCs in sentences or prepositional phrases. Paraphrasing is one way to exhibit the semantic relation which has been lost in compressing process, but the relation is not always obvious. This is caused by some features of technical texts as well as the development of language itself. Of those features, compression, lexicalization, and naming were frequently encountered in EEC.

4.2.1 Compression

Technical language is characterized by technicality; compression which was widely discussed in Chapter 1 is one factor which realizes the technicality of the specialized text. Because of the compression, the relation of two nouns is not explicitly exhibited, or there is no direct link between the two constituents of the combination. Take *leakage reactance* as an example. From the look at the linguistic surface of this combination, it could be interpreted as *reactance is derived from some leakage*, but leakage itself could not produce any reactance. Leakage is the activity; it is not the agent. In fact, *leakage reactance is the reactance of the leakage flux*. Therefore, the direct link in this case is between *reactance* and *flux*, but *flux* is omitted in the process of compression to form *leakage reactance.* The case of *power line* is similar. *Power line* is *the line which is used for transmitting or distributing power;* then, *power line* is formed from *transmission line* and *power transmission.* This makes it more difficult to identify the direct relation between *line* and *power.* In these cases, compression happens by omitting words in the formation.

4.2.2 Lexicalization

Another factor making semantic relation analysis difficult is the different degrees of lexicalization. Lexicalization is a historical process which may broaden or narrow meanings of the word or phrase (Downing, 1977:821). This process may result in many opaque compounds or combinations (Bauer, 1983:49). The difficulty of a various degree of lexicalization has been pointed out by researchers, and it was pervasively discussed in Downing (1977). The fact is that some studies about semantic relations in NNCs involve only novel combinations (Downing, 1977; Gagne, 2001) actually made up by the author especially for their analysis. The nature of those studies is different from this corpus – based study in which the combinations are extracted from the authentic text, but the different degree of lexicalization is unavoidable. According to Jones (1983), “a compound recalled as familiar by a human being, though it does not figure in an official lexicon, must be treated as lexicalized; and equally, a second occurrence of a novel compound in a text may be treated as lexicalized with respect to a lexicon generated by that text”. A familiar combination must be treated as lexicalized, but to what extent it is lexicalized or how close it is to the “official lexicon” is not easy to identify. The higher the lexicalized degree is, the less obvious the semantic relation. The reason may lie in the process of semantic broadening and narrowing throughout the development of language or in short in the historical process of any language. Examples could be found the data of the study. *Copper loss* is highly lexicalized; wires of some winding, for example the winding in transformers, were made of copper. When power flows in these wires, it dissipates and its value decreases, and the phrase *copper loss* was created. Even though the winding is not made of copper, *copper loss* still refers to the loss in the winding. Thus, the semantic relation exhibit in this example is location (loss in the winding), but *copper* itself does not exhibit the sense of location. Similarly, *air gap* refers to any nonmagnetic discontinuity regardless of whether the gap is filled with air or wood. As Downing (1977) said, there must be reasons for two noun joined together to coin a new combination; originally *air gap* may refer to *gap filled with air* and that is the reason for the combination *air gap* to be formed. However, after undergoing the historical process, the original reason is lost as well as the original relation. *Power factor* is also highly lexicalized and has a conventional use with the reference to a very specific concept, the ratio of real power and apparent power.

4.2.3 Naming

Offset current is one example of naming. As defined in Graf (1999), *offset current* is “the difference in current into the two inputs of an operational amplifier required to bring the output voltage to zero.” From its definition, *offset current* appears to be some way of naming as the relation between offset and current seems to be very weak. In another case, the surface linguistic feature of the combination shows only partly the concept the combination conveys. *Power factor* from the linguistic surface may bring us to the idea of a number which could be multiplied with another number to find out the value of power based on the meaning of factor “a factor of a whole number is a smaller whole number which can be multiplied with another number to produce the first whole number” (Collins COBUILD dictionary). However, this sense seems very far from how *power factor* is actually defined, the ratio of real power and apparent power. There is no linguistic clue to the different types of power (real power and apparent power) in the combination.

All those phenomena in specialized texts make it hard to interpret the NNC. Compression, the issue of technical language, different degree of lexicalization, the issue of corpus study, and naming often come together to make the semantic relation not obvious.

In conclusion, unpacking NNCs could be by means of clausal paraphrases or prepositional paraphrases. However, this process could be hindered by some factors which are common features of technical texts and also identified in this corpus including compression, lexicalization, and naming. These factors also caused difficulties for the analysis. On the basis of teaching purposes, the analysis resulted in a list of eight common relations in this set of data including *location, purpose, measure, representation, source, operation, objective,* and *structure.* Among them, *location, purpose,* and *measure* are three most common in all sub-groups: high frequency, medium frequency, and low frequency, but *purpose* is more frequent relation type in low frequency group than that in medium and high frequency groups.

The next chapter will discuss the teaching implication of the study, more specifically, different types of exercises in different formats and levels to practice students with semantic relations. In the next, limitations, recommendations, and a conclusion of the study will be presented.

CHAPTER 5

CONCLUSION

5.1 Implication for teaching

An eclectic teaching approach will be the guideline for the implication for teaching. The eclectic teaching approach is based on the fact that there are no “best” teaching approaches which can be applied in any classroom (Stern, 1992). Each teaching approach has its own strengths and weaknesses in each particular teaching situation. According to Verghese (2007:55), “a good teacher should, therefore, be eclectic in his approach; he need not accept any theory in toto; he should select what is best suited for his purpose in the classroom.” In the same way, the purpose of the study is not for describing the language, but it is for applications in classroom. Not all relations would be brought into the classroom. Only relations which are prominent and teachable will be included.

Teaching activities and exercises

The following tasks are adapted from Lewis (2002) and Master (2003) and added for the specific teaching purpose of NNCs. Some of tasks could be conducted in the class with teacher-student interaction and some could be individually carried out. By *Task,* in this study, it could be exercises which could be done individually by students or teaching activities which could be conducted with the interaction between teachers and students. All the tasks could be divided into 5 levels: selective attention, recognition, manipulation, interpretation, and production in a classification developed by Paribakht and Wesche (1996) (cited in Wesche and Paribakht, 2000). This classification was proved to facilitate learners' vocabulary growth and originally developed for the single-word item use. Therefore, for case of multiple-word items or NNCs in this study, the classification was adapted to be appropriate for the purpose of the study.

5.1.1 Selective attention

Task 1 Text Search

The student could be provided with a text related to their study (e.g. *generator, electric circuit,* and *power transmission*) and asked to underline, circle or list all NNCs they find in the text. The student can discuss and compare with other students' responses or with the teacher's one. The purpose of this exercise is to make students notice or raise students' awareness of such combinations in technical texts. Such awareness makes this category of lexical item, NNCs in particular, become salient the next time students encounter them (Nation, 2001).

This type of task could be equivalent to the most basic level in Gass's (1988) classification of vocabulary exercise and could be labeled as selective attention type according to Paribakht and Wesche's (2000) topology of text-based vocabulary exercises.

The following is an example of such kind of texts.

Electric generator, also called dynamo: any machine that converts mechanical energy to electricity for transmission and distribution over <u>power lines</u> to domestic, commercial, and industrial customers. Generators also produce the electrical power required for automobiles, aircraft, ships, and trains.

The mechanical power for an electric generator is usually obtained from a rotating shaft and is equal to the shaft torque multiplied by the rotational or angular velocity. The mechanical power may come from a number of sources: hydraulic turbines at dams or waterfalls; wind turbines; steam turbines using steam produced with heat from the combustion of fossil fuels or from nuclear fission; gas turbines burning gas directly in the turbine; or gasoline and diesel engines. The construction and the speed of the generator may vary considerably depending on the characteristics of the mechanical prime mover.

Nearly all generators used to supply electric power networks generate alternating current, which reverses polarity at a fixed frequency (usually 50 or 60 cycles, or double reversals, per second). Since a number of generators are connected into a power network, they must operate at the same frequency for simultaneous generation. They are therefore known as synchronous generators or, in some contexts, alternators.

Source: Electric generator. (2011). In *Encyclopædia Britannica*. Retrieved from http://www.britannica.com/EBchecked/topic/182624/electric-generator

In particular cases, this task could be divided into further steps. First, students are asked to highlight or list all noun phrases in the text including nouns modified by adjectives and by nouns (e.g. *electric generator, mechanical energy, power lines*, *industrial customer, shaft torque, wind turbines,* and *steam turbines*). Second, students are asked to classify those noun phrases into two groups (Y is X, and Y is not X). For example:

Y is X: *electric generator (the generator is electric), mechanical energy (the energy is mechanical)*

Y is not X: *power lines*, *industrial customer, shaft torque, wind turbines,* and *steam turbines*

Third, students are asked to find out the similarity within and difference between groups. It could be noticed that all of "Y is X" group are adjective + noun and most of Y is not X group are noun + noun (except *industrial customer*). It should be pointed out that some NNCs could be interpreted as "*Y is X"* as in the case of *girl friend*, but they do not account for a large number in electrical engineering texts (at least in this corpus). In the case of *industrial customers*, this noun phrase could be interpreted as *customers who work in the industrial sector.* The semantic relation could be identified as the same as the case of NNCs. This noun phrase exhibits an occupation relation. This group of adjectives (named non-predicative adjective by Levi) which cannot be predicative (e.g. *mechanical engineer*) or convey different meanings if they are predicative is not common in the corpus and in many cases they convey the occupation relation with the head noun.

The purpose of this way of conducting the task is to make noun phrases salient and emphasize on the difference between adj + noun and noun + noun. Both ways of carrying out the task could be used as a classroom activity and as an individual exercise.

Task 2 Directed interpretation (Follow-up)

With a list of NNCs extracted from such text as the previous one, the teacher can ask students questions to guide them to interpret the meaning of such combinations. The kind of activity is to direct students' attention to relations underlying between the two constituent nouns and to raise awareness that even the surface structure of those NNCs is similar, but different relations exist. One simple way to explain these different relations is asking "what questions do these combinations answer?" For example, with *power line,* the question which could be asked is *what does the line do?* and one answer which could come up is *distributing power*.

Further examples:

Shaft torque: Where is the torque? In the shaft

Wind turbine: How does the turbine work? It works by means of wind

Fossil fuel: How is the fuel created? Fossil

It should be noted that to answer the question *what does the line do?* is not as direct as to answer *where is the torque?* Or *How is the fuel created?* There is a compression in this combination which is formed from *power transmission* and *transmission line*. The teacher may need to give hints and further explanation in this case.

The purpose of this task is also awareness raising but different from Task 1 in the way that it focuses on underlying semantic relation instead of lexical items itself as in Task 1**.** Because of the interactive nature between teachers and students in this task, it should be a classroom activity.

Task 3 Text Search (Follow – up of Task 1 and 2)

In this activity, the student is asked to group the NNCs from the text based on the question derived from them. This activity helps students to notice differences among different groups and similarities among NNCs of the same group.

Example:

Questions	NNCs
What is the purpose of Y?	*Power line, power network*
Where is Y?	*Shaft torque*
How does Y work?	*Wind turbine, gas turbine, diesel turbine*
How is Y created?	*Fossil fuel*

The list of questions in the table can be extended depending on the text. Therefore, the text should be carefully selected for the purpose of teaching and should make emphasis on the prominent relations. Similar to Task 1, it is a noticing level task and could be an activity in the classroom with student-teacher interaction or an exercise for practice.

Task 4 Categorizing the given NNCs into the table based on the question they answer (Y stands for the second word):

Output voltage, communication system, *core loss*, induction motor, harmonic filter, *input voltage*, *circuit diagram*, *armature current*, *drain current*, *line voltage*, correlation receiver

Questions	NNCs
Where is Y?	
What is the purpose of Y?	
How does Y work?	
What does Y represent?	

As in Task 3, categorizing could help students to see the difference among groups and the similarity within group or may help them to recognize the pattern of members of the same group. This may facilitate the process of identifying and interpreting NNCs later on. Different from Task 3, the list of NNCs in this text is not extracted from a text but a list of independent NNCs which could be from a corpus. Because Task 3 is a follow – up task and is facilitated by Task 2 (directed interpretation with the help of teachers), it is at noticing level. In this task, it requires students to ask and answer by themselves, so it is a higher level in comparison with Task 3.

5.1.2 Recognition

Task 5 Deleting the NNC which is not the same group with the others

Example:

Location: input signal, output signal, source current, logic circuit, field circuit, gate oxide

Purpose: harmonic filter, communication system, voltage source, correlation receiver, magnetizing current

This exercise has the same focus as Task 3. This helps students to be familiar with the process of perceiving different types of relations existing among NNCs. The student is provided with a list of NNCs of the same relation such as location or purpose as in the example. What the student needs to do is to underline or by any means to point out the NNCs which do not have the same relation as others in the list. In the example, to identify location relation, the question "Where is Y?" could be asked. The NNCs which give appropriate answers belong to this group of relation and vice versa.

5.1.3 Manipulation

Task 6 Match NNCs with their appropriate explanations. There are more explanations than NNCs.

1. Source voltage a. Wind for running turbine a.

2. Voltage source b. A source for voltage

3. Wind turbines *c. Voltage for a source*

4. *Fossil fuel* d. Fossils for fuel

e. Fuel from fossils

f. a turbine run by wind

This manipulation task (Paribakht and Wesche's typology) focuses on the order of the word in the combination. The order changes could make the meaning of the whole combination change. This task gets students familiar with indentifying the role of head nouns and modifiers regarding to the combination meaning. This may be more helpful for students whose L1 NNC structure is different from that in English.

5.1.4 Interpretation

Task 7. Choose the correct interpretation of the following NNCs

1. Output voltage

   a. voltage for the output,

   b. *voltage at the output*

   c. voltage which produced by the output

   d. output providing the voltage

For one NNC, there are several interpretations provided but only one interpretation is correct due to the semantic relation between the two nouns and the word order of the NNC. This task requires students to identify the correct interpretation (both prepositional paraphrases and clausal paraphrases could be used as in the example) among all the options of possible ways to interpret it.

Task 8 True or False

a. Input resistance is the resistance provided to the input (F)

b. A diesel turbine is a turbine which produces diesel (F)

c. A power plant is a plant which produces power (T)

d. A power network is a network distributing the power (T)

e. A voltage regulator is a regulator which regulate voltage (T)

f. Control power is power which is controlled (F)

On the same basis as Task 7 in a different format, this task familiarizes students with the idea of semantic relation determining the meaning of combinations. The interpretation of the NNC is provided based on the correct relation or incorrect interpretation. The student needs to identify whether the interpretation is correct or incorrect.

Task 7 and Task 8 are the initial step at the interpretation level. The student has to decide the correct interpretations of the NNC. These two tasks introduce the student to the compression aspect of NNCs and let students learn that there are many ways of interpreting a NNC, but only one is correct.

Task 9 Use one of these prepositions to paraphrase the given NNC

At, in, for, from

Baseband filter, current gain, input signal, electron current, contact potential, gate voltage, line current, transmission line, temperature rise, field current

(*Filter for baseband, gain in the current, signal at input, current from electron, potential in contact, voltage at gate, current in the line, rise in temperature, current in the field*)

Task 9 practices students with prepositional paraphrase of the NNCs. Prepositional paraphrase could be the short form of clausal paraphrase. For example, the clausal paraphrase of *transmission line* could be line which is used for transmitting power; it could be paraphrased as line for transmission with preposition *for.* This task requires students themselves to interpret the NNCs with a given list of prepositions, which could require students' higher interpretation ability.

Task 10 Translate the following NNC into L1:

Power system (*Hệ thống điện*), transmission line (*đường truyền dẫn*), parallel circuit (*mạch điện song song*), depletion region (*vùng hao cạn*), frequency band (*dải tần số*)

This exercise may be helpful for students whose L1 NNC structure is different from English such as Vietnamese. From the writer's own experience, Vietnamese students often fail in terms of identifying head and modifier of the NNC. It may results from the interfere of their first language in which the head noun comes first which is opposite to English. As a result, the modifier becomes the head noun and vice versa when it is translated into Vietnamese.

Task 11 Use a sentence to interpret the following NNCs

Signal source (*source is used for providing signal*)

Voltage regulation (*Voltage is regulated*)

Circuit breaker (*The breaker is used to break the circuit*)

Node voltage (*Voltage is located at the node*)

Quantization noise (*Nose is produced during quantization*)

Diffusion current (*The current is produced by diffusion*)

Compound generator (*The generator is structured in compound manner*)

Relay coil (*The relay contains a coil or the coil is a part of the relay*)

The exercise has a higher requirement of students compared to Task 9. No preposition is provided and a full paraphrase by using sentence is required. A pre-task to train students in clausal paraphrasing could be provided by matching NNCs with their clausal paraphrases or providing the generic paraphrase such as *Y is located in/at X, Y is used to ....X* and asking students to paraphrase the NNCs following that format.

5.1.5 Production

Task 12. Form the correct NNC

a. the current in the winding: *winding current*

b. the circuit which is a part of the field: *field circuit*

c. the current which is produced by the diffusion of carriers: *diffusion current*

d. the carrier which is in a minority: *minority carrier*

This task is the simple exercise at production level. This may help student not only with interpreting the NNC but also producing them in at a low level.

Task 13 Paragraph Headings

Recent research has shown that NNCs are very common use in title of dissertation and journal articles. This activity helps to build up the student's ability to form NNCs and familiarizes them with the convention of forming NNCs as headings. In this task, the student is provided with short passages and asked to write the heading for them with NNCs.

Example:

The mechanical power may come from a number of sources: hydraulic turbines at dams or waterfalls; wind turbines; steam turbines using steam produced with heat from the combustion of fossil fuels or from nuclear fission; gas turbines *burning gas directly in the turbine; or gasoline and diesel engines. The construction and the speed of the generator may vary considerably depending on the characteristics of the mechanical prime mover.*

One possible heading: *mechanical power sources,*

Task 14 Domino games

As its name, it is a game which could be carried out in a competitive manner. Each group is asked to form a chain of NNCs of which the ending word of this NNC is the beginning word of the next one. It could be a time-limited and topic-constrained game (maximum time is specified and all NNCs formed should be under one topic like electric devices, signal modulation) and one requirement is that the group must be able to interpret the NNC they form. The focus of this activity is on the student's production of NNCs.

For example: *output current – current source – source voltage – voltage distortion – distortion current – etc.*

5.2 Conclusion

Despite some shortcomings, in general, this study initially achieved its original goals. As have been discussed in the introduction chapter, NNCs are commonly used in technical English as a means of compressing information. As the result of that process, the semantic relation between the constituent nouns of the combination is lost, and to interpret the combination means to recover this lost semantic relation. The purpose of the study was to investigate the semantic relation of NNCs in electrical engineering in order to create a list of common relations underlying the formation of NNCs in the field with the idea that the knowledge of such common relations could facilitate the process of engineering students' interpretation of NNCs in their textbooks. All NNCs were extracted from the list of noun phrases in an existing electrical engineering textbook corpus and then divided into three groups of high frequency types, medium frequency types, and low frequency types.

There were some difficulties encountered in the process of creating the NNC list which were not anticipated earlier. The original list consists of all types of noun phrases but the target of the current study was only the combination of noun and noun. Therefore, all combinations of adjective nouns needed to be left out. The problems lay in the conversion word-formation which is common English, which means the same word - form could be an adjective in this case or could be a noun in other cases. The predicative test was conducted to make to decision whether to leave out or include the phrase based on the predicative characteristics of the adjective. If the first word of the combination could be predicative, that word is an adjective. The second problem was that the original phrase list includes phrases of which the hyphens have been knocked out. It would cause some difficulties for a NNC to be included or excluded. For example, commonly the adjective would be deleted and the left NNC would be included. However, in some case when there was a hyphen between the adjective and the first noun, the semantic relation was not between the first noun and the second noun, but between the adjective – noun and the head noun. In those cases, the adjective could not be eliminated and the phrase would be classified as three word combinations and left out. Dictionaries could be helpful to determine the existence of hyphens.

After the NNC list was divided into three groups based on their frequency, samples were taken from different groups. The process of semantic relation analysis of these combinations encountered the difficulties brought about by the compression, lexicalization and naming characteristics of such combinations and also the polysemy of each noun constituents as well as the combination as a whole which could be clarified by the context the combination occurs. Another difficulty in creating the relation taxonomy was the conflict between coverage of the relations and the teachability of those relations. The relation categories should be general enough to cover a good number of NNC, but they should be specific enough to be teachable in the classroom.

After all those factors were taken into consideration, a list of 8 relation categories was created. Of them, *location* was the most frequent, and *structure* was the least frequent. The high frequent types cover all of the relations; the analysis of the medium and low frequency combination types help to have better understanding of the commonness of each relation. Each relation went with a question to facilitate the interpretation process. There were some conclusions which could be drawn from the result of the study.

First, the common semantic relations based on which the NNC was created were different among different kinds of texts. The nature of the subject matter may be an important factor which influence on the common semantic relation available on its text. That could be the reason why Master's taxonomy which was designed for technical texts in general could not be applied in this particular text of electrical engineering textbook, and Levi's taxomony for general English could not also have a high coverage. Measure may be common in electrical engineering but no common in chemistry which mostly describes the properties of the chemicals.

Second, underlying semantic relations are variable, and there is no taxonomy of relations which could cover all NNCs. Even the vaguest taxonomy as the Levi's one could not include all the NNCs. One reason could be that it was not created for electrical engineering texts but it could not be denied that it evidences that there is no existing taxonomy which could include in itself all kinds of NNCs and the effort to create such kind of taxonomy seems hopeless and wasteful.

Finally, for the purpose of helping engineering students going through the difficulties caused by NNCs, some common relations were pointed out. Using this relation taxonomy in exercise design could help to familiarize students with the idea of an underlying semantic relation of NNCs and facilitate their process of interpreting NNCs. However, because of the indefiniteness of semantic relations, the relation taxonomy should not be used as a checklist for all NNCs.

5.3 Limitations of the study

This study is somewhat interdisciplinary. It requires both the knowledge of language in terms of semantic relations and the knowledge of the subject matter concerned in this case, it is electrical engineering. Dictionaries or other kinds of references such as textbooks, encyclopedia, and online references could help, but they cannot entirely compensate the poor background knowledge of the writer in the field of electrical engineering. It is a time-consuming process to get both backgrounds in language and in subject matter.

The study much depends on intuition, so the subjectivity of the study is not totally eliminated. There is no reference for any reliability checking instrument used in the literature. Even the rating scale which was used by Chung and Nation (2003) is adapted into use in the study. The reliability could not be ensured because the study requires both linguistic and specific subject matter expertise. The linguistic expert needs to be provided the background in electrical engineering and the subject matter expert needs to be provided the background in language. Both approaches requires immense of time but the reliability is still not be ensured, because the nature of the study is subjective.

Another limitation of the study is that the teaching activities and exercises are merely designed as samples and guidelines. They have been not proved to be efficient in the classroom, which requires a higher level of the study. If the tasks are used as treatment, then the whole experimental research design needs to be set up to identify whether or not there is any significant improvement with or without the treatment, and it requires a whole different study.

In addition, the limitations of the corpus could also bring limitations to the study. As the corpus builder stated, the corpus was built in a small size due to the manual analysis and limited in the books of one publisher due to the copyright issue. The small size of the corpus could limit the chance of a good NNC sample to be taken and affect the occurrence frequency and relation frequency. The books from only one publisher could also have language use bias which could affect the result of the study as well.

5.4 Recommendations

This is only a preliminary study about the semantic relation in electrical engineering textbooks. Further studies could be done with a larger corpus compiled from textbooks of different publishers (the current corpus contains only extracts from books of the same publisher) or in different subfields of electrical engineering. It should be noted that the corpus of electrical engineering used in this study consists of some sub-areas including power transmission, electric devices, communication, system control. Corpora for further study could be compiled from textbooks of such sub-areas as automation, instrumentation or from textbooks of only one sub-area to identify whether or not there are differences among different areas of electrical engineering itself.

As stated earlier, the nature of a subject area could influence the semantic relation. The study could be conducted in different fields such as electronics, chemical engineering, and mechanical engineering. Because the nature of those subject areas is different, the language use will be also different. This fact may affect the source of common semantic relations brought to be used. Those semantic relations to some extent reflect the nature of the area and may help students of the field to improve their reading compression.

One point which should be mentioned is that the current corpus is compiled from textbooks of which the language use is proved to be different from the language use in research articles. It is more fundamental than the language use in research articles, which may lead to one fact that there are more number of high lexicalized NNCs (familiar NNCs and used conventionally without being asked how it was formed). Then, is there any difference between NNCs in textbooks and in research articles regarding to semantic relations?

The study did not cover the combinations of more than 2 noun members because of the recursion of relations in more-than-two-noun combinations. The difficulty of those combinations is in bracketing. More insight into bracketing rules of long combination should be brought into light. It is important to properly interpret or recover the relation of NNC consisting of more than two nouns.

One weakness of the studies is that the exercises have not been tried out with students. A small scale experimental study could be conducted to find out how much they could help student both in terms of improving their interpretation of NNCs and their reading comprehension especially. The teaching activities and exercises could be revised to be more useful in classroom.

REFERENCES

Aderson, R. C., & Freebody, P. (1981). Vocabulary Knowledge. In J. T. Guthrie (Ed.), *Comprehension and Teaching: Research Reviews* (pp. 77 - 117). Newar: International Reading Association.

Aktas, R. N., & Cortes, V. (2008). Shell nouns as cohesive devices in published and ESL student writing. *Journal of English for Academic Purposes, 7*(1), 3-14.

Anwar, A. A. (1994). The effect of knowledge-map and underlining training on the comprehension of scientific texts. *English for Specific Purposes*, *13(1),* 35–45.

Banks, D. (2001). The reification of scientific process: The development of grammatical metaphor in scientific discourse. In F. Mayer (Ed.), *Language for special purposes: perspectives for the new millennium* (Vol. 1, pp. 555-563). Tübingen: Gunter Narr.

Banks, D. (2005). On the historical origins of nominalized process in scientific text. *English for Specific Purposes, 24*(3), 347-357.

Bauer, L. (1983). *English word-formation*. Cambridge: Cambridge University Press.

Bauer, L. (1998). When is a sequence of two nouns a compound in English. *English Language and Linguistics 2*(1), 65-86

Bauer, L. (2006). Compound. In K. Brown (Ed.), *Encyclopedia of Language and Linguistics* (Vol. 2, pp. 719 - 726). Oxford: Elsevier.

Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes, 26*(3), 263-286.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

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

Cohen, A., Glasman, H., Rosenbaum-Cohen, P., Ferrara, J., & Fine, J. (1988). Reading English for specialized purposes: Discourse analysis and the use of student informants. In P. Carrell, J. Devine & D. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 152–167). Cambridge: Cambridge University Press.

Downing, A., & Locke, P. (2002). *A university course in English grammar*. London: Routledge.

Downing, P. (1977). On the creation and use of English compound nouns. *Language, 53*(4), 810-842.

Fabre, C. (1996). *Interpretation of nominal compounds: combining domain-independent and domain- specific information.* Paper presented at the Proceedings of COLING-96, Copenhagen.

Farrell, P. (1990). *Vocabulary in ESP: A lexical analysis of the English of electronics and a study of semi-technical vocabulary.* CLCS Occasional Paper No. 25 Trinity College.

Fender, M. (2001). A Review of L1 and L2/ESL Word Integration Skills and the Nature of L2/ESL Word Integration Development Involved in Lower-Level Text Processing. *Language Learning, 51*(2), 319-396.

Fender, M. (2003). English word recognition and word integration skills of native Arabic- and Japanese-speaking learners of English as a second language. *Applied Psycholinguistics, 24*(02), 289-315.

Fender, M. (2008). Arabic literacy development and cross-linguistic effects in subsequent L2 literacy development. In A. M. Z. Keiko Koda (Ed.), *Learning to read across languages: cross-linguistic relationships in first- and second-language literacy development* (pp. 101 - 124). New York, NY: Routledge.

Fernández-Domínguez, J. (2009). Productivity in English word-formation: an approach to n+n compounding. Bern: Peter Lang.

Finin, T. W. (1980). *The semantic interpretation of compound nominals.* University of Illinois at Urbana-Champaign, Champaign, IL.

Flynn, S. (1989). The role of the head-initial/head-final parameter in the acquisition of English relative clauses by adult Spanish and Japansese speakers. In S. M. Gass & J. Schachter (Eds.), *Linguistic perspectives on second language aquisition* (pp. 89-108). Cambridge: Cambridge University Press.

Gagné, C. L. (2001). Relation and Lexical Priming During the Interpretation of Noun-Noun Combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(1), 236-254.

Gagné, C. L. (2002). The competition-among-relations-in-nominals theory of conceptual combination: implications for stimulus class formation and class expansion. *Journal of the Experimental Analysis of Behavior, 78*(3), 551–565.

Gagné, C. L., & Shoben, E. J. (1997). Influence of Thematic Relations on the Comprehension of Modifier-Noun Combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(1), 71-87.

Gass, S.M. (1988). Second language vocabulary acquisition. *Annual Review of Applied Lingusitic, 9,* 92-106.

Gass, S. M. (1989). How do learners resolve linguistic conflicts? In S. M. Gass & J. Schachter (Eds.), *Linguistic perspectives on second language acquisition* (pp. 183-199). Cambridge: Cambridge University Press.

Girju, R., Moldovan, D., Tatu, M., & Antohe, D. (2005). On the semantics of noun compounds. *Comput. Speech Lang., 19*(4), 479-496.

Graf, R. F. (1999). Modern Dictionary of Electronics. Boston: Newnes.

Greenbaum, S. & Quirk, R. (1990). A student's grammar of the English language. Essex: Longman.

Halliday, M. A. K. (1985). *An introduction to functional grammar*. London: Arnold.

Halliday, M. A. K. (1993). The analysis of scientific texts in English and Chinese. In M. A. K. Halliday & J. R. Martin (Eds.), *Writing science: literacy and discursive power* (pp. 124 - 132). Oxon: Routledge.

Halliday, M. A. K. (2004). Some grammatical problems in scientific English. In C. Coffin, A. Hewings & K. O'Halloran (Eds.), *Applying English grammar: functional and corpus approach* (pp. 77 - 94). London: Arnold.

Hiebert, E. H., & Kamil, M. L. (2005). *Teaching and learning vocabulary: bringing research to practice*. Mahwah, NJ: Lawrence Erlbaum.

Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes, 27*(1), 4 - 28.

Jones, K. S. (1983). "*Compound noun interpretation problems*." Technical Report 45, Computer Laboratory, Cambridge University, Cambridge.

Koda, K. (1989). The Effects of Transferred Vocabulary Knowledge on the Development of L2 Reading Proficiency. *Foreign Language Annals, 22*(6), 529-540.

Koda, K. (2005). *Insights into second language reading*. Cambridge: Cambrigde University Press.

Lauer, M. (1995). Designing statistical language learners: experiments on noun compounds. Macquarie University, Australia.

Laufer, B. (1992). Reading in a foreign langage: how does L2 lexical knowledge interact witht the reader's general academic ability? . *Journal of Research in Reading, 15*(2), 95 - 103.

Laufer, B., & Sim, D. D. (1985). Measuring and explaining the threshold needed for English for Academic Purposes texts. *Foreing Language Annuals, 18*, 405 - 413.

Lees, R. B. (1960). *The grammar of english nominalizations*. The Hague: Mouton de Gruyter.

Levi, J. M. (1978). *The Syntax and Semantics of Complex Nominals*. New York: Academic Press.

Lewis, M. (2002). The lexical approach : the state of ELT and a way forward. Boston, MA: Thomson Corporation.

Lewis, M., & Gough, C. (2002). Implementing the lexical approach : putting theory into practice. England: LTP Language.

Maguire, P., Maguire, R., & Cater, A. (2008). *Factors Influencing the Interpretation of Noun-Noun Compounds.* Paper presented at the Cognitive Science Conference, Washington DC.

Master, P. (2003). Noun compounds and compressed definitions. *English Teaching Forum, 41*(3), 2-9.

Mezynski, K. (1983). Issues Concerning the Acquisition of Knowledge: Effects of Vocabulary Training on Reading Comprehension. *Review of Educational Research, 53*(2), 253-279.

Moon, R. (1997). Vocabulary Connection: multi-word items in English. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, Acquisition and Pedagory* (pp. 40 - 63). Cambridge: Cambridge University Press.

Neter, J., Wasserman, W., & Whitmore, G. A. (1988). *Applied statistics*. Boston: Allyn and Bacon.

Nguyen, H.-C. (2006). Tình hình dạy và học ngoại ngữ chuyên ngành ở các trường đại học. *Thông tin khoa học xã hội, 11*, 13-16.

Norman, J. G. (2003). Consistent naming in scientific writing: sound advice or shibboleth? *English for Specific Purposes, 22*(2), 113-130.

Pastor, L. C. (2008). English complex noun phrase interpretation by Spanish learners. *RESLA, 21*, 27-44.

Pritchard, R. M. O., & Nasr, A. (2004). Improving reading performance among Egyptian engineering students: principles and practice. *English for Specific Purposes, 23*(4), 425-445.

Pueyo, I. G., & Val, S. (1996). The construction of technicality in the field of plastics: A functional approach towards teaching technical terminology. *English for Specific Purposes, 15*(4), 251-278.

Qian, D. D. (1999). Assessing the Roles of Depth and Breadth of Vocabulary Knowledge in Reading Comprehension *Canadian Modern Language Review/ La Revue canadienne des langues vivantes, 56*(2), 282-308.

Quirk, R., Greenbaum, S., Leech, G., & Svartvi, J. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.

Raffray, C. N., Pickering, M. J., & Branigan, H. P. (2007). Priming the interpretation of noun-noun combinations. *Journal of Memory and Language, 57*(3), 380-395.

Ravelli, L. J. (1988). Grammatical Metaphor: An initial analysis. In E. Steiner & R. Veltman (Eds.), *Pragmatics, discourse and text: some systemically-inspired approaches* (pp. 133-147). London: Pinter.

Rutherford, W. (1989). Interlanguage and pragmatic word order. In S. M. Gass & J. Schachter (Eds.), *Linguistic perspectives on second language acquisition* (pp. 163-182). Cambridge: Cambridge University Press.

Ryder, M. E. (1994). *Ordered chaos: the interpretation of English noun-noun compounds*. Los Angeles, CA: University of California Pres.

Sáenz, F. S. (2000). Halliday's grammatical metaphor, conceptualization and linguistic construal. *EPSO, 16*, 497-511.

Salager, F. (1983). The lexis of fundamental medical English: classification framework and rhetorical function--a statistical approach. *Reading in a Foreign Language 1*(1), 54-66.

Salager-Meyer, F. (1984). Compound nominal phrases in scientific-technical literature: proportion and rationale. In A. K. Pugh & J. M. Ulijin (Eds.), *Reading for professional purposes – studies and practices in native and foreign languages* (pp. 138–149). London: Heinemann.

Schmitt, N. (2000). *Vocabulary in Language Teaching*. Cambridge: Cambirdge University Press.

Stern, H. H. (1992). *Issues and options in language teaching*. Oxford: Oxford University Press.

Ulijn, J., M., & Strother, J., B. . (1990). The effect of syntactic simplification on reading EST texts as L1 and L2. *Journal of Research in Reading, 13*(1), 38-54.

Verghese, C. P. (2007). *Teaching English as a Second Language*. New Delhi: Sterling publisher

Wang, M., Koda, K., & Perfetti, C. A. (2003). Alphabetic and nonalphabetic L1 effects in English word identification: a comparison of Korean and Chinese English L2 learners. *Cognition, 80*(1), 129-149.

Ward, J. (2007). Collocation and technicality in EAP engineering. *Journal of English for Academic Purposes, 6*(1), 18-35.

Ward, J. W. (2005). The lexical aspect of reading English as a foreign language for engineering undergraduates. Unpublished doctoral thesis, Birmingham University

Warren, B. (1978). Semantic Patterns of Noun-Noun Compounds. *Gothenburg Studies in English 41*, Goteburg, Acta Universtatis Gothoburgensis.

Wasuntarasophit, S. (2008). *Techinical and Academic Vocabulary in Electrical Engineering Textbooks.* Suranaree University of Technology, Nakhon Ratchasima.

Weiskopf, D. (2007). Compound Nominals, Context, and Compositionality. *Synthese, 156*(1), 161-204.

Wesche, M. B., & Paribakht, T. S. (2000). Reading-Based Exercises in Second Language Vocabulary Learning: An Introspective Study. *The Modern Language Journal, 84*(2), 196-213.

Wilkins, D. A. (1972). *Linguistics in Language Teaching*. Cambridge, MA: MIT Press.

Willis, D. (1990). The lexical syllabus : an new approach to language teaching. London: HarperCollins.

Yang, H. (1986). A New Technique for Identifying Scientific/Technical Terms and Describing Science Texts *Literary and Linguistic Computing, 1*(2), 93-103.

Zhang, L. J., & Anual, S. B. (2008). The Role of Vocabulary in Reading Comprehension: The Case of Secondary School Students Learning English in Singapore. *RELC Journal, 39*(1), 51 - 76.

CURRICULUM VITAE

Nguyen Mai Linh is a lecturer who teaches ESP at the Faculty of Information Technology, Thai Nguyen University in Thai Nguyen, Vietnam. She got her B.A. in English for Science and Technology from Hanoi University of Technology, Hanoi, Vietnam in 2008. In the same year, she started her M.A. in English Language Studies at School of English, Institute of Social Technology, Suranaree University of Technology in Nakhon Ratchasima, Thailand. Her research interests include vocabulary, corpus linguistics, technology-enhanced language learning.