CREATING A KOREAN ENGINEERING ACADEMIC VOCABULARY LIST (KEAVL):

COMPUTATIONAL APPROACH

A Thesis Submitted to the

College of Graduate and Postdoctoral Studies

In Partial Fulfillment of the Requirements

For the Degree of Master of Arts

In the Department of Linguistics

University of Saskatchewan

Saskatoon

By

OLGA KRIUKOVA

Permission to use

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis or, in their absence, by the Head of the Department or the Dean of the College in which my thesis was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use, which may be made of any material in my thesis.

Requests for permission to copy or to make other uses of materials in this thesis in whole or part should be addressed to:

Department of Linguistics
College of Arts & Science
University of Saskatchewan
908 Arts Building, 9 Campus Drive
Saskatoon, Saskatchewan S7N 5C9
Canada

Or

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan S7N 5A5
Canada

Abstract

With a growing number of international students in South Korea, the need for developing materials to study Korean for academic purposes is becoming increasingly pressing. According to statistics, engineering colleges in Korea attract the largest number of international students (Korean National Institute for International Education, 2018). However, despite the availability of technical vocabulary lists for some engineering sub-fields, a list of vocabulary common for the majority of the engineering sub-fields has not yet been built. Therefore, this study was aimed at creating a list of Korean academic vocabulary of engineering for non-native Korean speakers that may help future or first-year engineering students and engineers working in Korea.

In order to compile this list, a corpus of Korean textbooks and research articles of 12 major engineering sub-fields, named as the Corpus of Korean Engineering Academic Texts (CKEAT), was compiled. Then, in order to analyze the corpus and compile the preliminary list, I designed a *Python*-based tool called *KWordList*. The *KWordList* lemmatizes all words in the corpus while excluding general Korean vocabulary included in the Korean Learner's List (Jo, 2003). Then, for the remaining words, KWordList calculates the range, frequency, and dispersion (in this study deviation of proportions or DP (Gries, 2008)) and excludes words that do not pass the study's criteria (range $\geq$ 6, frequency $\geq$ 100, DP $\leq$ 0.5).

The final version of the list, called Korean Engineering Academic Vocabulary List or KEAVL, includes 830 lemmas (318 of intermediate level and 512 of advanced level). For each word, the collocations that occur more than 30 times in the corpus are provided.

The comparison of the coverage of the Korean Academic Vocabulary List (Shin, 2004) and KEAVL based on the Corpus of Korean Engineering Academic Texts showed that KEAVL covers more lemmas in the corpus. Moreover, only 313 lemmas from the Korean Academic

Vocabulary List (Shin, 2004) passed the criteria of the study. Therefore, KEAVL may be more efficient for engineering students' vocabulary training than the Korean Academic Vocabulary List and may be used for the engineering Korean teaching materials and curriculum development. Moreover, the *KWordList* program written for the study can be used by other researchers, teachers, and even students and is open access (https://github.com/HelgaKr/KWordList).

Acknowledgments

     I would like to thank my supervisor, Dr. Zhi Li, for his continuous support and invaluable guidance throughout the research.

     I am also grateful to my committee members Dr. Veronika Makarova, Dr. Jesse Stewart, and Dr. Hyunjung Shin, for their insightful comments, recommendations, and feedback.

     Last but not least, I am thankful to my family and especially to my husband for their support and encouragement in every step.

Table of contents

List of Tables

List of Figures

List of Abbreviations

AV – Academic Vocabulary

AWL – Academic Word List

CKEAT – Corpus of Korean Engineering Academic Texts

DECL – Declarative marker

DP – Deviation of Proportions

EFN – Declarative marker (*Kkma* tag)

GSL – General Service List

HON – Honorific marker

KAP – Korean for Academic Purposes

KAVL – Korean Academic Vocabulary List

KCI – Korean Citation Index

KEAVL – Korean Engineering Academic Vocabulary List

KLL – Korean Learners' List

LOC – Locative marker

MAG – Adverb (*Mecab* tag)

NNG – Noun (*Kkma* and *Mecab* tag)

PST – Past tense marker

POS – Part of Speech

PVL – Physics Vocabulary List

SK – Sino-Korean

TOP – Topic marker

VA – Adjective (*Mecab* tag)

VV – Verb (*Mecab* tag)

VX – Auxiliary predicate (*Mecab* tag)

XSA – Adjectival derivational suffix (*Kkma* tag)

XSV – Verbal derivational suffix (*Kkma* tag)

1 Introduction

There are more and more students worldwide who decide to receive their degree abroad. While the most substantial number of international students choose English-speaking countries, the Republic of Korea has recently increased in popularity for international students due to the variety of governmental and university grants and relatively affordable costs. According to the Korean National Institute for International Education (NIIED), the number of international shows a stable increase since 2010 and almost doubled since 2014 (from 80,000 to almost 150,000) (Korean National Institute for International Education, 2018). Moreover, the target of the Korean government is to enroll 200,000 international students by 2023.

However, for international students in Korea, studying in Korean universities requires not only general language skills but also academic language proficiency. Therefore, such an increase in the number of non-native Korean students and the government's further plans to attract even more cannot but give a significant impetus to the development of Korean language education, especially, Korean for academic purposes (hereafter KAP) (Lee, 2018). Indeed, in recent years, many researchers emphasized the growing need for KAP materials (Kim, 2003; Kim, 2005; Lee, 2004), and as a result, more and more research studies related to KAP, KAP teaching methodology, and study materials development have been published (Han, 2010; Lee, 2018).

An integral part of the materials development for language for academic purposes is the construction of word lists (Coxhead, 2000), and more importantly, field-specific word lists (Hyland & Tse, 2007). Word lists have significant implications for language learning and teaching for several reasons. First, academic word lists prioritize particular vocabulary, so students know which words they should learn first and which one they can learn later. Second, they may set a vocabulary focus for textbooks and study materials, and, as a result, help their authors make these

materials more efficient. Third, they help teachers construct their curricula and lesson plans with attention to the vocabulary their students need the most. Hence, they are essential for effective language training (Coxhead, 2011). Thus, this thesis is an attempt to contribute to the KAP by compiling the Academic Word List of engineering Korean.

The field of engineering was chosen for two main reasons in this study. First, according to the statistics reported in a Case Study on Support of International Students in Korea (Korean National Institute for International Education, 2018), engineering programs are in the highest demand among international students in Korea. However, it is essential to mention that after the Korean government decided to pursue a policy of increasing global competitiveness by the introduction of English as a Medium of Instruction (EMI), most of the programs in engineering departments are taught in English. In this case, it is natural to ask why we need a list of Korean words for the engineering field then. The answer is that even though the teaching is mostly conducted in English, many studies report that teachers quite often explain certain topics in Korean, as well as use Korean words in their presentations (Kim, Kweon, & Kim, 2017). Park and Min's (2014) study showed that 50% of teachers in engineering programs used both English and Korean during their EMI classes, and 38.2% of teachers used Korean more than English, because of the insufficient English proficiency of some Korean students to cover all the materials without any comments made in their L1. Moreover, another study shows that Korean is often being used by lectors to tell engineering-related stories, jokes, or to create a better connection with students (Oh & Kim, 2012). Therefore, although it is stated that the main language of instruction in engineering classes is English, instructors often switch to Korean. Hence, international students undoubtedly require some engineering vocabulary training in Korean as well to understand some explanations and comments and to feel included in the discussion. Furthermore, even students with high

proficiency in general Korean still require some additional language training because of the very field-specific vocabulary (Jeong, 2018; Shin, 2003).

The second reason for choosing the engineering field is the fact that Korea has such engineering giants as *Samsung*[1], *Hyundai*[2], *Hanwha*[3], *Daewoo*[4], *Dohwa*[5], and many others (all these conglomerates have affiliates in different engineering fields). Therefore, it is natural that some international students after graduation would like to work in these world-famous companies, where knowing Korean engineering vocabulary may be essential. Moreover, while working in *Samsung*, I met many engineers working in the abroad office (*Samsung Research and Development Center* in Moscow, Russia in this case) who wanted to know more engineering vocabulary in Korean in order to be able to communicate more clearly with Korean colleagues during frequent business trips to Korea. Hence, such a vocabulary list may be beneficial for people who are willing to work in Korea in engineering companies and/or research institutes or for international scientists in Korean engineering company branches.

However, despite the large number of international students and researchers who may be benefited from the field-specific language training, there are still few materials on the engineering vocabulary of Korean. Existing word lists are focused on narrow fields of engineering (e.g., mechanical engineering, see Lee, 2014, textile engineering, see Nergui, 2015). At the same time, there is no engineering word list that combines the essential vocabulary for the field of engineering as a whole (like the Student Engineering Word List for the English language, see Mudraya, 2006 or Basic Engineering English Word List, see Ward, 2009).

---

[1] South Korean conglomerate with many industrial affiliates like Samsung Electronics, Samsung Heavy Industries, Samsung Engineering.

[2] South Korean automotive manufacturer. Third largest vehicle manufacturer in the world.

[3] South Korean conglomerate with many such affiliates as Hanhwa Chemical, Hanhwa Aerospace, Hanhwa Engineering and Construction and others.

[4] Large South Korean electronics company.

[5] Large engineering and construction company.

Although word lists that have narrow vocabulary focus are also crucial, the academic vocabulary list for the whole field may be very useful for those who start their education in engineering or only preparing to apply to an engineering college and thus do not have a narrow specialization yet. Moreover, it can help engineering students/professionals who work in several engineering sub-fields. The development of the word list of engineering academic Korean may not only assist current international students and engineers but, more importantly, contribute to further development of the engineering language training and studies of engineering vocabulary of Korean.

To conduct the study, computational linguistic methods were used. First, the corpus of engineering academic texts was created. Then the corpus was processed with the script written with the *Python* programming language that produced the word list.

This thesis uses the *Revised Romanization of Korean*[6] for the transliteration of Korean names and article titles and *Yale Romanization*[7] for the transliteration of linguistic data.

In the following sections, the literature review, rationale of the study, detailed methodology explanation, report, and discussion of results will be presented.

---

[6] Revised Romanization of Korean is an official romanization system in South Korea developed by National Institute of Korean Language.

[7] Yale Romanization is a romanization system of Korean that was developed in Yale University by the professor S. E. Martin and his colleagues. This romanization is a standard in linguistics for the Korean language because of its emphasis on morphophonemic structure of words.

2   Literature Review

A large body of literature on vocabulary types and word list compilations for English and Korean provides a basis for the present research. In this chapter, the main studies in this field will be examined. Moreover, since the current study aims to build word lists for the Korean language, it is essential to provide brief details about Korean language specifics crucial for this study.

2.1   Korean language features

Since Korean is an agglutinative language, words often attach multiple morphemes that modify its meaning. Its inflectional system is very rich. For instance, nouns can attach postpositions, prefixes, and suffixes, and predicates can attach tense markers, honorifics, and clause type markers (Sohn, 1999).

The word can attach several morphemes in a time building complex inflections. In the case of nouns, morphemes attached to the stem, for instance, *kyosil (教室)*[8]*:*

1) *kyosil - ey*
   classroom - LOC
2) *kyosil - ey - nun*
   classroom - LOC - TOP
3) *kyosil - ey - to*
   classroom - LOC - also

In the case of predicates, morphemes are attached to the stem after the removal of infinitive marker *-ta*, for instance, the verb *cek-ta*:

4) *cek - ess - ta*
   write down - PST - DECL
5) *cek-usi-ess-ta*
   write down - HON - PST - DECL

---

[8] In the present study, the etymology of the Sino-Korean vocabulary will be presented in the parentheses.

The number of such possible inflection combinations in Korean is quite high, and for effective lemmatization of Korean, morphological tagging of words is first required (Han & Palmer, 2005). More information on Korean language taggers is provided in Chapter 2.

Another feature of the Korean language that is essential for this study is Sino-Korean (hereafter SK) verbs and adjectives, which are usually called verbal and adjectival nouns (Sohn, 1999). Their main feature is that they cannot function as predicates in Korean without certain modifications. To make them functional, they must be compounded with Korean predicates such as verb *hata* "to do/to be", verb *toyta* "to become", and the copula *-ita* or causative affix *-sikhita*. Hence, these manipulations make verbal and adjectival SK nouns function as predicates in Korean (Sohn, 1999), for example:

6) *yenkwu (研究)* 'research' – *yenkwu-**hata*** 'to research',

7) *nonuy (論議)* 'discussion, debate' – *nonuy-**hata*** 'to discuss, to debate',

8) *cwunpi (準備)* 'preparation' – *cwunpi-**toyta*** 'to be prepared',

9) *hayngpok (幸福)* 'happiness' – *hayngpok-**hata*** 'happy'.

Among the examples when the copula *ita* is used with SK nouns, it is necessary to mention forms ending with *-cek (的)* suffix. The SK nouns that have this suffix attached to the stem change their meaning to descriptive that can be translated like "related to the [noun]/having properties of [noun]". In order for them to function as a predicate in Korean, the copula *ita* is being added to the stem, for instance:

10) *mwunhwa (文化)*'culture' -> *mwunhwa-**cek*** 'SK-be cultural' ->

   *mwunhwacek-**ita*** 'be cultural',

11) *kyengcey (經濟)* 'economy' -> *kyengcey-**cek*** 'SK-financial' -> *kyengceycek-**ita*** 'financial'.

6

Hence, considering the fact that a learner has a basic knowledge of Korean grammar, it is easy for them to translate a predicate form with *hata*, *toyta*, *sikhita, ita* by knowing the SK noun or SK verbal/ adjectival noun meaning.

The notable exception is the single-syllable SK adjectival and verbal nouns, which in Korean usually appear in the form of a Korean predicate that is not used as a noun (with some rare exceptions) (Sohn, 1999), for instance:

12) *yak-hata (弱)* 'weak',

13) *phyo-hata (表)* 'to express, to show'.

These features of SK words are important in the light of the topic of this thesis because the existing Korean language taggers always divide the Korean verb or copula and SK noun (with >1 syllables), while Korean predicates based on one-syllable SK nouns are treated by the taggers as a verb. For instance, the results of tagging with *Kkma* tagger[9] will show (done in *Python* console):

14) kkma.pos("연구한다") => [('연구', **'NNG'**), ('하', 'XSV'), ('ㄴ다', 'EFN')][10]
      yenkwuhanta => [('yenkwu', **'NNG'**), ('ha', 'XSV'), ('nta', 'EFN')].
15) kkma.pos("행복하다") => [('행복', **'NNG'**), ('하', 'XSA'), ('다', 'EFN')]
      hayngpokhata => [('hayngpok', **'NNG'**), ('ha', 'XSA'), ('ta', 'EFN')],

but:
16) kkma.pos("표한다") => [('표하', **'VV'**), ('ㄴ다', 'EFN')]
      phyohanta => [('phyoha', **'VV'**), ('nta', 'EFN')],
17) kkma.pos("약하다") => [('약하', **'VA'**), ('다', 'EFN')]
      yakhata => [('yakha', **'VA'**), ('ta', 'EFN')].

This feature of taggers is essential when comparing a list of words lemmatized by the tagger to other lists because they may contain SK nouns in predicative form, while the tagger lemmatized

---

[9] **Part-of-speech tagger or tagger** – is a tool that assigns a particular part of speech to each word in a piece of text (or corpus).
[10] The following *Kkma* tagger abbreviations are used in the examples 14-17: NNG – noun, XSV – verbal derivational suffix, EFN – declarative marker, XSA – adjectival derivational suffix, VV – verb, VA – adjective.

words will contain them in noun form. Thus, in this kind of situation, the comparison will give incorrect results. Therefore, while comparing two Korean word lists to each other (computationally), it is crucial to ensure that both lists use the same form for SK words.

2.2 Vocabulary types

As mentioned in the previous section, it is important to understand the specific features of the language for which the word list is being created. Nevertheless, it is also crucial to know for which purposes and for whom this list is being developed since not all vocabulary is equally necessary for different language needs. That is, some vocabulary that is important for our daily life at the same time may be useless if we are reading professional literature. Hence, language speakers are continually dealing with different types of vocabulary. Therefore, linguists have categorized vocabulary. The frequency of a word and the range of its occurrence were considered the main factors of distinguishing vocabulary levels (Nation, 2013). Nation (2013) is often accredited for the most widely used division, where he states three overall major groups of vocabulary, which are high-frequency, mid-frequency, low-frequency, and two groups of specialized vocabulary, which are academic and technical in nature. Usually, the first three groups are considered when working with general corpora, and the last two are considered when working with certain kinds of texts (academic articles, newspaper articles etc.).

Word lists usually aim to cover basic vocabulary (or high-frequency words), academic vocabulary, or technical vocabulary. In the following sections, detailed information about these three major vocabulary types is provided.

2.2.1 *Basic Vocabulary (high-frequency words)*

It seems natural that researchers tried to identify what is the basic lexicon that every language learner must know. This concept is usually referred to as *basic vocabulary* or *core*

vocabulary (Carter, 1998). Nation and Waring (1997) refer to it as "common vocabulary items that frequently occur across different texts". Im (1991) also mentions basic vocabulary as minimal essential words that are frequently used in our daily life and are widely distributed. Moreover, Im points out that compound words are often built out of basic vocabulary. A basic vocabulary usually covers a substantial proportion of the texts' tokens (Nation, 2001). Therefore, it is considered that a lack of knowledge of a basic vocabulary may seriously affect language understanding and production (Nation & Waring, 1997).

In 1953, West had developed the "General Service List" (GSL) that aimed to state the most frequent and, hence, useful English words for a learner. The list contained 2,000- word families and was based on frequency criteria. It became a foundation that teachers and students could rely on for English teaching and learning, and a starting point for other vocabulary lists construction (academic, technical vocabulary lists), playing the role of a filter of general words. (more about this list in a Section 2.3.1)

2.2.2   Academic Vocabulary

The second type of vocabulary that was mentioned earlier (Section 2.2) is an academic vocabulary. Academic vocabulary is a "formal, context-independent words with a high frequency and/or wide range of occurrence across scientific disciplines, not usually found in basic general English courses" or "words with high frequency across scientific disciplines" (Farrell, 1990, p. 11). Coxhead and Nation (2001) also specify academic vocabulary as words that may be frequently observed in academic texts, but that are rare in other kinds of texts. Some researchers also referred to academic vocabulary as *semi-technical* vocabulary (Farrell, 1990), *subtechnical* vocabulary (Anderson, 1980), or *specialized nontechnical* lexis (Cohen, Glasman, Rosenbaum-Cohen, Ferrara, & Fine, 1979).

Scholars usually speak about academic vocabulary in the context of the language for academic purposes, or language for specific purposes. One of the reasons why scholars and teachers pay attention to academic vocabulary is that it is considered to help students, especially non-native speakers, to adapt to academic life in a university. Naturally, a lack of confidence in the academic genre of language, and especially its' vocabulary, may lead to mistakes in their writings, misunderstanding in readings, stress and even difficulties in identifying themselves as a part of a university/academic social group (Wray, 2002). Moreover, Santos's (1988) study shows that lexical errors in English were considered the most serious by professors and that evaluation of writing content may be affected by word choice mistakes.

Considering the above, we can state that academic vocabulary is essential for students and should be taught. For these purposes, many special teaching materials, curricula, and learning strategies have been constructed. Usually, these materials are focused on the academic vocabulary common for different scientific disciplines and the academic vocabulary typical for specific fields (Tajino, Dalsky, & Yosuke, 2009).

Nevertheless, some scholars warn that despite the fact that the concept of academic vocabulary is convenient to specify a general difference between academic register and general language, it may also lead to an opinion that there is only a small set of vocabulary that may be used in academic language regardless of the academic field (Hyland & Tse, 2007). Hyland and Tse also highlight that it is better to have a "variety of subject-specific literacies"(Hyland & Tse, 2007, p. 247), because there are many words that have different meanings across registers (Biber, Johansson, Leech, Conrad, & Finegan, 1999). Moreover, different fields of science have their own writing styles, and ways of explaining and argumentation (Hyland, 2001), in addition to collocational patterns (Martínez, Beck, & Panza, 2009). Wang, Liang, and Ge (2008) mention this particular field's academic vocabulary type as an academic vocabulary of a single discipline.

Nowadays, AV lists for certain disciplines are being built more often (see more details in Section 2.3.1). Moreover, it is important to mention that the above-mentioned lists contain exactly the academic vocabulary and not the technical vocabulary that is described in the next section (2.2.3).

2.2.3   Technical vocabulary

The third vocabulary type is highly specialized and called *technical vocabulary*. While academic vocabulary is common across most academic texts or a certain discipline, technical vocabulary is used only in certain scientific fields or even sub-fields (Chung & Nation, 2004). For instance, certain words may be used in the field of architectural engineering often, but may not occur as frequently in the field of computer engineering and vice versa. Therefore, this type of vocabulary is commonly referred to as *vocabulary for specific purposes* (Chung & Nation, 2004), *specialized lexis* (Baker, 1988), *technical words* (Farrell, 1990).

However, even though the technical vocabulary is a narrow subtype of vocabulary, it is not easy to tell if a word is technical or not as it may overlap with academic or even basic vocabulary (Chung & Nation, 2004; Mudraya, 2006; Nation, 2013, p. 20). Therefore, many linguists are still trying to identify more precise criteria to distinguish these vocabulary subtypes.

2.3   Previous studies on the development of word lists

Word lists aim to represent a set of a specific vocabulary. In order to make a word list, researchers use a corpus often divided into sub-corpora. Lists may be based on one or more of the following criteria: frequency, range, and/or coverage (Nation, 2013). The criterion of frequency shows how many times a given word appears in the corpus; the criteria of the range shows in how many sub-corpora the word appeared or, in some studies, in how many texts of sub-corpora the word appears; the criterion of the coverage shows the percentage of the text that was represented

by these words. Additionally, researchers may apply their own criteria for words in a given list (e.g., dispersion (Gardner & Davies, 2013), field experts' opinion (Lei & Liu, 2016; J. Park, 2018)).

Moreover, for corpus size description terms 'token' and 'lemma' are often used. A number of tokens (or running words) shows how many words are in the corpus. If some word form occurs more than once, each occurrence is counted separately (Nation, 2013). For instance, in the sentence "This student does not get along with this group of students", each word will be counted as token (11 tokens) despite two occurrences of "this". A number of lemmas show how many word lemmas are in the corpus. Different forms of one lemma are counted as one (Nation, 2013). For instance, 'student' and 'students' will represent one lemma.

In this section, the most well-known and influential English and Korean lists will be reported. Moreover, different approaches and methods for word list compilation will be described.

2.3.1 *Word Lists for the English language*

The English language is the most studied in the world. Therefore, although this thesis is about word lists for the Korean language, I believe that it is important to take the experience of word list compilation for the English language into account.

As already mentioned in Section 1.2.1, the most well-known attempt to create a basic vocabulary list was made by West (1953). It contained 2,000 words that were identified on the basis of a corpus of 2.5 million words. With these words, learners are supposed to understand more than 80% of the vocabulary used in English texts and colloquial speech. Besides the fact that the list was used for the development of English language study materials, it has also become a basis or a reference for other lists creation. However, the use of GSL in research was criticized by some scholars (Browne, 2014; Engels, 1968; Richards, 1974) for its low relevance for modern English because the texts in the corpus for GSL were published before 1930.

Subsequently, since 2010 there have been two attempts to update the list. First of them was made by Browne, Culligan, and Phillips (2013), who used a corpus containing 273 million words based on nine Cambridge English Corpus (hereafter CEC) sections. As a result, they developed a list of approximately 2800 words. Sections of the CEC were carefully selected based on their size, in order to avoid the prevalence of vocabulary from the largest sub-sections. Their goal was to provide better coverage with fewer words, and the test conducted on the CEC showed that in comparison with West's list, the new one covers 90.34% of the corpus. It is also important to mention that for this list, researchers used a modified lexeme approach[11] that covers all the inflected forms from different parts of speech (Browne, 2014), instead of a common word-families approach[12], described by Bauer and Nation (1993). For the computational work, the authors used *Sketch Engine* (http://www.sketchengine.eu) and *AntConc* (Laurence, 2018).

Another new edition of GSL was also compiled in 2013 by Brezina and Gablasova (2015). However, their methods were different from those mentioned above. First of all, they relied on quantitative methods only. The study was drawn on data from four English language corpora (*The Lancaster-Oslo-Bergen Corpus*, *The British National Corpus*, *The BE06 Corpus of British English*, and *EnTenTen12*). Using Sketch Engine, the researchers examined the overlap between most frequent words in these corpora and compiled a list containing 2,494 lemmas based on the overlapping words with the most stable frequency and distribution indicators. Although the resulting list consists of lemmas and not word families as West's GSL (1953), its coverage is fully comparable to or even higher (in some corpora) than GSL's coverage. Thus, as can be seen from

---

[11] **Modified lexeme approach** "count the headword in all its various parts of speech and include all inflected forms"(Browne, 2014). For instance, *clear*, *cleared*, *clears*, *clearing*, *clearer*, *clearest* are in one group, while *unclear* and *clearly* presented as heads of separate groups.

[12] **Word families approach** unites a word stem and all its closely related forms (i.e. "all inflections and most frequent, productive, and regular prefixes and suffixes" (Bauer & Nation, 1993, p. 255)) into one word family. For instance, words *detect*, *detective*, *detecting*, *detectors* are part of one word family.

the results of the two attempts to upgrade the GSL, the modified-lemmas approach can be as efficient as the word families' approach. Moreover, a comparison of new editions and the GSL provided important information about the changes in core vocabulary that have occurred over time.

However, the development of such lists is not limited to general vocabulary lists. Some linguists (Campion & Elley, 1971; Coxhead, 2000; Lynn, 1973; Praninskas, 1972; Xue & Nation, 1994) also showed an interest in covering academic vocabulary as well. Such lists soon became an integral part of the vocabulary for the academic purposes' education process. Moreover, different methods for deriving such lists were developed, and different opinions about what words could be considered 'academic' were formed.

There have been several attempts to create an AV list for English language learners. The main difference between them is the corpus they used to create a list. For instance, one corpus made by Campion and Elley (1971) contained textbooks, published lectures, and some university examination papers. Another list created by Praninskas (1972) was made with the corpus consisting of ten textbooks for first-year students of American University in Beirut. Moreover, there was a more field-oriented list created by Ghadessy based on a corpus built of twenty chemistry, biology, and physics textbooks and list created by Lynn (1973) based on a corpus containing textbooks and handouts of students of accounting, economics and business administration fields. The next one was a so-called University Word List (UWL) created by Xue and Nation in 1984 from Campion, Elleys and Praninskas' lists mentioned above. It contained 808 words grouped into the 11 "levels" (or groups), where first-level words are the most frequent and last-level words are least frequent accordingly. A range was also considered in the process of level division.

Nevertheless, probably the most recognizable English AV list was made by New Zealand scholar Averil Coxhead and was called Academic Words List (hereafter AWL) (Coxhead, 2000).

Relying on corpus design principles stated by Sinclair (1991), she created a corpus that has approximately 3.5 million tokens (while all the previous corpora had no more than 500 thousand). The corpus covered 28 subject areas that were grouped into seven general areas of arts, commerce, science, or law (Coxhead, 2000).

As a result, she developed the AWL of 570-word families. According to Nation and Waring, word families are groups that include "a base word, its inflected forms, and a small number of reasonably regular derived forms" (Nation & Waring, 1997). To identify these 570-word families, Coxhead followed a three-stage process: 1) create a corpus of 3.5 million tokens; 2) count and sort word frequency with a corpus analysis program called *Range* (Heatley & Nation, 1996); 3) select words for the list based on specific criteria. Coxhead's criteria were as follows: word family should not be presented among the first 2000 most frequent English words (based on a West's GSL of English (West, 1953)); it also should be presented ten or more times in each of the four main areas and in at least 15 out of 28 subject areas; the overall frequency should not be less than 100 times per corpus (Coxhead, 2002). As a result, the AWL was created and is still widely used, despite of the fact that basic vocabulary that was excluded from the corpus to identify academic vocabulary was from GSL, that is already considered to be not the best representation of the current basic English words (Browne, 2014).

Among more field-specific academic vocabulary, lists there are lists for medicine (Lei & Liu, 2016; Wang et al., 2008), nursing (Yang, 2015), computer science (Lam, 2000), engineering (Mudraya, 2006), chemistry (Valipouri & Nassaji, 2013) and for other specific fields. These lists aimed to support the development of curriculum and teaching materials tailored to the specific purposes of students. Thus, as already mentioned in Section 2.2.2, these types of lists may be more efficient than a universal academic vocabulary (Hyland & Tse, 2007).

Corpora for these lists have many similar criteria like the IMRaD format of the articles included, fixed articles length, authors' first language (English in this case), years of publications, the reputation of academic journals, or the importance of the textbook for the students. However, corpora may be based on different academic materials (as it was aforementioned, corpora may contain textbooks, research articles, lecture texts, and notes). The majority of the field-specific academic vocabulary lists are based on the research articles' texts only.

Nevertheless, having very similar criteria for a corpus, the resulting lists, and the process of list construction may vary in four different ways:

- whether lists contain word families or lemmas;

- whether lists were used to filter the general and/or academic vocabulary form the corpus;

- which programs were used for the list creation;

- whether the lists were edited after the filtration processes by authors or specialists in the list's academic field or not.

Many lists contain word families like famous GSL and AWL (Coxhead, 2000; Lei & Liu, 2016; Martínez et al., 2009; Mudraya, 2006; West, 1953). Nevertheless, some scholars argue that the word-families approach may not be efficient for students since the concept of word family is usually acquired late, and morphological analysis skills range depending on a student's metalinguistic knowledge (Gardner & Davies, 2013). Therefore, some scholars considered individual lemmas as the best form of the final output (Brezina & Gablasova, 2015; Browne, 2014; Gardner & Davies, 2013; Lei & Liu, 2016).

To create the list computationally, different programs were used. Many studies (Coxhead, 2000; Liu & Han, 2015; Valipouri & Nassaji, 2013; Yang, 2015) used the *Range* software, which calculates a distribution of words across uploaded English or French language corpora or sub-

corpora automatically (Heatley & Nation, 1994). Additionally, some researchers used code written on Python (Lei & Liu, 2016), while some used the *Sketch Engine* corpus manager software (Brezina & Gablasova, 2015), and other used Wordsmith Tools (downloadable at www.lexically.net/wordsmith/) (Martínez et al., 2009).

The last point of difference between the above-mentioned vocabulary lists is whether they were edited by their authors (Browne, 2014; Lei & Liu, 2016) or by specialists in the field. The second method naturally has its advantages because professors and teachers use field-specific vocabulary often and can evaluate the importance of certain words for the students. Nevertheless, the authors who edited the lists without experts' assistance also provide a detailed explanation of a word's inclusion criteria.

To sum up, there were many studies that aimed to create word lists for the English language from elementary to very specialized lists. Additionally, several major differences may be identified, for instance: use of word family or lemma concepts; excluding of general and academic word families from a list or their use in lists comparisons; frequency and range as general selection criteria or consultation with the discipline experts to identify the correctness of the frequency-based list. With increased linguistic experience in the field of word list compilation, researchers apply more efficient methods that provide shorter lists with better coverage (like dispersion measurements).

Moreover, the overview of English word lists also shows that it is important to update word lists (especially general vocabulary lists) because the world surrounding us is constantly changing; thus, some words may fall out of use, while others may appear, or evolve in meaning becoming more frequent (Brezina & Gablasova, 2015). For instance, the words 'Internet', 'website', 'email' have become general vocabulary, while 'milkmaid', 'telegraph', 'servant' fall out of the daily use (Brezina & Gablasova, 2015).

2.3.2 Word Lists for the Korean Language

Since interest in the Korean language among international students has grown sharply in recent years (Looney & Lusin, 2018), the creation of Korean word lists, as a response to the need for KAP, has attracted much attention (Han, 2010; Kim, 2003; Lee, 2004). Consequently, many academic and technical vocabulary lists have been developed (e.g., J. Park, 2018; Shin, 2004; Yeo, 2018). However, the methods and designs used for their construction differ from English ones. These differences in the methodologies caused by the different morphological structures of the languages and different tools for their processing (e.g., the aforementioned program *Range*, for instance, was tailored for the English and French languages only), while the difference in the list design seems to be influenced by different trend-setters in both languages. In the following section, the most well-known general and academic vocabulary lists and their construction processes will be described. Moreover, the existing technical vocabulary word lists for engineering sub-fields will be examined. A summary of these lists is presented in Table 2.1.

There have been several attempts to create a list of general Korean vocabulary (Gang, 2012). However, the creators of these lists set themselves with different objectives. Some researchers tried to identify the entire vocabulary set necessary for a Korean language learner (N. Jo, 2003), while others attempted to create the shortest and the most efficient list (H. Jo, 2000). Moreover, some authors made their lists for certain language proficiency levels (Im, 2002) or with divisions of vocabulary in the list based on three proficiency levels (beginner, intermediate, advanced) (N. Jo, 2003). Therefore, these lists significantly differ in length.

The *Korean Learner's list* (hereafter KLL) is typically considered the most influential general vocabulary list. This list was constructed by the National Institute of Korean language and contained 5,965 lemmas of general vocabulary (Jo, 2003) and was based on the Korean Words Frequency List (Jo, 2002). Words included in the list were selected in two main steps. In the first

18

step, six specialists in the education field selected general vocabulary from the first 10, 000 most frequent Korean words; then, these specialists separately grouped the words by the level of difficulty. It was expected that the first-level group (or beginner level) should be around 1,000 words, second-level (or intermediate level) – around 2,000 words, third-level (or advanced level) – 3,000 words. All the words that specialists have not included in these groups were stored separately. As a result, some of the levels' content was very similar, while some of them differed substantially from each other. Therefore, discussions and regroupings (with the creation of additional "boarder" groups) continued until the final three groups were made (where A - is a beginner level words, B - is intermediate, and C - is advanced). The second step was to compare the list with Jo's (H. Y. Jo, 2000) list and with the contents of the Korean Learners' Dictionary. Words that were not in the list were inspected, and some of them were included in the list, even though some of them were not among the most frequent 10,000 Korean words. It is also important to mention that some decisions were made about certain word groups:

- initial forms of Korean-origin numerals were included in the list, but not Sino-Korean;

- some closely related words appeared to be in different level-groups because some of them were much more frequent than others;

- word forms with the *-cek* ending were not included since they are represented by their SK noun form;

- verb forms with *-sikhita* were also excluded, even if some of them were frequent.

Except for the words themselves, the list also contains important information about overall frequency (taken from the Korean Word Frequency List), part of speech, Chinese equivalent for Sino-Korean vocabulary, and word of origin for the loanwords from other languages (mostly from English). In the case of homonyms, the meaning of the word was marked with a number as in the

Standard Korean Language Dictionary [13] . The dictionary can be accessed online (https://stdict.korean.go.kr/main/main.do).

Perhaps, one of the first researchers who started to talk about the need for a full Korean AV list was Shin (2003). He reported that academic vocabulary is challenging to acquire without a special approach and that academic vocabulary lists are essential to increase the language competence of international students in Korea. Therefore, he compiled the Korean Academic Vocabulary List (hereafter KAVL) that contained 1,404 words grouped into the 957 modified-lexeme families (Shin, 2004). Moreover, the author compiled a list of 926-word family heads with the members of the families. For these purposes, a corpus of 3,634,809 tokens was built. It covered fields of humanities, social science, art, natural science, and engineering. The data were analyzed with the *SynKDP* corpus analysis software. This program can count the occurrences of a particular word form (not lemma) in the text and calculate the frequency of the form in the corpus. The lemmas range and frequency parameters were calculated as a separate step. The range parameter was set to the occurrence in each of the fields more than ten times, and the frequency parameter was set to not less than 100 occurrences across the whole corpus. The words that passed these criteria were added to the list. In addition, an extended version of the list with word families was also built.

Since then, not only the general vocabulary lists but also subject-oriented academic ones have been developed, such as the lists of vocabulary for Economics and Business (Kang, 2007) and for Natural Science (Choe, 2016). These studies have different approaches and corpus building principles.

---

[13] In the Standard Korean Language Dictionary, words that have the same written representation but several meanings (and different etymology) have a number assigned to each meaning.

For instance, for the Economics and Business field, Kang (2007) chose 16 related textbooks. Information about these books was retrieved from the course syllabi that were available online at Korean universities. Originally, Kang collected 76 study materials, but then excluded materials that did not meet the following selection criteria: 1) a book has a length more than 300 pages; 2) there are no more than two books from the same author; 3) the book is written by a Korean author (not a translation); 4) the books that were mentioned in different universities syllabi considered to be more relevant. The size of the final corpus was 777,402 tokens. All words in the list were put in the initial form:

- all prefixes and suffixes on nouns were removed;

- verbs with -*sikhita* were changed to the standard -*hata* form;

- adjectives built with -*cek* from SK nouns were returned to their SK noun form;

- personal names (of researchers, books, etc.) were excluded;

- words with several meanings (homonyms) were checked based on their context.

Then to create the final word list, Kang filtered the corpus through the list of basic Korean words and the list of basic academic words. These manipulations were made with Microsoft Excel functions. As a result, Kang (2007) developed a list of 771 Business and Economic academic vocabulary.

In 2016, Choe made vocabulary lists for three fields of natural sciences: chemistry, biology, and physics. The process of the creation of these lists differed from that of Kang's (2007). Firstly, her corpus was based on testing materials (2011-2015) from the College Scholastic Aptitude Test (CSAT) in Korea, from which she compared words to the science glossary of the Korean Ministry of Education for university admission. Then, words present in both sources were compiled into one list and then divided into separate three field-specific lists. These two sources were selected

because the CSAT is a required exam in Korea for university applicants, and the science glossary was designed for publishers of Korean school textbooks. Therefore, Choe assumes that these words are essential for future students of natural sciences programs since both the Ministry of Education and the CSAT committee consider them necessary. For these purposes, Choe used the Adobe Reader, to convert test papers to a digital format, and the JAVA programming language, to compare corpora, and Microsoft Excel, to calculate word frequency and distribution. The final list sizes were 356 (chemistry), 491 (biology), 313 (physics) words.

Nevertheless, more importantly, for the current study, the list of Physics Vocabulary (hereafter PVL) for the first-year international engineering students was compiled in 2018 (J. Park, 2018). The list contains 624 nouns and noun phrases. It aims to help international students in engineering colleges to be prepared for the first year physics classes since this subject is obligatory for engineering students. To create the list, Park built a corpus of Korean physics textbooks. The chosen textbooks were not originally written in Korean, but are the editions translated from English. In order to compile the final version of the list, the author consulted with physics professors. Some words from Intermediate and Advanced level KLL are presented in the list as well. The difficulty level of the words and the frequency of their occurrence are also marked in the list.

There are also different studies of technical vocabulary. For instance, Yeo (2018) made a list for Politics and Diplomacy majoring students. The process of list creation is close to Kang's process. First, the author chose universities with the highest enrollment of international students. Then she found five books that are in use for mandatory courses for Political Diplomacy. During the book selection process, the author also paid attention to the relevance of these books in the library and four on-line bookstores. As a result, corpora with 530,000 tokens were made. The author used ABBYY Fine Reader to scan books for the corpus, 'Intelligent morphological analyzer 2.0' (Kor.: *Jineunghyeong hyeongtaeso bunseogi 2.0*) to tag POS, 'UTagger 2.1.2' to work with

the tagged corpus, and Microsoft Excel to exclude all parts of speech except nouns and compile the final list. All basic and semi-technical Korean words were deleted from the corpora prior manipulations. As a result, a list of 582 words was compiled.

However, more importantly, for the current study, there are technical vocabulary lists that were made for engineering sub-fields. For example, there are lists for automobile engineering (Kim, 2015), electrical and electronics engineering (Park, 2017), mechanical engineering (Lee, 2014), and textile engineering (Nergui, 2015). These lists have some similarities; for instance, all of them consider frequency and the opinion of the experts in the field as inclusion criteria while only Lee (2014) had frequency and range; the authors mostly used textbooks to build corpora (only Nergui (2015) added research articles as well), and the lists were divided into several levels by importance for the students (or frequency of occurrence in study corpus). Furthermore, all of them compared their final list with the technical dictionary from the researched field.

However, the output lists are organized differently in these studies. Park (2017) presents results in the form of a list of nouns. Lee (2014) also provided a list of nouns, but, in addition, suggested teaching methods for these words. Nergui (2015) developed not just the list of the nouns but also the table with verb collocations, while Kim (2015) provided verb collocations in addition to a list of the compound words that may be built from words in the list. Hence, in the case of the last two lists, they perform two functions: 1) report on the most necessary technical words in the field and 2) show examples of their usage.

Summarizing the above, we can observe several different approaches to Korean vocabulary lists creation. The main difference is the design of the final list (simple lists, lists divided into three groups by the difficulty level, lists of nouns with collocations). However, there are also some differences in the data processing. For instance, SK nouns compounded with *-hata* in some cases were just lemmatized (Kang, 2007), while others were presented in the list with just noun form (H.

Lee, 2014). Furthermore, researchers tend to use different methods of data processing. The diversity in the tools used for POS-tagging and lemmatization is also apparent: Intelligent morphological analyzer 2.0, *Geuljabi II, Hanmaru* 2.0 (these programs may be downloaded after the registration and identity authentication from the website of National Institute of Korean Language https://ithub.korean.go.kr/user/corpus/programManager.do).

In comparison to the English word list compilation processes, Korean word lists require more steps. For example, for word lemmatization and token counting, the POS-tagging step is first required. As for the presentation of lists, Korean lists include more details because the information about etymology and polysemy is sometimes essential for learners to understand a Korean word. Moreover, studies on the Korean technical vocabulary tend to provide collocation lists for the selected words, or if the list is noun-based, to provide their verb collocations and required case marker. Some studies (e.g., J. Park, 2018; Kim, 2015) also developed recommendations for teaching these words. A summary of the described word lists is presented in Table 2.1.

Table 2.1

Summary of the described Korean word lists

Title or Field and Length	Author	Type[14]	Corpus Size	Sources[15]	Content	Word Inclusion Criteria	Tools Used
Korean Learner's List (5,965)	Jo (2003)	G	-	WFL	Lemmas	Frequency, experts' opinion	*Microsoft Office Excel*
Korean Academic Vocabulary (916-lemmas; 1,404)	Shin (2004)	A	3,634,809	TB	Modified-lemmas, word families	Frequency, range	*SynKDP*
Natural Science (Physics -313, Chemistry - 356, Biology - 491)	Choe (2016)	FSA	-	Science glossary, College Scholastic Aptitude Tests	Nouns, noun phrases	Frequency, range, dispersion	*JAVA programming language, Microsoft Office Excel*

---

[14] In this column the following abbreviations were used: G – general vocabulary, A – academic vocabulary, FSA – field-specific academic vocabulary, T – technical vocabulary.

[15] In this column the following abbreviations were used: TB – textbooks, RA – research articles, WFL – words frequency list (Jo, 2002) that was developed based on a 21st Century Sejong Corpus of Korean Language.

Title or Field and Length	Author	Type[14]	Corpus Size	Sources[15]	Content	Word Inclusion Criteria	Tools Used
Economics and Business (771)	Kang (2007)	FSA	777,402	TB	Lemmas	Frequency	Intelligent morphological analyzer 2.0, *Microsoft Office Excel*
Physics (624)	Park (2018)	FSA	>1,500,000	TB	Nouns, noun phrases	Frequency, experts' opinion	*Hanmaru* 2.0, AstroGrep
Political Diplomacy (582)	Yeo (2018)	T	528,393	TB	Nouns	Frequency, range	Intelligent morphological analyzer 2.0, UTagger 2.1.2, *Microsoft Office Excel*
Electrical and Electronics Engineering (89)	Park (2017)	T	485,469	TB	Nouns	Frequency, experts' opinion	ABBYY Fine Reader, Intelligent morphological analyzer 2.0, *Geuljabi II, Microsoft Office Excel*
Mechanical Engineering (372)	Lee (2014)	T	934,103	TB	Nouns	Frequency, range	Intelligent morphological analyzer 2.0, UltraEditv19, *Microsoft Office Excel*
Automobile Engineering (716)	Kim (2015)	T	266,652	TB	Nouns	Frequency, experts' opinion	*ABBYY Fine Reader, Ttolmangsae* 2.5, *Microsoft Office Excel*
Textile Engineering (203)	Nergui (2015)	T	857,895	TB, RA	Nouns, verb collocations for nouns.	Frequency, experts' opinion	*Geuljabi II, Microsoft Office Excel*

2.4 Word lists implications

Although word lists may be used directly for language teaching and learning as a reference for the most frequent words (for certain fields/purposes), their primary role is to be a base for the development of other language teaching materials (Coxhead, 2000). First of all, word lists are essential for textbook development. They set the primary vocabulary focus for authors so that they can design textbook chapters and all types of activities (listening, reading, writing, speaking, etc.) in accordance with the most necessary vocabulary.

Furthermore, word lists may be used for the development of language testing materials, since they provide a cross-section of the most important vocabulary for a specific purpose/field and give information on word frequency. Thus, testing material developers can balance the vocabulary in a test accordingly.

Word lists may also be used by language teachers since they highlight vocabulary that requires more attention in the classroom. Especially, classroom activities that are designed to use particular words from lists may be extremely helpful for the learners (Nation, 2001).

Additionally, it should not be forgotten that word lists may be used by language learners directly. First, they can use them as a vocabulary "check-list". Second, not all students may have access to field-specific textbooks or language courses. In this case, word lists (available online) may become a basis for the self-education and self-preparation for university or other purposes.

Academic and technical word lists can also make readings in very narrow fields more reachable for language learners since they provide the words that appear the most in the exact field.

To summarize, although word lists have various implications, their main goal is to provide vocabulary necessary for the construction of other teaching materials.

2.5   Rationale for the study

The literature review has shown that word lists are now in great demand for the Korean language, and methods for their development are being actively discussed and improved using modern language processing technologies and special mathematical and statistical formulas. Nevertheless, in the case of Korean word lists, there is still a larger field for further development. With a growing number of international students, more and more academic fields may require special language preparation for the students. Thus, these fields will require word lists to base the

academic materials on. At the same time, the word list creation methods themselves may also require further development due to the high demand.

It was also observed that there are different opinions about how word lists should be designed, and what output form is more efficient for the students (word families vs. lemmas). Moreover, materials that were used for the corpora varied from one study to another.

As was mentioned in Chapter 1, engineering is currently a field in high demand among international students in Korea (Korean National Institute for International Education, 2018). Hence, word lists for this field may contribute language for academic purposes education and international students in engineering colleges directly. A survey of the literature review shows that there are several lists of the technical vocabulary for some engineering sub-fields. Nevertheless, there was no study done on the academic vocabulary common for the main engineering sub-fields as a whole. Therefore, the research questions I address in this thesis are:

1. Which lexical items frequently occur across engineering sub-fields (but are not included in the Korean Learner's List), and what are their features?

2. Which words frequently co-occur with these lexical items, and what are the overall tendencies of their collocations? In other words, what are their frequent collocations if they have such.

3. Are there high-frequency words among engineering academic texts that were not identified in the KAVL?

According to the research questions, the goal of this thesis is to develop a Korean Engineering Academic Vocabulary List (hereafter KEAVL).

The objectives of the study are:

- to build a corpus of engineering vocabulary in Korean based on textbooks and research articles published in the main sub-fields of engineering;

- to automate the way of corpus data analysis and word list compilation;

- to organize KEAVL results with collocations in a table;

- to analyze the content of the resulting lists;

- to make KEAVL available online.

3 Methodology

This chapter explains the methods used in completing the study. It gives special emphasis to the process of data collection, corpus building, and data processing. In the process of constructing the methodology for the current study, much attention was paid to the design of previous studies mentioned in Chapter 2.

3.1 Data collection and corpus establishment

For the purposes of the current study, a corpus of textbooks and research articles in the Korean language published in the field of engineering was created. Unlike previous studies that contained only research articles (Lei & Liu, 2016; Yang, 2015) or only textbooks (Kim, 2015; J. Park, 2018; S. H. Park, 2017) in a corpus, in this study I decided to include both types because such a combination provides a complete representation of the field, and thus provides a larger variety of writing styles and vocabulary use (Sutarsyah, Nation, & Kennedy, 1994).

Naturally, engineering is a vast field and has specific sub-fields. In order to identify the main sub-fields, I first reviewed articles in various engineering sub-fields on the Korean research articles database *DBpia* (http://www.dbpia.co.kr/). This database has more than 2.5 million research articles written in Korean or English languages from over 2300 Korean journals and considered to be one of the largest and representative collections of Korean research articles in different disciplines from humanities to engineering. There were 12 sub-fields of engineering at the time of the study: general engineering, mechanical engineering, aerospace engineering, chemical/biological engineering, electrical/control and instrumentation engineering, civil/environmental engineering, electronics/communications engineering, architectural engineering, industrial engineering, marine/ocean engineering, resources/materials science and engineering.

Since the current research goal is to provide a word list for international students of engineering departments, I also confirmed if the sub-fields above correlate with the engineering departments in the following top Korean engineering universities: Korea Advanced Institute of Science and Technology (KAIST), Pohang University of Science and Technology (POSTECH), Ulsan National Institute of Science and Technology (UNIST), to be sure that all fields are covered. As a result, only the sub-field of nuclear and quantum engineering was not mentioned on the *DBpia* as separate. Nevertheless, it was presented in the sub-field of the resources/materials science and engineering. Therefore, the *DBpia* library sub-fields match the engineering departments of the top Korean universities.

To identify the most important journals for each engineering sub-field, I relied on the *DBpia* top-journals rating and the Korean Citation Index (hereafter KCI) rating (https://www.kci.go.kr/kciportal/landing/index.kci). For each sub-field, I collected two journals for the corpus (a total of twenty-four). The majority of the journals chosen are in the *DBpia* TOP-5 journals for their sub-field and are published by specific field engineering associations or research institutes. Moreover, all journals are licensed by KCI. All selected journals were open-source and made available on the Korean Open Access Journals website (https://www.kci.go.kr/kciportal/landing/index.kci) or through the Korean National Digital Science Library (http://www.ndsl.kr).

Then from each journal, fifteen research articles were randomly selected after meeting the following criteria:

- length criterion: since the length of the articles in the identified journals ranges from 2 to 15 pages (average length is 5-7 pages), it was decided to set a length criterion of a 6-page minimum;

- <u>author L1 criterion</u>: since some articles are written in co-authorship with international researchers, it was decided to limit research articles to the articles that written by Korean authors only for the sake of corpus consistency (the judgment was based on the information about authors that is usually presented in the beginning or in the end of the Korean research articles);

- <u>author presence criterion</u>: to avoid the influence of personal writing style, only one article per same author/group of authors was included to the corpus;

- <u>year of publication criterion</u>: to limit the size of the possible articles, I set the publication year limit between 2000-2020;

- <u>structure criterion</u>: only articles with identifiable Introduction – Methodology – Results – Conclusion (typical structure for Korean engineering articles) structure were selected.

As a result, 360 research articles were collected for the corpus with a total of 829,562 tokens (or running words). The average article length in the corpus is 2,330 tokens, the shortest article contains 886 tokens, and the longest article contains 5,441 tokens. Detailed information about each journal selected and article used for the current research may be found in Appendix A and B accordingly.

The process of textbook selection was a bit more complex. Previous studies mostly considered textbooks that were mentioned in universities' syllabi for their corpora. However, official syllabi for engineering departments' courses of top Korean universities mostly mention English language textbooks. Therefore, only Korean textbooks that were mentioned in the engineering syllabi as recommended literature were considered. Moreover, the rating of books in a particular field on the biggest Korean bookstore website, *Kyobo* (www.kyobobook.co.kr), was examined. As mentioned in Section 2.3.2, Yeo (2018) also considered the *Kyobo* ratings in the

book selection process. After a comparison of books in recommended reading lists and ratings of *Kyobo* for each engineering sub-field, a shortlist of three books for each field was compiled. Then, the detailed information about each textbook, their introductions, the author's affiliation, publisher's reputation, and e-book format availability were checked. As a result, of three books, only one was chosen to be included in the corpus. During the process of final selection, attention was also paid to the level of the textbook, e.g., introductory, advanced, etc. The preference was given to the introductory textbooks; for example: "Introduction to mechanical engineering" rather than more advanced and narrowly focused textbooks. Moreover, the author's L1(not translated textbooks) and year of publication (2000-2020) criteria were applied. All books were purchased in the e-book format from the already mentioned bookstore *Kyobo* and online bookstore *Yes24* (www.yes24.com). As a result, 11 books were collected (see Appendix C). Textbooks for general engineering were not chosen due to the fact that there are only textbooks in the specific engineering sub-fields. Unfortunately, I was not able to collect more books for each sub-field because of the library logistical difficulties. Nevertheless, the number of books collected seems to be sufficient to meet the goals of the current study.

When the process of data collection was finished, all the research articles that were in PDF format were converted to .txt files with ABBYY FineReader 14 (ABBYY Production LLC, 2017). Then, all references, appendices, footnotes, acknowledgments, information about authors, and other information that was not included in the articles' main structural parts were manually removed. Unfortunately, despite of attempts to automate this process, the design of the articles varied greatly; thus, it was impossible to write the script that can remove unnecessary parts from all of the articles.

The process of preparing the textbooks was similar to the research articles' preparation, except the e-books were first converted to PDF format from the mentioned bookstores e-book

formats and only then were converted to .txt files. Each chapter of the textbooks was stored as a separate .txt file.

Furthermore, each corpus file was checked manually to be sure that ABBYY FineReader (ABBYY Production LLC, 2017) optical recognition gave high-quality results. The completed corpus consists of 360 articles and 11 books from 12 engineering sub-fields. The corpus was titled Corpus of Korean Engineering Academic Texts (hereafter CKEAT).

The size of the corpus, regarding both tokens and lemmas, was calculated with the program described in Chapter 3.2. The textbooks and research articles collected for the corpus contained many English and Chinese words, numbers, and mathematical/ physical/ chemical formulas that are usually counted as words by word counting tools. Hence, in order to provide definite information about corpus size, non-Korean words and other symbols and the Korean particles attached to them were excluded from the corpus (in a separate file) prior to the tokens' counting like in Shin (2004). It was done with the *Regular Expressions* formulas in *EditPad Lite 7* text editor. The formulas searched for all of the non-Korean and non-punctuational *Unicode* symbols. For instance, to find all of them in the .txt file the following formula was used: [^\uAC00-\uD7AF\u0020-\u0023\u0027-\u002A\u002C-\u002F]. The matches were removed with the *EditPad Lite 7* built-in tools. The particles attached to them were also removed. For instance, there were cases when the word written in Korean was following by its English translation (e.g., 텔레매틱스 (telematics)). In these cases, the English translation was removed to not affect the number of words written in Korean in the corpus. All manipulations to count tokens and lemmas were done in a separate file, and the original corpus was not changed.

In order to count the tokens, the *Mecab* (Kudo, 2013) POS-tagger was used. The reason for this choice is that in some cases, authors may write some noun phrases without spaces while

some authors may write them with spaces (etc. *hwankyengoyempangci* VS *hwankyeng oyem pangci (環境汚染防止)* 'prevention (of) environmental pollution'). Hence, ordinary word counting tools (like in *Microsoft Office Word*) cannot give accurate information about a number of tokens used in the corpus. On the contrary, Korean language pos-taggers like *Mecab* and *Kkma* divide such compounded words while pos-tagging, hence they can provide an accurate number of the words used in the corpus.

As a result, the compiled corpus has 1,303,272 tokens and 29,896 lemmas. Following the modified lexeme approach (Gardner & Davies, 2013), Sino-Korean adjectival and verbal nouns forms with *-hata*, *-toyta*, *- sikhita*, and *-cok* were counted together with the noun forms. Other word forms that have the same root were counted as separate words. For instance, *silhem (實驗)* 'experiment', *silhem-hata* 'conduct an experiment', *silhem-cek* 'experimental' are counted as one unit, while *silhemchey* 'subject of the experiment' is counted as a separate word.

The function that counts tokens and lemmas was also added to the program that was written for the current research (see Section 3.2). The corpus size with regards to token and type counts is organized by sub-field in Table 3.1, and detailed information about the minimal, maximal, and an average number of tokens in research articles and textbooks is presented in Table 3.2.

Table 3.1

The Corpus of Korean Engineering Academic Texts

Sub-field	Total Tokens	Total Lemmas
General	63,382	4,902
Mechanical	75,990	6,698
Aerospace	121,763	5,289
Computer	128,802	7,943
Chemical/Biological	125,249	8,310
Electrical/Control	73,823	5,330

Civil/Environmental	114,997	8,845
Architectural	148,519	11,142
Industrial	124,231	6,568
Marine/Ocean	97,967	5,936
Resources/ Materials Science	106,152	6,436
Electronics/ Communications	108,647	6,626
Total	**1,303,272**	**29,896**

Table 3.2

Comparison of the sub-corpora units' size (in tokens)

Sub-field	Textbook Tokens		Research Article Tokens	
	Min/Max	Average	Min/Max	Average
General	-	-	1,055/3,086	2,154
Mechanical	1,178/9,587	4,841	914/3,919	1,768
Aerospace	1,427/7,285	4,651	1,113/3,749	2,043
Computer	1,257/8,727	4,322	1,901/4,424	2,606
Chemical/Biological	2,471/6,184	5,102	1,038/3,920	2,176
Electrical/Control	3,336/6,075	4,376	1,054/4,476	2,350
Civil/Environmental	2,817/5,141	3,874	1,214/4,901	2,454
Architectural	1,084/7,781	3,624	1,718/4,727	2,864
Industrial	894/3,602	2,428	1,750/5,411	3,130
Marine/Ocean	2,382/6,364	3,839	900/4,610	2,284
Resources/ Materials Science	1,506/6,390	4,132	1,087/4,282	2,203
Electronics/ Communications	1,123/3,433	1,982	886/3,213	1,912

3.2   Data processing

Data processing was entirely performed with a program written with the *Python* programming language, which I call *KWordList*. *Python* was chosen for the purposes of the current thesis because of its high efficiency for natural language processing (NLP). The main packages

that were used in the process of program creation are *NLTK* (Natural Language Toolkit)(Bird &

Loper, 2009) and *KoNLPy* (Korean Natural Language Processing in Python) (Park & Cho, 2014).

The program creates a word list in seven main stages:

1. lemmatization;

2. general vocabulary extraction;

3. academic vocabulary extraction (optional stage);

4. range calculation;

5. frequency calculation;

6. dispersion measurements;

7. collocations search.

In addition, the program counts tokens and lemmas in the text, percentages of tokens and lemmas

covered by general and academic vocabulary lists, and according to the compiled list.

The lemmatization process is the most important for the program because the result directly

depends on the quality of the lemmatization. For the Korean language, lemmatization may only be

performed with the part-of-speech taggers (Han & Palmer, 2005), as it was mentioned in Chapter

2.1. Currently, the *KoNLPy* package has five taggers: *Hannanum* (KAIST Semantic Web Research

Center, n.d.), *Kkma* (Lee, Yeon, Hwang, & Lee, n.d.), *Okt* (previously *Twitter*) (Ryu, n.d.),

Komoran (Shineware, n.d.)*,* and *Mecab* (Kudo, 2013) (was used on the *Ubuntu* machine). Each of

these taggers has specific pros and cons. For instance, the *Okt* and the *Mecab* have the best loading

time and execution time (the speed of POS-tagging). In contrast, the *Kkma* shows better in

performance in difficult cases, e.g., it handles POS-tagging of the sentences without spaces and

compound words better than the other three. Nevertheless, it shows an increase in the processing

time while the large corpora are loaded.

For the first versions of the code, the *Kkma* tagger was used. Nevertheless, although it worked with a testing part of the corpus when the full version of the corpus was uploaded, *Kkma* processing time rose significantly (even using it on powerful computers). Therefore, despite *Kkma*'s excellent accuracy of tagging, the preference in this study was given to the more speed/accuracy balanced tagger – *Mecab*. It has high accuracy (under normal conditions). At the same time, it is written in the *C++* programming language that allows it to work much faster than Kkma that is written on *JAVA*. However, it is important to mention that the accuracy of tagging depends on the corpus compiled. For instance, if a text has no spaces, the accuracy of POS tagging may decrease (L. Park, 2018). Hence, the choice of a tagger is highly dependent on the research objectives and on the format of the corpus used in the study. Since the corpus for the current study has spaces and was compiled very carefully, *Mecab* tagger was chosen.

To extract all nouns from the corpus, a special *KoNLPy Mecab* function (mecab.nouns()) was used. However, for predicate and adverb extraction, a separate code based on the tags was written. In the case of predicates, the code identifies the words tagged as 'VA', 'VV', 'VXV', and then lemmatize them by adding infinitive form ending *-ta* to the stem. In the case of adverbs, the code is looking for the words tagged as 'MAG'. All the tags that *Mecab* uses are presented in Appendix D. Since adverbs in Korean do not attach morphemes, they also do not require lemmatization. The results of the lemmatization of each sub-corpus file were stored separately by the program.

During the second stage, the general vocabulary was extracted from the lemmatized words list. For this purpose, words from the KLL mentioned earlier were added to the code. The KLL was downloaded from the Korean Language Institute Website (https://www.korean.go.kr/). As mentioned in Section 2.3.2, the list contains not just words but also important information about them. For the program, only the column with the words was used. However, instead of one list, I

chose to divide them into three lists by the level of difficulty stated in another column (A, B, C). Since word lists may be focused on the students with different levels of proficiency in Korean, it was concluded that it would be better to use three separate lists (i.e. it may help to build lists, for instance, for intermediate or advanced learners separately).

It is essential to mention that all three lists were pre-processed before they were added to the code. More precisely, the manipulations with the predicates were done. This was important because of the taggers' features that were mentioned in Chapter 2.1. The SK adjectival and verbal nouns compounded with Korean predicates *hata*, *toyta*, and *ita*, when being processed through the tagger (both *Kkma* and *Okt*), are divided by nouns and predicates accordingly. Thus, it was essential to remove SK predicative forms from the *Korean Learners'* lists to avoid the situation when the program cannot find a match in the list (because one word is represented by different POS). Nevertheless, one-syllable SK adjectival and verbal nouns in the form of predicate were not manipulated because they will not be divided by the taggers. All the manipulations mentioned above were done with the regular expressions (or *Regex*) formulas in the *EditPad Lite 7* text editor.

Moreover, originally the KLL contained the number of the meaning (for homonyms) as it is presented in the Standard Korean Language Dictionary. These numbers were removed from the lists. Therefore, the program itself does not take into account multiple meanings of the words. Nevertheless, this problem is solved because the words are presented with their collocations, thus, the meaning of the word (for homonyms) may be identified from the context.

To summarize, the program uses three lists that represent vocabulary for the different levels of proficiency in Korean. Depending on the preferences set at the beginning, the program compares the list of lemmatized words of the uploaded corpora with the vocabulary lists and removes all the words that occur in both. For instance, the word *mul* ("water") will be removed by the program, because this word is in the A-level vocabulary list. As a result, the program compiles a new list

that has only those words that are not presented in the general vocabulary lists (in all three of them, or only in some of them, depending on the preferences. For the current study, word lists for different levels of proficiency learners were created. Thus, for the first lists, A- and B-level words were excluded, while for the second list (for the more proficient students), all the general vocabulary was excluded.

The next step is the exemption of academic vocabulary. This step is optional and required only for those cases when the aim is to compile a technical vocabulary list. Hence, this study did not require this step. Notwithstanding, this option was added to the program to make it more useful in the future for different kinds of research. The process is similar to the previous one: the program compares words that left after the exclusion of the general vocabulary with the academic vocabulary word list and removes the words that present in both. As a list of academic vocabulary, the list made by Shin (2004) was used.

The fourth stage is the range measurements. For the case of the current study, it is important for words to appear across different engineering subfields and not just in one. Like in some previous studies (Wang et al., 2008), the range criterion was set to 50%. Hence, only a word that appeared in ≥ 50% of sub-fields can pass the range criteria.

The fifth step is the frequency measurements. For the words that passed the range criterion, the program measures their frequency. Then, it arranges them by the frequency from the most frequent to the less frequent. The words with the frequency above 100 were included in the lists.

The sixth step was the variability value (or statistical dispersion) calculation. There are several ways to count specifically the dispersion of a lexical item. For this study, the deviation of proportions (hereafter DP) measurement introduced by Gries (2008) was used. The reason for this choice was that the DP measurement was geared to handle corpora with different sizes of sub-corpora. Hence, the DP value is very sensitive to the disbalance of the word occurrence in

39

correlation to sub-corpus sizes that makes it perfect for the unequal CKEAT sub-corpora size (see Table 3.1). In other words, this value can show whether the word is evenly spread among sub-corpora, or it is mostly concentrated on a specific sub-corpus and rarely occurs in others. For instance, a word may occur in 8 out of 12 sub-corpora passing the range criterion, and occur more than 500 times across the corpus, passing the frequency criterion. However, if this word occurred 400 times in one sub-corpus and 100 times in the others, it makes its value questionable since it is more common for one specific sub-field than for a broader field of engineering as a whole. The DP calculation reflects the disbalance in the frequency of occurrences and in the size of the sub-corpora of word occurrence and, thus, can filter these words out of the list.

To calculate DP, a special function was written. The formula for the DP calculation is given in Formula 1 (from Brezina, 2018):

$$\text{DP} = \frac{Sum\ of\ absolute\ values\ of\ (observed-expected\ proportions)}{2} \tag{1}$$

First, following Gries' (2008) steps, this function counts the size of each sub-corpus, normalizes these numbers against the entire corpus. Second, the code counts the occurrences of the word in each sub-corpus and normalizes these numbers against the total frequency of the word. Third, for each sub-corpus, the difference between the normalized occurrence and normalized size is counted. The resulting numbers are then summed and divided by two. This number is the DP of the word.

The closer DP to the 0, the more evenly word is spread across the corpus, while a DP closer to 1 shows that a word occurrence is imbalanced. For this study, the maximum DP value for a word to be included in the list was set to 0.5.

After completing these five stages, two lists containing engineering academic vocabulary were compiled: one for the intermediate level learners of Korean and one for advanced level learners of Korean.

The last step is a search of collocations for each word in the lists. For these purposes, the *NLTK* package for *Python* was used. This package has a special tool that can search for bi-grams and tri-grams. The number of collocations and their frequency may be set individually. These tools can provide the most frequent collocations for the whole corpus, and for particular words as well. The program was set to search for the collocations for each word in the final lists; the frequency of collocation occurrence was set to ≥ 30; the maximum number of collocations was set to the twenty most frequent collocations. The program looked for both bi-grams and tri-grams. The collocations were derived from the tagged version of the full corpus that was prepared separately for this purpose. The tagged version contains not just lemmas but also suffixes, endings, etc. Therefore, all the cases and the most frequent verbal morphemes were set with NLTK as corpus parts to be ignored.

All the word data received (word, range, frequency, DP value, and collocations information) is stored by the program to the *Python* nested dictionary[16] forming the database. The dictionary format was used because it allows storing data in 'pairs'. In the case of the program presented in this study, the pairs contain a word as a key and all the information about the word as its values. Hence, it makes it easier to read and change information. Moreover, it allows converting all the data to the conveniently organized table (examples of output are presented in Chapter 4). Then, this dictionary is converted to a CSV-file (Comma Separated Values), which is the final output of the program. This type of file may be opened easily in Microsoft Word Excel and edited.

In addition, the program also creates a separate CSV-file with the number and percentage of tokens and lemmas covered by each list. Moreover, it also saves all words excluded by the KLL (by levels) and by KAVL if it is required in separate .txt files.

---

[16] Dictionary is a format of data storage in *Python* where multiple values may be assigned to the unique key (word, in case of this study). Nested dictionary may contain dictionaries inside dictionaries.

To summarize, in order to compile KEAVL, the corpus of Korean academic texts in 12 main engineering fields was built. Afterward, all the data was processed with the *Python* code, which selected the words and their collocations that were meeting the criteria of the study. The process is summarized in Figure 3.1. The script is fully available online at *GitHub*, a special platform to share the code (https://github.com/HelgaKr/KWordList).

Corpus input	Calculation of word frequency	Calculation of DP value (for words that passed range and frequency criteria)
Deletion of all non-Korean and non-punctuation symbols with Regex	Calculation of word range	Collocation search for the words in the list
POS-tagging and lemmatization of the corpus (with Mecab)	Deletion of the general and (if required) academic vocabulary	Output and coverage data calculation.

Figure 3.1 Process of word list creation with KWordList

4   Results

As stated in Chapters 1 and 2, the study reported here examined in detail the academic engineering vocabulary of Korean and used computational methods to develop word lists for it. This chapter is organized in accordance with the three specific research questions posed in Section 2.5. It first presents the word lists that were compiled based on the built Korean Corpus of Engineering Academic Texts and their coverage; it then presents the collocations found for the lemmas in the list; finally, it compares the built lists to the Academic Vocabulary List of Korean and explains its role among other word lists for engineering students.

4.1   Korean Engineering Academic Vocabulary List

This study focused on the most-frequently-used vocabulary in Korean academic texts (textbooks and research articles) in the field of engineering. The first research question asked which lexical items frequently occur across different sub-fields of engineering but are not presented in the KLL (in A and B levels for intermediate level learners and A, B, C levels for advanced level learners).

In order to answer this question, a corpus of 1,303,272 tokens and 29,896 lemmas was compiled. Then, the corpus was processed with the program described in detail in Chapter 3.2. Since the goal of the study is to compile two lists of academic engineering vocabulary (for intermediate and advanced level learners), the results for two lists will be described below.

For inclusion into the KEAVL (with both intermediate and advanced words), a word had to occur at least in six engineering sub-fields, that is, half of the twelve selected sub-fields, and at least 100 times across the corpus (0.01% of corpus or 77 times per million words). More importantly, word had to have a DP value less than 0.5 (i.e., a word had to be evenly presented across sub-corpora). Furthermore, general beginner- and intermediate-level vocabulary (from the

KLL) was excluded prior to the compilation of the list. Words that were filtered out from the corpus by the beginner- and intermediate- level basic vocabulary lists are presented in Appendix E and F accordingly.

As a result, 1,189 words passed the range and frequency criteria, and 830 words passed all three, including dispersion. The program output provided the .csv file with all 1,189 words and their range, frequency, dispersion value, and frequent collocations (that appeared more than 30 times in the corpus). The words that have not passed the dispersion criteria were provided in the output file to give a researcher the possibility to analyze the words these words. Since the output file may be easily edited with Microsoft Office Excel, the words that have not met dispersion criteria may be deleted from the list after sorting the word list by dispersion value. The example of the output is shown in Figure 4.1.

	Column1	range	frequency	collocations	dispersion
2	시스템	12	4096	결재 시스템, 모니터링 시스템, 탐지 시스템, 제어 시스템, 데이터베이스 시스템, 파일 시스템,	0.255
3	증가	12	3293	급격히 증가	0.19
4	데이터	12	3161	데이터 마이닝, 데이터 버전, 데이터 링크, 계측 데이터, 데이터 전송, 데이터 통신, 입력 데이터	0.41
5	구조	12	3143	외피 구조, 바닥 구조, 구조 부재, 콘크리트 구조, 네트워크 구조, 구조 설계, 구조 해석, 구조 시	0.19
6	설계	12	2744	내진 설계, 참여 설계, 설계 기준, 설계 단계, 건축 설계, 설계 변수, 기본 설계, 구조 설계, 설계 고	0.24
7	특성	12	2677	절연 특성, 특성 곡선, 거동 특성, 유동 특성, 제어 특성	0.155
8	수행	12	2599	수행 능력	0.185
9	구성	12	2509	구성 요소, 구성 성분, 시스템 구성	0.195
10	방식	12	2460	변조 방식, 전송 방식, 통신 방식, 신호 방식, 생산 방식	0.385
11	영향	12	2441	정의 영향, 의한 영향	0.195
12	평가	12	2354	성능 평가, 평가 항목, 평가 단위, 평가 기준, 평가 대상, 평가 방법, 평가 결과, 평가 시스템, 평가	0.28
13	실험	12	2331	실험 체의, 실험 결과, 실험 장치, 실험 데이터, 실험 방법	0.285
14	내	12	2295		0.16
15	측정	12	2290	압력 측정, 측정 장치, 측정 방법, 측정 결과	0.28

Figure 4.1 Program output example (for the full KEAVL)

For the case of advanced part of KEAVL, the words had to pass the same range, frequency, and DP criteria but could not belong to the Korean Learner's List general vocabulary. As a result, 512 words passed all four criteria. The example of the output is presented in Figure 4.2. The lemmas that were excluded by the KLL advanced list are presented in Appendix G.

	Column1	range	frequency	collocations	dispersion
1					
2	데이터	12	3161	데이터 마이닝, 데이터 버전, 데이터 링크, 계측 데이터, 데이터 전송, 데이터 통신, 입력 데이터	0.41
3	설계	12	2744	내진 설계, 참여 설계, 설계 기준, 도록 설계, 설계 단계, 건축 설계, 설계 변수, 기본 설계, 구조 설	0.24
4	측정	12	2290	압력 측정, 측정 장치, 측정 방법, 측정 결과, 측정 한다	0.28
5	장치	12	2016	단말 장치, 기억 장치, 장치 정비, 보호 장치, 저장 장치, 입력 장치, 처리 장치, 제어 장치, 실험 장	0.385
6	성능	12	1984	내진 성능, 내화 성능, 성능 보증, 조종 성능, 성능 향상, 성능 평가, 성능 개선, 제어 성능	0.2
7	효율	12	1906	변환 효율, 에너지 효율, 보다 효율	0.14
8	제어	12	1881	제어 이득, 열차 제어, 유동 제어, 전자 제어, 제어 시스템, 제어 장치, 산업 제어, 제어 기능, 공정	0.35
9	공정	12	1821	포토 공정, 생물 공정, 제조 공정, 분리 공정, 공정 조건, 공정 설비, 생산 공정, 처리 공정, 공정 저	0.49
10	요소	12	1633	유한 요소, 투입 요소, 구성 요소, 산출 요소, 위험 요소, 요소 기술	0.23
11	대상	12	1490	자극 대상, 대상 선박, 평가 대상, 분석 대상, 연구 대상	0.18
12	기존	12	1488	기존 건축물, 기존 연구	0.185
13	공학	12	1337	공학 회지, 토목 공학, 생물 공학, 인간 공학, 유전 공학, 화학 공학, 산업 공학, 전자 공학, 기계 공	0.375
14	하중	11	1316	익면 하중, 하중 조건	0.455
15	기반	12	1306	규칙 기반	0.24

Figure 4.2 Program output example (for advanced level of KEAVL)

Afterward, the contents of the lists were carefully examined. First, words that were excluded by the program were checked. It was done to prevent cases when a word with several meanings (usually SK words that are written the same in Korean but are represented by different Chinese origin words) is excluded by the program in one meaning, but in the corpus, it had another one. For instance, the word *phwungsok* was returned to the advanced part of KEAVL after being excluded by the advanced general vocabulary filter of the program. There, word *phwungsok* was used to mean 'custom'(風俗), while in the corpus *phwungsok* was used to mean 'wind speed'(風速).

Second, the words selected by the program were inspected in the corpus to be sure that they are not used in the different meanings in the corpus. Consequently, some changes were made to the lists. For instance, the frequency for the word *thongsang* was changed, and the DP was recalculated because, in two cases out of 119, this word was used to mean 'trade', while in other 117 cases, the word was used to mean 'usually, normally, generally'. Thus, in the list, the word is written as *thonsang 02* because, in the Standard Korean Language Dictionary, this meaning is presented under the number two. Similarly, range, frequency, and dispersion values of the words with the same issue were recalculated and changed.

Third, one case of different spelling versions of the English borrowing 'algorithm' was identified (*alkolicum* and *alkolitum*). The only *alkolicum* passed all three selection criteria, however, *alkolitum* was also frequent across the corpus. Some previous studies on technical engineering vocabulary also reported the same issues (Lee, 2014; Nergui, 2015). In their cases, the decision of what version to include to the list was based on the checking of the correctly built borrowing transliteration. That is, if the borrowing is transliterated into Korean incorrectly, the version was not included in the list. In the case of this study, since both versions are presented in the Standard Korean Language Dictionary, it was decided to include both spelling variants as one entry. Moreover, their parameters were recalculated for the final lists.

However, I had some doubts about whether to include verbs *pihata (比-)*'compare' and *inhata (因-)* 'be caused by'. The reason for this is that, usually, these verbs are being taught as grammatical units (*-lo/ulo inhae(se); -ey pihay(se)*) and presented in grammar sections of textbooks or in grammar books (An & Seon, 2013). Hence, students likely encounter them during general Korean classes. However, it was decided to include them to the engineering lists for several reasons: 1) advanced KLL has them included as words; 2) it is still possible that students have not encountered these verbs; 3) they are frequent across academic texts. Therefore, they are essential for intermediate level students.

As a result, the length of the lists after the manipulations stated above have changed. The final size of the full list is 830 lemmas where 318 lemmas are of the intermediate level and 512 lemmas are of the advanced level. The words are presented in alphabetical order in Tables 4.1 and 4.2.

Table 4.1

Full KEAVL (in alphabetical order)

가공 가동 가상 가설 가속 가시 가열 가중치 가치 가하다 각도 간격 간섭 간접 감성 감소 감쇠 감시 감지 강성 강재 강하 강화 갖추다 개념 개방 개별 개선 개소 개수 개시 개요 개정 개체 객체 거동 거래 건 건조 검증 검토 결함 결합 경계 경계층 경과 경로 경사 경쟁 경쟁력 경향 계수 계열 계층 고압 고온 고유 고정 고찰 곡선 공격 공공 공극 공급 공급자 공법 공식 공유 공정 공진 공통 공학 과도 과정 관계없이 관점 관측 광 광학 교차 구간 구동 구리 구매 구분 구성 구속 구역 구조 구조물 구체 구축 구현 구획 국부 규격 규명 규모 규정 균열 균일 균형 그래프 극대 극복 근거 근접 급격히 긍정 기관 기구 기기 기능 기동 기둥 기반 기법 기상 기업 기여 기인 기존 기체 기판 기후 난류 낮추다 내 내구 내부 내장 냉각 널리 노동 노드 노이즈 노출 누적 다루다 다소 다중 단계 단독 단말기 단면 단위 단일 단축 단층 달리 달성 당시 대가 대기 대략 대량 대류 대비 대상 대수 대안 대역폭 대응 대입 대책 대체 대체로 대칭 대형 더불다 더욱 데이터 도달 도식 도입 도출 독립 독성 동영상 동일 동적 동향 두께 등급 등장 디젤 라인 레벨 로봇 루프 링크 마이크 망 매립 멀티미디어 메모리 메커니즘 메탄 면적 명시 모니터링 모델링 모듈 모드 모멘트 모바일 모터 무선 무인 문서 문헌 물류 물리 물성 물체 미만 미세 민감 밀도 바이오 바탕 박리 반경 반도체 반면 반사 반영 반응기 발사 발현 방법론 방사 방식 방안 방정식 방지 방출 배수 배열 배출 배치 범위 법칙 베이스 벡터 벽면 변경 변동 변수 변위 변형 변환 별도 병렬 보급 보상 보수 보완 보유 보장 보정 보조 보존 복귀 복원 복합 볼트 부가 부담 부문 부식 부여 부위 부재 부착 부품 부피 부하 분류 분리 분말 분배 분산 분야 분포 분할 분해 불량 붕괴 블록 비대칭 비례 비율 비중 비하다 빈도 사건 사례 사료 사슬 사업자 사용량 사이클 사이트 사항 산업 산정 산출 산화 삶 상관 상기 상당 상당히 상부 상세 상수 상승 상업 상용 상위 상태 상호 상황 샘플 샘플링 생명 생물학 생산 생산량 생성 서버 석탄 선 선도 선로 선정 선행 선형 선호 설계 설문 설비 설정 설치 섬유 성능 성분 성장 성질 세대 세부 센서 셀 소모 소비 소성 소요 소자 소형 속성 손실 송수신 송신 수단 수렴 수리 수립 수면 수명 수분 수소 수송 수식 수신 수압 수요 수용 수익 수정 수준 수지 수직 수집 수축 수평 수행 순 순위 순차 순환 스마트 스마트폰 스펙트럼 승객 시각 시기 시나리오 시멘트 시뮬레이션 시설물 시스 시스템 시점 시편 시행 식별 신경망 신뢰 신뢰도 신속 실린더 실시 실시간 실용 실정 실제로 실질 실행 실험 실현 안정 알고리즘 알루미늄 알아보다 알칼리 압 압축 압축강도 압축기 앞서 액체 야기 양상 억제 언급 업체 여부 역량 역학 연간 연계 연관 연구원 연령 연료 연비 연성 연속 열전달 열화 영역 영향 영향력 예방 예비 예측 오염 오차 온라인 와류 완료 완성 완전 외기 외부 외피 요소 요약 요인 용기 용도 용량 용액 용어 용적 용해 우려 운반 운영 운용 운행 원가 원격 원료 원리 원소 원인 원자 원형 원활 위상 위협 유기물 유도 유동 유량 유무 유발 유사 유속 유용 유의 유입 유전 유지 유체 유출 유해 유형 유효 육상 음성 응답 응력 응용 의존 이내 이론 이산화 이온 이중 인가 인공 인력 인류 인쇄 인식 인자 인장 인접 인증 인체 인터페이스 인프라 인하다 일치 임계 임의 입력 입자 입출력 자극 자세 자원 자율 작동 작성 작업 작용 잔류 장기 장력 장비 장애 장애물 장착 장치 재생 재질 저감 저장 저하 저항 적응 적재 적정 전개 전극 전단 전략 전망 전반 전산 전용 전이 전진 전파 전형 전환 전후 절감 절단 절차 점성 점토 접근 접지 접촉 접합 정규 정량 정밀 정부 정수 정의 정적 정지 정책 정체 정확도 제거 제도 제시 제안 제약 제어 제어기 제외 제작 제정 제조 조류 조립 조성 조작 조절 조정 조직 조합 조화 존재 종료 좌우 주기 주도 주목 주민 주의 주파수 주행 줄어들다 중단 중력 중앙 증가 증대 증폭 지니다 지름 지배 지상 지속 지수 지식 지연 지원 지정 지지 지침 지향 지형 직경 직관 직류 직선 진단 진동 진동수 진입 진폭 질량 질소 집단 집합 차별 차원 차지 참고 참여 참조 창출 채취 채택 처리 척도 천연 천연가스 철골 철도 체계 체적 초과 초기 초래 초점 촉진 최소 최적 최적화 최종 추가 추구 추세 추정 추진 추출 축 축소 축적 출력 충돌 충족 취급 취득 취약 측 측면 측정 치수 침투 칩 컴퓨팅 케이블 코드 코어 클러스터 탄성 탄소 탐색 탑재 터빈 토대 통계 통과 통상 통신망 통제 통합 투입 투자 특성 특수 특이 특정 특허 파괴 파라미터 파악 파장 파형 판 패턴 펌프 편의 편차 평가 평균 평균값 평면 평행 평형 폐기물 포트 포화 폼 표면 표본 표적 표준 품질 풍속 프레임 프로젝트 플랜트 플랫 피드백 피크 피해 필드 필수 필터 하류 하부 하위 하중 하천 학문 한계 한정 할당 함량 함수 함유 합금 합리 합성 항공 항목 해당 해상 해양 핵심 행렬 행위 행하다 향상 향후 현상 현실 현장 현황 협력 형성 형식 형태 혼합 홉 화면 화물 화학 화합물 확률 확보 확산 확장 환기 활발히 회로 회수 회전수 획득 횟수 횡 효율 훈련 흐름 흡수 흡입

Table 4.2

Advanced KEAVL (in alphabetical order)

가공 가동 가설 가속 가시 가열 가중치 각도 감성 감쇠 감시 감지 강성 강재 강하 개념 개소 개수 개시 개요 개정 개체 객체 거동 거래 검증 결함 결합 경계층 경과 경로 경사 계수 계열 고압 고온 고유 고정 고찰 곡선 공극 공급자 공법 공유 공정 공진 공학 과도 관측 광 광학 교차 구간 구동 구리 구배 구조물 구체 구축 구현 구획 국부 규격 규명 균열 균일 그래프 극대 근접 긍정 기동 기반 기상 기인 기존 기체 기판 난류 내구 내장 냉각 노드 노이즈 노출 누적 다중 단말기 단면 단일 단축 단층 달성 대류 대상 대수 대안 대역폭 대칭 데이터 도식 도출 독성 동영상 동적 동향 등급 디젤 레벨 루프 링크 망 매립 멀티미디어 메모리 메커니즘 메탄 명시 모니터링 모델링 모듈 모드 모멘트 모바일 모터 무선 무인 문헌 물류 물리 물성 미세 민감 밀도 바이오 박리 반경 반도체 반사 반응기 발사 발현 방법론 방사 방정식 방출 배수 배열 배출 베이스 벡터 벽면 변수 변위 변형 변환 병렬 보급 보유 보정 복원 복귀 복합 볼트 부가 부식 부여 부착 부하 분류 분말 분배 분산 분할 분해 불량 붕괴 블록 비대칭 비례 빈도 사료 사슬 사용량 사이클 사이트 사항 산정 산출 산화 상기 상부 상세 상수 상승 상용 상위 상호 샘플 샘플링 생물학 생산량 생성 서버 석탄 선도 선로 선행 선형 설계 설비 설정 섬유 성능 성분 세부 센서 셀 소모 소성 소요 소자 속성 송수신 송신 수렴 수립 수분 수소 수송 수식 수신 수압 수용 수익 수정 수지 수직 수축 수치 수평 순 순차 순환 스마트 스마트폰 스펙트럼 시뮬레이션 시설물 시스 시편 시행 식별 신경망 신뢰 신뢰도 실린더 실시간 실용 실질 실행 알고리즘 알칼리 압 압축 압축강도 압축기 액체 야기 양상 억제 언급 여부 역량 역학 연계 연료 연비 연성 열전달 열화 오차 와류 완료 완전 외기 외피 요소 요인 용량 용액 용적 용해 운용 원가 원격 원료 원리 원소 원자 원형 원활 위상 유기물 유도 유동 유량 유무 유속 유용 유입 유전 유체 유출 유해 유효 응력 응용 이론 이산화 이온 인가 인력 인자 인장 인접 인증 인터페이스 인프라 임계 임의 입자 입출력 작동 잔류 장기 장력 장애 장애물 장착 장치 재질 저감 저장 저하 저항 적재 적정 전극 전단 전략 전산 전이 전진 전파 전형 절감 절단 점성 점토 접지 접합 정규 정량 정밀 정수 정의 정적 정책 정체 정확도 제도 제어 제어기 제정 제조 조류 조립 조성 조작 조합 조화 종료 주기 주도 주목 주파수 주행 중력 증대 증폭 지름 지상 지속 지수 지연 지정 지침 지향 지형 직경 직관 직류 진동수 진입 진폭 질량 질소 집합 차원 차지 창출 채취 채택 척도 천연 천연가스 철골 체계 체적 초과 초래 촉진 최적 최적화 추구 추세 추정 추출 축 축적 출력 충족 취급 취득 취약 측 측면 측정 치수 침투 칩 컴퓨팅 케이블 코어 클러스터 탄성 탄소 탐색 탑재 터빈 통상 통신망 투입 특허 파라미터 파장 파형 패턴 펌프 편차 평균값 평면 평행 평형 폐기물 포트 포화 폼 표본 표적 프레임 프로젝트 플랜트 플랫 피드백 피크 필드 필터 하류 하부 하위 하중 학문 할당 함량 함수 함유 합금 합리 합성 항목 해상 해양 행렬 행하다 향후 현황 혼합 홉 화물 화합물 확률 확산 환기 회로 회수 회전수 획득 횡 효율 흡수 흡입

The complete list is presented in Appendix H. The first 318 words of the table are those words that are necessary only for the intermediate level learners. The following 512 words filled with grey colour are necessary for both intermediate and advanced level learners. Both parts are sorted in alphabetical order (see Appendix H).

There are six columns in the resulting table (e.g., Table 4.3). The first and the second columns contain the number of the word and the word itself word as well as the number of the meaning if applicable. The third column presents the origin word for SK vocabulary and other borrowings for additional clarification of the word meaning and etymology. The fourth column includes all the modified lemmas for SK nouns (versions with *-hata*, *-toyta*, *-sikhita*, *-cek*). The

fifth column contains collocations found by the program (see Section 3.2). The version of the table with all identified collocations is published online (see detailed information below). The last column provides translation of the word and its modified lemmas to the English language. Each translation was added manually and were taken from the ET-house Neungyule Korean-English Dictionary (Park, 2006) and National Institute of Korean Language's Korean-English Learner's Dictionary (https://krdict.korean.go.kr/eng/mainAction). For SK nouns which occurred in the corpus in both noun and predicative forms, the translation of both forms to English is given.

Table 4.3

Example of the resulting table

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
1	가상	假想	가상적 가상하다	가상 계측, 가상 경로점	imagination, virtuality; (-적) imaginary; (-하다) to imagine
2	가치	價値		가치 사슬	value, worth
3	가하다 02	加--			to give, to inflict, to deliver, to deal, to apply

In order to make work with the lists more convenient and to make them more accessible, the multiple versions of both lists were published online (www.topik-tips.com). The variants with words sorted by frequency of occurrence, by DP, and in alphabetical order are available. Furthermore, both lists are published on the learning platform *Quizlet* (https://quizlet.com/class/13685648/).

As for the contents of the lists, the most frequent words in the full list are *sisutheym* 'system', *cungka (增加)* 'increase, growth', *teyithe* 'data', *kwuco (構造)* 'structure, organization', *selkyey (設計)* 'plan, map, design', *thukseng (特性)* 'characteristic', *swuhayng (遂行)* 'fulfillment, performance', *kwuseng (構成)* 'composition, organization', *pangsik (方式)* 'method'. The least

frequent words in the full list are *halyu (下流)* 'low-stream', *sungkayk (乘客)* 'passenger', *kayso (個 所)* 'place, site', *cwuuy (注意)* 'attention, caution', *phwungsok (風速)* 'wind speed', *kongsik (公式)* 'rule, formula', *taycheylo (大體-)* 'generally, mostly', *mwullyu (物流)* 'distribution', *sayngmyeng (生命)* 'life', *cekung (適應)* 'adaptation, adjustment'.

For the advanced part of KEAVL the most frequent words are *teyithe* 'data', *selkyey (設計)* 'plan, map, design', *chukceng (測定)* 'measurement', *cangchi (裝置)* 'installation', *sengnung (性能)* 'performance', *hyoyul (效率)* 'efficiency', *ceye (制御)* 'control', *kongceng (工程)* 'process', *yoso (要素)* 'element, factor', *taysang (對象)* 'object'. Among the least frequent words were *wenca (原 子)* 'atom', *myengsi (明視)* 'clear statement', *thongsinmang (通信網)* 'communicational network', *phothu* 'port', *tanchung (單層)* 'cross-sectional', *kwuli* 'copper', *polthu* 'bolt', *khemphyuthing* 'computing'.

Although the lists contain mostly SK vocabulary, English borrowings and native words have also occurred (see Figure 4.3). Among frequent English borrowings there are already mentioned *sisutheym* 'system', *teyithe* 'data', and *khonkhulithu* 'concrete', *pullok* 'block', *simyulleyisyen* 'simulation', *seynse* 'sensor'. As mentioned above, the etymology of the words is indicated in the third column of the final list (see Appendix H).
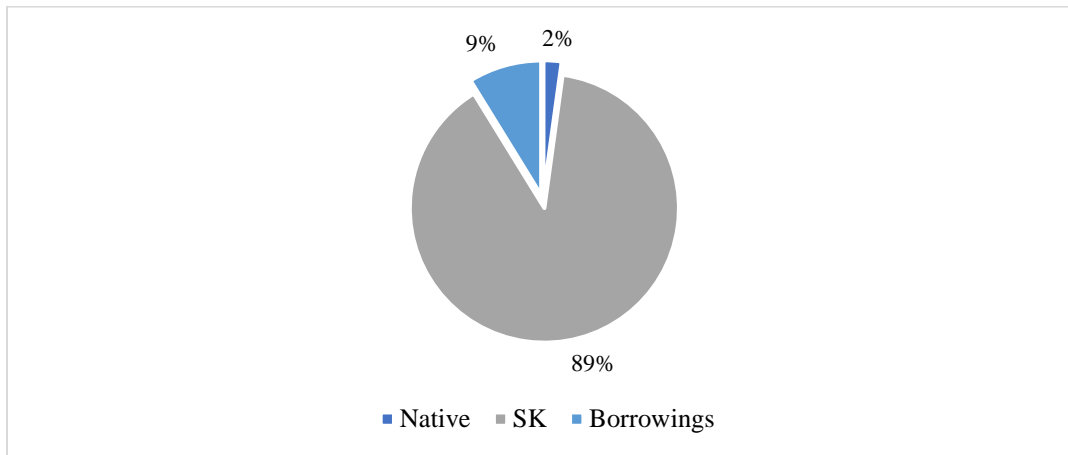
Figure 4.3 Vocabulary types in the full KEAVL

Regarding parts of speech, the KEAVL contains mostly nouns, but also has seven native predicates (e.g., *talwuta* 'handle, deal with', *nacchwuta* 'reduce, drop') and a few adverbs (e.g., *tewuk* 'more and more, increasingly', *nelli* 'widely') as shown in Figure 4.3. However, all these verbs and adverbs are not presented in the advanced version of the lists because they were excluded by the filter. Additionally, for interpreting these numbers, it is important to remember that the majority of the SK adjectival and verbal nouns may be represented as predicates as well (see the fourth column of Appendix H).
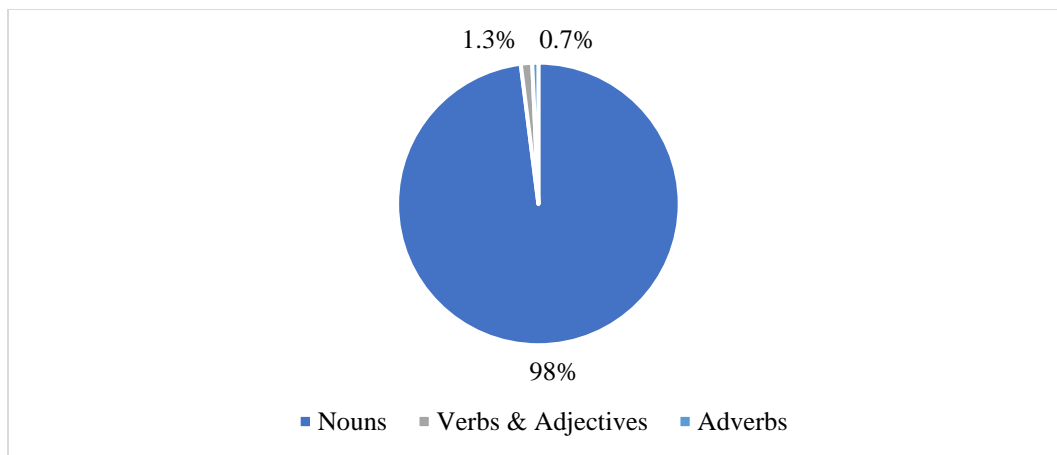
Figure 4.4 Parts of speech in the full KEAVL

According to the calculations made by the program, the coverage of the KEAVL in the corpus was 24.6% among the tokens and 2.7% among the lemmas. Although the minor

51

manipulations were done to the list, the coverage results almost have not changed (Table 4.4). The coverage by the A, B levels of KLL is also shown in Table 4.4.

Table 4.4

Coverage of CKEAT by KLL and KEAVL

List	Tokens covered	% of tokens covered	Lemmas covered	% of lemmas covered
KLL listA	308,537	23.6%	715	2.39%
KLL listB	349,219	26.8%	1,502	5.03%
KEAVL	321,674	24.6%	830	2.7%

Thus, by mastering intermediate vocabulary and words from Korean Engineering Academic Vocabulary List, students are expected to understand up to 75% of words in Corpus of Korean Engineering Academic Texts. In order to provide a more illustrative example of the coverage, two random passages from the corpus were selected. The words covered by the KEAVL are bold-faced, the words covered by A, B levels of KLL are underlined. KEAVL covers 21 out of 59 words in the first passage and 25 out of 72 Korean words in the second passage. The coverage of the KEAVL in the selected excerpts is 35.6% and 34.7% accordingly. Hence, the coverage of these passages even higher than the average coverage of the corpus.

Passage 1

"축류식 선풍기에서 발생하는 압력 **손실**은 **회수** 가능한 부분과 **회수** 불가능 부분으로 **구분할** 수 있다. **회수** 가능한 **손실**은 선풍기에서 토출되는 **기체 내의 와류** 및 회전에 따른 **손실** 부분이며 안내 날개의 **최적 설계**로 **회수** 가능하다, 선풍기의 베어링, 마찰 **손실**이나 케이싱, 회전 차의 허브, 지지빔, 날개 자체에서의 **손실**은 **회수** 불가능하다. 이와 같은 기계적 에너지의 **손실**은 열 에너지로 **전환**되어 **유용**한 일 에너지의 **손실**을 **초래**한다."

'Pressure losses from axial fans can be divided into recoverable and nonrecoverable parts. The recoverable loss caused by the vortex and rotation in the gas discharged from the fan and is recoverable by the optimal design of the guide wing. The loss of friction or casing of the fan, loss of the hub, support beam, and the wing itself from the rotating car is not recoverable. This loss of mechanical energy is converted into thermal energy, resulting in a loss of useful work energy.'

Passage 2

"다공성 멤브레인 **기반 바이오 센서**의 경우 셀룰로오스로 **구성된 샘플** 패드는 **샘플**의 불순물을 걸러내고, 분석 완충액 (assay buffer)을 **건조 상태**로 **저장하고** 있다. 유리 **섬유로 구성된 결합** 패드는 라벨용 건식 시약 **저장**소로서 사용되며, 이 위치에서 라벨용 시약과 **바이오** 마커의 **특이 결합** 반응이 일어난다. 니트로셀룰로오스로 이루어진 검출 패드에서는 캡쳐 시약이 **고정**되어 있고, 캡쳐 시약은 라벨용 시약과 **결합한 바이오** 마커와 다시 **결합한 후** 신호를 **출력하게** 된다. 마지막으로 셀룰로우스 **필터**는 **흡수** 패드로 사용되며, 이를 통해 **유체를 회수하는** 역할을 **수행한다**."

'In the case of a porous membrane-based biosensor, a sample pad composed of cellulose filters out impurities in the sample and stores an assay buffer in a dry state. A binding pad consisting of glass fibers is used as a dry reagent reservoir for the label, where a specific binding reaction of the labeling reagent and the biomarker takes place. In the detection pad made of nitrocellulose, the capture reagent is fixed, and the capture reagent recombines with the biomarker combined with the reagent for labeling and then outputs a signal. Finally, the cellulose filter is used as an absorption pad, which serves to recover the fluid.'

Even without the words from KLL, the advanced part of KEAVL covers 12.1% of CKEAT. The coverage of the full KLL in the corpus is 77.1%. Thus, knowing the full KLL and advanced part of KEAVL may help to understand up to 89% of the CKEAT lemmas (see Table 4.5).

Table 4.5

Coverage of CKEAT by the KLL and Advanced part of KEAVL

List	Tokens covered	% of tokens covered	Lemmas covered	% of lemmas covered
KLL listA	308,537	23.6%	715	2.39%
KLL listB	349,219	26.8%	1,502	5.03%
KLL listC	347,779	26.7	1844	6.2
Adv KEAVL	158,700	12.1	512	1.7

For the comparison, the coverage of the advanced AVL and full KLL in the text is presented in Passage 3 that was also randomly selected from the corpus. Eight out of 62 Korean words in the passage are in the advanced level of KEAVL. The coverage in this passage is 13%, supporting the results of the study.

Passage 3

라지 스케일 인지 **시뮬레이션**은 최근 들어 연산 뇌 과학, **시뮬레이션 방법론** 및 슈퍼 **컴퓨팅** 분야의 학제간 연구 분야이다. 두뇌 인지 컴퓨터 분야에서 인지 시뮬레이터는 두뇌 구조, 동작 및 기능에 대한 **가설** 실험을 가능하게 하는 중요한 기술이다. **시뮬레이션**은 또한 새로운 시냅스 나노 디바이스를 사용한 컴팩트 저전력 뉴로모픽 시냅트로닉 **칩**들의 혁신적 시스템을 불러일으키고자 하는 야심 찬 목표를 갖고 있는 DARPA′s SyNAPSE (Systems of Neuromorphic Adaptive Plastic Scalable Electronics) 프로그램 같은 최첨단 연구 **항목**들의 집합체이다.

'Large scale cognitive simulation has recently been an interdisciplinary field of computational brain science, simulation methodology, and supercomputing. In the field of

brain cognition computers, cognitive simulators are an important technology that enables hypothesis testing of brain structure, behavior, and function. Simulation is also a collection of cutting-edge research items such as the DARPA's SYNAPSE (Systems of Neuromorphic Adaptive Plastic Scalable Electronics) program, which has an ambitious goal to generate innovative systems for compact, low-power, neuromorphic synaptic chips using new synaptic nanodevices.'

Regarding the range of occurrence of the words across the sub-fields, the majority of words in both lists occurred in all 12 sub-fields. The minority of the words occurred only in 6 to 7 sub-fields. The detailed comparison is presented in Figure 4.5.

Furthermore, the coverage of the words by their frequency was inspected. According to the numbers shown in Table 4.6, the hundred most frequent lemmas in the corpus cover 11% of CKEAT that is almost half of the total coverage of KEAVL in CKEAT.
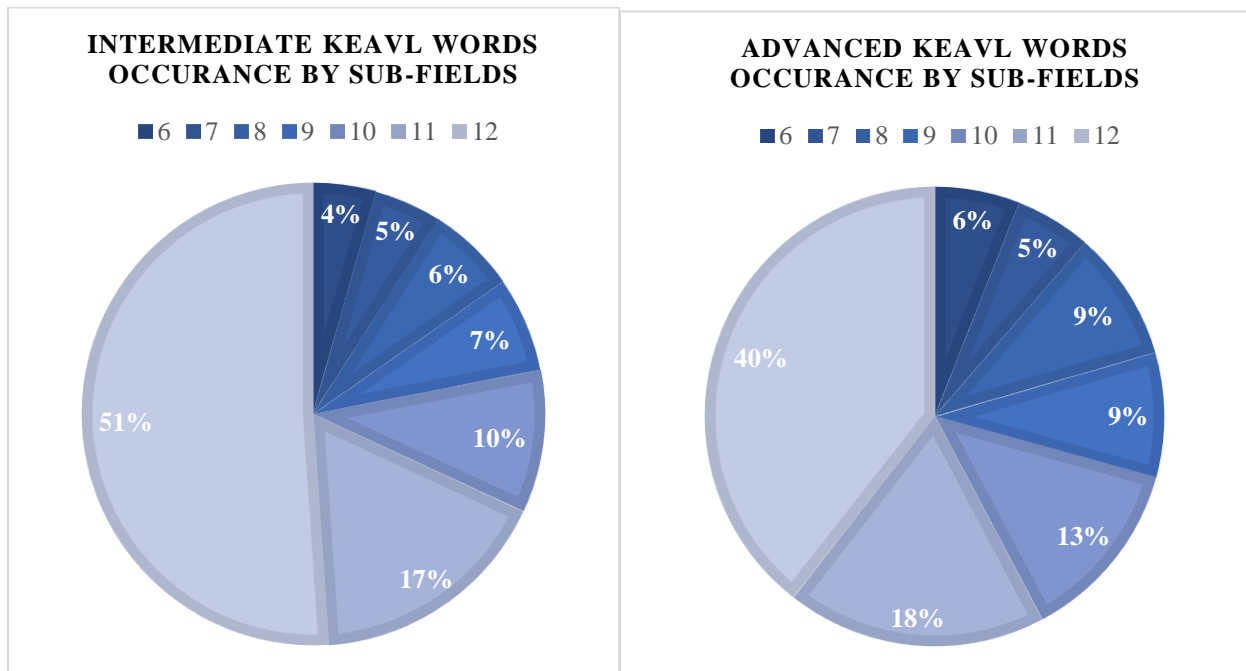
Figure 4.5 Comparison of the range of words occurrence for both KEAVL versions

Table 4.6

Coverage of KEAVL lemmas in CKEAT

Words (frequency order)	Frequency	% of CKEAT covered
1-100	137,865	10.6%
101-200	57,932	4.5%
201-300	35,635	2.7%
301-400	25,912	2%
401-500	20,639	1.6%
501-600	16,278	1.3%
601-700	13,318	1.02%
701-830	14,245	1.09%

To conclude, this research has identified 830 words that frequently occur across academic texts in the field of engineering. Among them, 502 words are not represented in the list of the general vocabulary of Korean (KLL). The coverage of the full KEAVL in the compiled CKEAT is 24.6% and the coverage of the advanced part only is 12.1%. KEAVL, along with A, B levels of KLL, can cover up to 75% of CKEAT. The advanced part of KEAVL, along with full KLL, can cover up to 88%.

4.2 Collocations for the KEAVL words

The second research question of the present study aimed to identify the lexical items that often co-occur with the words in KEAVL and provide a description of the overall tendencies. To do so, the function was added to the program to search collocations of the words identified for KEAVL. As was stated in Section 3.2, the collocations were selected according to the frequency criterion that was set on ≥30 occurrences in CKEAT. Not only bi-grams but also tri-grams were searched by the program. It was also decided to look for both left-side and right-side collocations.

As a result, the collocations for 311 lemmas in the list were found. Mostly, the more frequent a word in the corpus, the more collocations it takes. Some lemmas like *allwuminyum* 'aluminum' are used in different contexts each time, thus, some words in CKEAT naturally do not have frequent collocations. The number of collocates per lemma varies from 1 to 18. The majority of the words have only one frequent collocation found. Nevertheless, some lemmas in CKEAT have more than six collocations. The detailed number of collocations is presented in Figure 4.6.

In some cases, words collocate with other words from the list. In these cases, the word collocation is marked with the number of it, as is in Appendix H. For instance, word *kyeyswu (係 數)*'coefficient, modulus' (number 351 in Appendix H) collocates with the word *thanseng (彈性)* 'elasticity' (number 757 in Appendix H) that is also presented in the list. Therefore, in the 'collocations' column for *kyeyswu*, the phrase "thanseng (757) kyeyswu" has a number of the *thanseng* entry in the list (indicated both in Table 4.6 and Appendix H).
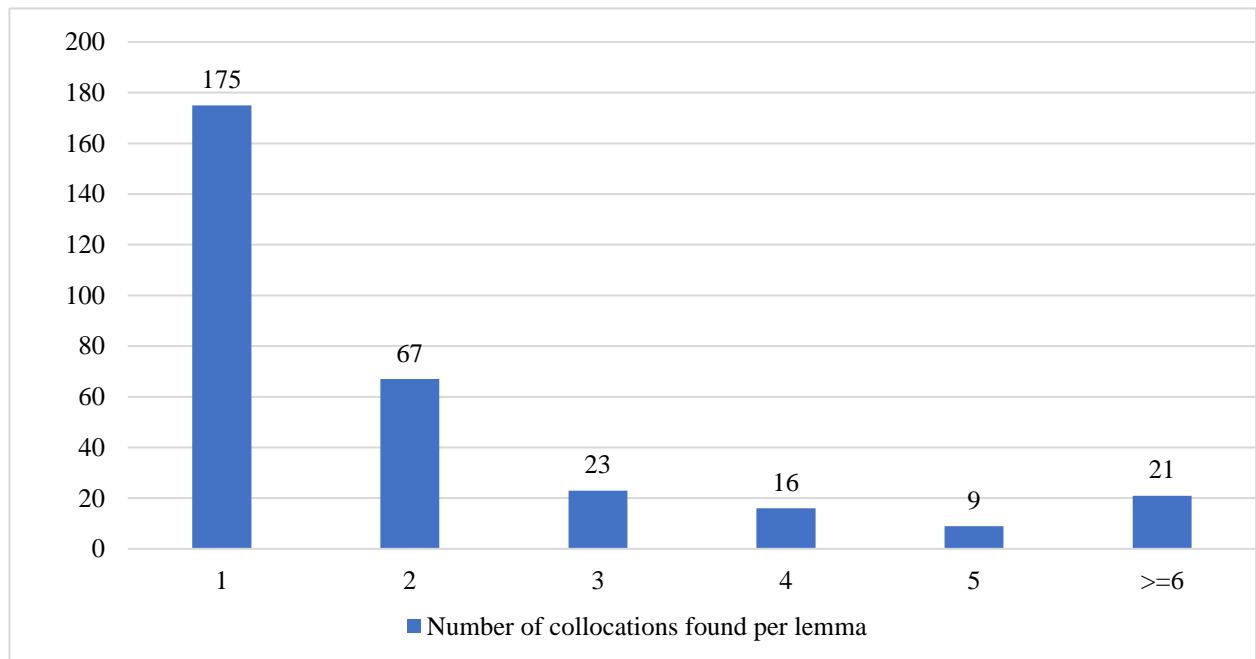
Figure 4.6 KEAVL lemmas with the highest number of collocations in CKEAT

In Table 4.7, the collocations of the top-10 frequent words on KAEVL are presented with the English translations of the phrases. Analyzing the examples given in Table 4.7, some KEAVL words which collocate with each other can be noticed. Especially, it is the case of the most frequent words in the list.

Summarizing the above, the present study was able to find lexical items that often co-occur with more than a third of the lemmas in KEAVL. The information about collocations is presented in Appendix H and online (www.topik-tips.com).

Table 4.7

Collocations of 10 most frequent words in the KEAVL

Word	Collocations
시스템 'system'	결재 시스템 'authorization system', 모니터링 시스템 'monitoring system', 탐지 시스템 'detection system', 제어 시스템 'control system', 데이터베이스 시스템 'database system', 파일 시스템 'file system', 방지 시스템 'prevention system', 관리 시스템 'management system', 통신 시스템 'system of communication', 컴퓨터 시스템 'computer system', 전체 시스템 'whole system', 생산 시스템 'production system', 시스템 개발 'system development', 정보 시스템 'information system', 시스템 구성 'system configuration', 구조 시스템 'structural system', 시스템 설계 'system design', 평가 시스템 'evaluation system'
처리 'processing, handling, disposal'	하수 처리 'wastewater disposal', 폐수 처리 'wastewater disposal', 중앙 처리, 처리 시설, 처리 장치 'processing unit', 신호 처리 'signal processing', 처리 과정 'management process', 처리 공정 'handling process', 정보 처리 'information processing', 데이터 처리 'data processing', 처리 속도 'processing speed', 처리 시간 'processing time', 처리 방법 'processing method', 처리 기술 'processing technology'
공정 'process'	생물 공정 'bioprocess', 제조 공정 'manufacturing process', 분리 공정 'separation process', 공정 조건 'process condition', 공정 설비, 생산 공정 'production process', 처리 공정 'handling process', 공정 제어 'process control'
장치 'equipment, device, installation'	단말장치, 기억 장치 'memory unit', 장치 정비 'service device', 보호 장치 'protective device', 저장 장치 'storage device', 입력 장치 'input device', 처리 장치 'processing unit', 제어 장치 'control system', 실험 장치 'experimental equipment', 측정 장치 'measuring equipment'

Word	Collocations
공학 'engineering'	공학 회지 'engineering journal', 토목 공학 'civil engineering', 생물 공학 'bioengineering', 인간 공학 'human engineering', 유전 공학 'genetic engineering', 화학 공학 'chemical engineering', 산업 공학 'industrial engineering', 전자 공학 'electronic engineering', 기계 공학 'mechanical engineering', 공학 분야 'engineering field', 자동차 공학 'automotive engineering', 환경 공학 'environmental engineering'
데이터 'data'	데이터 마이닝 'data mining', 데이터 버전 'version of data', 데이터 링크 'data link', 계측 데이터 'data measurement', 데이터 전송 'data transmission', 데이터 통신 'data communications', 입력 데이터 'input data', 데이터 처리 'data processing', 데이터 모델 'data model', 실험 데이터 'experimental data', 데이터 분석 'data analysis'
평가 'evaluation'	성능 평가 'performance evaluation', 평가 항목 'evaluation criteria', 평가 단위 'unit of valuation', 평가 기준 'valuation basis', 평가 대상 'subject of evaluation', 평가 방법 'method of evaluation', 평가 결과 'evaluation result', 평가 시스템 'evaluation system'
제어 'control'	제어 이득 'gain control', 열차 제어 'train control', 유동 제어 'flow control', 전자 제어 'electronic control', 제어 시스템 'control system', 제어 장치 'control device', 산업 제어 'industry control', 제어 기능 'control function', 공정 제어 'process control', 제어 성능 'control performance', 제어 특성 'control characteristics', 제어 가능 'control function', 제어 기술 'control technology'
성능 'performance'	내진 성능 'seismic performance', 내화 성능 'fire resisting capacity', 성능 보증 'performance warranty', 조종 성능 'maneuvering performance', 성능 향상 'performance improvement', 성능 평가 'performance evaluation', 성능 개선 'performance improvement', 제어 성능 'control performance'
품질 'quality'	품질 경영 'quality management', 품질 관리 'quality control', 서비스 품질 'quality of service', 품질 향상 'quality improvement', 품질 모형 'quality model', 품질 비용 'quality cost', 개발 품질 'development quality'
오염 'pollution'	대기 오염 'air pollution', 오염 물질 'pollutant', 수질 오염 'water pollution', 토양 오염 'soil pollution', 오염 배출 'polluting emissions', 환경 오염 'environmental pollution'

4.3 Comparison with the Korean Academic Word List

The third research question of the current study was whether there are high-frequency words in CKEAT that were not identified in the KAVL. In other words, to what extent the coverage of CKEAT by KEAVL differs from the coverage by KAVL. By comparing the coverage of the two lists, it is possible to determine whether KEAVL actually contains words specific to the academic texts in the engineering field and for academic language in general.

To answer this question, first, both lists were compared to measure overlap. The KAVL contains 916 lemmas, while the KEAVL contains 830 lemmas. There are 240 words that overlap between KAVL and KEAVL. Among the words that are both in KEAVL and KAVL but are not in KLL are, for instance, *selceng (設定)* 'set-up', *kyumyeng (糾明)* 'investigation', *nochwul (露出)* 'exposure, disclosure', *ceha (低下)* 'fall, decline', *kipan (基盤)* 'base', *yenkyey (連繫)* 'connection, link', *cangay (障㝵)* 'impairment, barrier', *hwaklyul (確率)* 'probability', *pilyey (比例)* 'proportion', *celcha (節次)* 'procedure'. However, the majority of the overlapping words are outside the hundred most frequent words of KEAVL.

The analysis of the coverage of CKEAT by both lists also suggests that KEAVL covers more tokens and more lemmas in the corpus in comparison to KAVL, as shown in Table 4.8. At the same time, the advanced KAVL has a smaller coverage in comparison to KAVL. Nevertheless, in order to interpret these results correctly, it is also important to remember that the KAVL has some intersections with the KLL words that were previously excluded partially or entirely from KEAVL.

Table 4.8

Comparison of KAVL and KEAVL coverage in CKEAT

List	Tokens covered	% of tokens covered	Lemmas covered	% of lemmas covered
KAVL	281,532	22.1%	635	2%
Full KEAVL	322,052	25.6%	830	2.7%
Adv KEAVL	158,700	12.1%	512	1.7%

As shown in Table 4.8, 635 out of 916 KAVL lemmas occurred in the corpus. Among these words, only 313 passed the criteria of the current study (range, frequency, and DP). Thus, the answer to the third research question is that present research has identified words that are frequent across CKEAT but are not included in KAVL. Nevertheless, there is a small overlap between the lists. The coverage of CKEAT by full KEAVL is also higher than KAVL's coverage. In other words, the majority of the selected words are frequent across different sub-fields of engineering but do not frequently occur in all sub-corpora of the Corpus of Academic Texts of Korean.

The results presented in this chapter will be discussed in detail in the next section of the present thesis.

5   Discussion and Conclusion

This study was conducted to compile the word lists for international students of engineering colleges in Korea. This final chapter of the thesis restates the research questions and reviews the major methods of the study. The major sections of this chapter summarize the results and discuss their implications.

As mentioned in Chapter 1, the current thesis was an attempt to contribute to the field of Korean for Academic Purposes by compiling the vocabulary list for the students of engineering majors. As described in Chapter 2, the list was built based on a Corpus of Korean Engineering Academic Texts (or CKEAT) from 12 main sub-fields of engineering made for this study. The corpus was then processed with a Python code, written by the author, that was made to generate the Korean language word lists based on the set criteria. For this study, the main criteria were the word range (occurrences in more than half of sub-fields), its frequency ($\geq$ 100 times across the corpus), and its variability (common across $\geq$ 6 sub-fields). Moreover, I prepared a list for intermediate-level and advanced level learners. For inclusion to the first list, a word must not occur in the beginner and intermediate general Korean vocabulary (A and B levels in Korean Learner's List). For inclusion in the second list, a word must not occur in the Korean Learner's List.

5.1   KEAVL

The current thesis presented KEAVL with intermediate and advanced parts. The full list contains 830 lemmas and covers 25% of the corpus. The advanced part contains 512 lemmas and covers 12% of the corpus. The list was compiled, taking into account such basic criteria as word range and frequency. The same criteria were used by the majority of the previous studies mentioned in Section 2.3.2. Nevertheless, unlike in the other studies, the present research used word dispersion value as an additional criterion for a word to be included in the list. The dispersion

value (in this study deviation of proportions proposed by Gries (2008)) allowed to identify how equally the word is distributed across sub-corpora. Hence, it provided more precise results for the word lists by selecting only words that appeared evenly between sub-corpora. Moreover, the calculation of DP allowed taking into account the frequency of occurrence of words in relation to the size of the sub-corpus where they appeared. It also made the results of the current study more accurate since the words that were frequent only in one or several sub-corpora were excluded.

The importance of the dispersion measurements for the word lists creation can be better demonstrated with an illustrative example: the output for the KEAVL compiled by the program without dispersion filter contained 1189 words, while with dispersion filter activated, it contains 830 words, thus, 359 words were excluded based on the dispersion criterion. Among the excluded words, for instance, was *phokpal (暴發)* 'explosion' that occurred in all sub-fields with a total frequency of 366. At the same time, the word *wenli (原理)* 'principle, fundamentals' having the same range and frequency was included in the list. However, if we look at Figure 5.1, which represents a word's variability across the sub-corpora, we observe a clear difference in word usage. Thus, the dispersion measurement is crucial for word list compilation and should be considered in future research devoted to vocabulary list compilation.
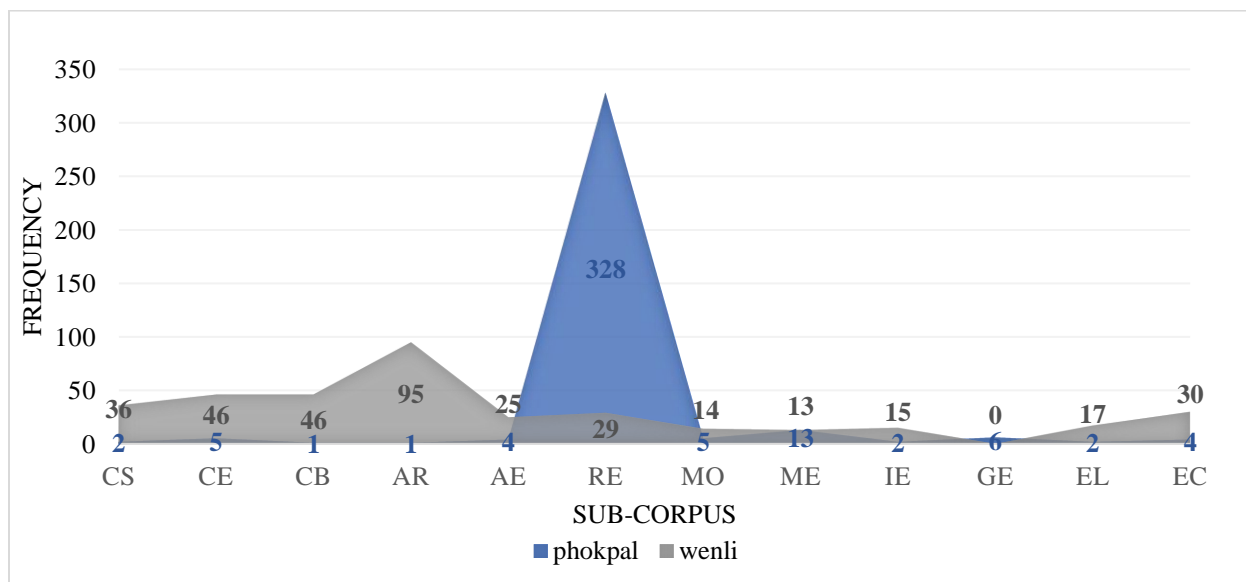
Figure 5.1 Difference in distribution between two words with the same frequency

Furthermore, the current study differs from the previous studies because it provides word lists for two different levels of proficiency. The decision to build two lists was motivated by the fact that not all students start their degree in Korea with an advanced level of proficiency. The requirement for language proficiency in the Korean-language programs is usually 3rd or 4th level (or *gup*) (B1 or B2, according to CEFR (Won, 2016)) which is equal to the intermediate level of proficiency. Therefore, if the general vocabulary was completely filtered out to build the list, students with an intermediate level of proficiency of Korean would have to learn some advanced general vocabulary words that frequently occur across engineering texts on their own. That is the reason why the KEAVL contains two parts: intermediate and advanced, where intermediate part has some general vocabulary from the advanced level of KLL. Hence, by preparing KEAVL, it is possible to give students required engineering vocabulary focus while at the same time, enrich their advanced general vocabulary level. The advanced part, at the same time, was prepared for

the highly proficient students who need to focus on the engineering-related words that rarely occur across general language materials.

Another important feature of the lists is that they contain not only nouns but also verbs and adverbs. Although the lists contain mostly Sino-Korean lemmas as it was described in Chapter 3, some native Korean nouns, verbs, and adverbs also appeared in them. Some previous researches considered only nouns (J. Park, 2018), and some added native verbs and adjectives as collocations of the nouns (Nergui, 2015; Yeo, 2018). Notwithstanding, although native Korean verbs and adjectives may rarely occur across academic texts, I believe that it is still important to take them into account while compiling word lists. Therefore, *KWordList*, which was used for the compilation of the lists, was not only set to retrieve nouns but also for the verbs, adjectives, and adverbs. As a result of other POS inclusion, the intermediate level of KEAVL contains eight native verbs. Furthermore, the native adverb *tewuk* 'more and more, increasingly', which was included in the KEAVL, was one of the most equally distributed words across sub-corpora (DP = 0.13). This once again shows the importance of considering other parts of speech for inclusion in the word lists.

Another advantage of the list is that collocations are also provided for the majority of the most frequent words in the list. It makes the list very useful since it gives not just a word but also a phrase in which this word frequently occurs. This, in turn, greatly simplifies the process of learning a new word because it gives a student a ready-made word sequence (Sinclair, 2004).

Concerning the implications of the lists, there are several areas for their application. First of all, lists may be used for the development of Korean teaching materials in the area of engineering. These words can potentially set up a vocabulary basis for special vocabulary workbooks and training materials and be a reference for the Korean for engineering purposes textbooks.

Second, these lists may be useful for curriculum developers. Since the number of international students in Korea is growing steadily, I believe that more pre-university Korean language preparation courses will appear. Therefore, these word lists may be used as materials for pre-university curriculum development (for future engineering college students) by setting the vocabulary focus of a program.

Third, the compiled lists may be essential for self-learners of Korean, and particularly for the self-learners who are getting ready to apply to engineering colleges in South Korea. Students who prepare for university admissions and study Korean academic vocabulary their given field on their own are often not taken into account by the word list creators as potential users lists' beneficiaries. At the same time, word lists can help such students no less than those who study Korean by the developed curriculum. Unfortunately, many existing word lists, although available online and in the appendices of the thesis, are difficult to find by the students who are not fluent in Korean. Therefore, it was decided to make these lists as accessible on the Internet as possible. KEAVLs are fully available online (www.topik-tips.com), and they are searchable both in English and Korean.

Furthermore, since many language students now search for word lists and learn words on the online platform *Quizlet*, for their convenience, words from the lists were divided into small groups and published as a part of a single course on a *Quizlet* platform (https://quizlet.com/class/13685648/). The most common word collocations were also included in the course. Therefore, the *Quizlet* course and its availability on the website make these word lists more available for future engineering students in South Korea from all over the world. Additionally, it helps students who prefer the "cards" method of vocabulary learning by providing them with the vocabulary in this format.

Forth, the lists may be used as small dictionaries containing the most frequently occurring engineering-related words not only by students but also by engineers who want to work or are already working in South Korea. Furthermore, lists may be very useful for translators who work with engineering companies and factories.

To sum up, the developed lists may be used for different purposes, from curriculum development to self-learning. Moreover, they were made as widely available online as possible.

5.2   KEAVL and other lists for engineering students

This study is the first attempt to build a list of the Korean academic vocabulary that is common for all main engineering subfields. Before, only separate vocabulary lists for engineering sub-fields and PVL for engineering students were compiled. Naturally, it is important to provide vocabulary help to the engineering students who have already claimed their specific major in engineering colleges. Moreover, it is equally important to provide adequate preparation for first-year engineering physics classes, since the physics vocabulary is highly specific. Nevertheless, as it was stated in Chapters 1 and 2, some students may require some general engineering language preparation. For instance, cases when a student is just preparing to go to the engineering college and is not yet sure which major they will choose. Hence, KEAVL can provide general vocabulary preparation for the international students to be able to read research materials from different engineering sub-fields and to enrich their engineering vocabulary even before they claim a specific major. Indeed, according to the coverage measurement results, KEAVL covers a higher percentage of lemmas in CKEAT in comparison to KAVL. Hence, it can be a good addition to the existing lists for engineering students.

The measurements of the coverage (done by the *KWordList*) of CKEAT by KAVL show a result of 22% that is lower than the KEAVL coverage while the coverage of advanced KAEVL is

lower. Nevertheless, it is important to keep in mind that KAVL has an intersection with the KLL words that were excluded from the advanced KAEVL.

5.3 The role of KEAVL among other vocabulary lists for engineering students

As mentioned in Chapter 1, there are List of Physics Vocabulary for first-year engineering students (Park, 2017), and technical vocabulary lists for automobile (Kim, 2015), electrical and electronics (Park, 2017), mechanical (Lee, 2014) and textile engineering (Nergui, 2015).

Physics is an obligatory class for all first-year engineering students regardless of the sub-field they will major in. Therefore, physics vocabulary is very important for engineering college freshmen. However, it was decided to check how often the physics vocabulary is used beyond physics classes, particularly across CKEAT. As a result, 11.4% of the corpus was covered by the vocabulary from the PVL (see Table 5.1). This percentage was calculated with the list as it is. Hence, it should not be summed with other lists coverage in the corpus because they have some overlap with PVL.

Table 5.1

Comparison of the PVL and KEAVL coverage in CKEAT

List	Tokens covered	% of tokens covered	Lemmas covered	% of lemmas covered
PVL	148,836	11.4%	485	1.6%
Inter KEAVL	321,674	25.6%	830	2.7%
Adv KEAVL	158,700	12.1%	512	1.7%

Among the words in PVL, only 116 occurred evenly across the corpus. Among these words there are *peykthe* 'vector', *taylyu* 'convection current', *kansep* 'interference', *songsin*

'transmission', *kikwu* 'apparatus, machine'. Therefore, even PVL words occur in CKEAT, not all of them are typical for the majority of the engineering sub-fields.

Nevertheless, the coverage of the conjoint KEAVL and PVL in CKEAT was also measured. It turned out that two lists cover 36% of all words in the corpus, which is relatively high. Thus, by learning both lists, students can improve their understanding of the engineering academic texts from different sub-fields. To illustrate the coverage of both lists together, a random passage (Passage 4) from the corpus was selected. The words that are covered by the KEAVL are bold-faced, and the words that are covered by the PVL are underlined. Out of 91 Korean words (compounded nouns were counted as two nouns), nine words are exclusively presented in PVL, seventeen words are exclusively presented in KEAVL, and four words are presented in both of them. Therefore, the coverage of the following passage by these lists together is 33%.

Passage 4

"일반적으로 염료 감응 태양 전지는 두 개의 **전극**(photo electrode 와 counter electrode), **반도체** 나노 **입자** (주로 이산화티타늄), 염료 그리고 **액체** 전해질로 **구성**되어 있고, 이중 **액체** 전해질 **구성 요소** 부분이 **소자**의 **장기** 안정성과 아주 밀접하게 관련되어 있다. 요오드가 포함된 용액 **상태**의 휘발성 전해액은 에너지 **변환 효율 측면**에서는 우수한 장점을 가지고 있지만, 사용기간 동안 전해액이 누출 또는 휘발되게 되면 **소자**의 안정성에 치명적인 문제를 일으킬 수 있다는 단점도 함께 가지고 있다. 특히 전해액의 요오드 **성분**은 장시간 구동 시 염료 분자의 **화학**적 분해를 **유발**할 수 있고, 소량의 산소와 수분과의 **작용** 으로 금속 **성분**의 module grid 를 심각하게 **파괴**하기도 한다. "

'In general, dye-sensitized solar cells consist of two electrodes (a photoelectrode and a counter electrode), semiconductor nanoparticles (mostly titanium dioxide), dye, and liquid electrolytes, with a dual liquid electrolyte component closely related to the long-term stability of the device. The volatile electrolyte solution containing iodine has an excellent advantage in terms of energy conversion efficiency but also has the disadvantage that if the electrolyte leaks or

evaporates during the use period, it may cause a fatal problem in the stability of the device. In particular, the iodine component of the electrolyte may cause chemical decomposition of the dye molecule when it is operated for a long time, and may seriously destroy the module grid of the metal component by the action of a small amount of oxygen and water.'

The technical vocabulary lists for the engineering sub-fields were made by exclusion of KLL and KAVL lemmas from the corpora. Nevertheless, as it was shown in Section 4.2, KEAVL has better coverage of CKEAT in comparison to KAVL. Hence, it can be a good alternative to KAVL for engineering technical vocabulary list creation. This is because, it can filter out the vocabulary that is typical not just for all academic texts, but particularly for the academic texts in engineering. Thus, it can provide more precise results on a technical vocabulary of a particular field.

The technical vocabulary list for mechanical engineering, for instance, has 137 lemmas overlapping with KEAVL. Thus, knowing words from KEAVL before starting technical vocabulary education may benefit students and provide them with basic engineering vocabulary.

To sum up, KEAVL could become an alternative to KAVL for future engineering students and also be very helpful for engineering technical vocabulary lists creation. Moreover, KEAVL may be a significant addition to the PVL for students to learn at the beginning of their education. Both lists, together with basic and technical vocabulary, can potentially increase the number of known words in a subfield's texts up to 95%.

The measurements of the coverage of the Physics Vocabulary List for the engineering students in CKEAT also showed a low percentage of coverage. Therefore, it is possible to say that KEAVL could be a good vocabulary basis for the international students who have started their degree in engineering, and are having classes related to different sub-fields of engineering. The

results of the PVL and KEAVL joint-coverage in the corpus show a number of 36% of CKEAT. Thus, even PVL alone has a relatively low coverage in the corpus, it can be a good start for engineering students, and KEAVL can be the next stage of their vocabulary learning process. This is to say, since physics classes are obligatory for first-year engineering students, first, their vocabulary focus should lie in the PVL. Nevertheless, for the following stage of their education, when they are attending more specific classes in different fields of engineering, their vocabulary focus could be switched to KEAVL. Having both vocabulary lists covered, students may be able to understand almost 50% of textbooks and research articles, while general vocabulary covers the rest.

Then, after a decision to focus on a specific engineering sub-field (i.e., mechanical engineering), a student may focus on the technical engineering vocabulary lists. For instance, the Mechanical Engineering Vocabulary List, Textile Engineering Vocabulary List, Electronics and Electrical Engineering Vocabulary list. Therefore, KEAVL may be a good "transition list" from a vocabulary of physics to the more narrow-focused engineering vocabulary.

5.4   The program

The equally important outcome of the current study is the *KWordList* script that can create word lists for the Korean Language. The current study is not the first to use programming in the word list creation process (Choe, 2016; Lei & Liu, 2016). Nevertheless, unlike the previous studies, the script developed in this study is fully reusable. It was written to create not only KEAVL but to be used by other researchers in the future. The script is fully available on *GitHub*, the online platform to share code (https://github.com/HelgaKr/KWordList).

The *KWordList* is able to compile lists based on the corpus and take into account such criteria as range, frequency, and variability. It also can search for the collocations of the words in

71

the list (bi-grams and tri-grams). Additionally, it can process large corpora, a corpus with many sub-corpora, and even some small texts in the .txt format and UTF-8 encoding. For this study, a corpus with sub-corpora was processed, however, another type of corpus may be set at the beginning with the program.

As well as a corpus type, a range and frequency parameters can be set individually depending on the researchers' objectives. For the collocations, frequency of their occurrence and a maximum number of them can also be set. Moreover, *KWordList* can count the dispersion of the word distribution in the case of the sub-corpora. Nevertheless, it does not work like a filter for the words, and the final decision whether to include the word with low variability or not is up to the researcher.

Finally, the *KWordList* also can calculate the size of the corpus by tokens and lemmas, coverage of KLL and KAVL, and coverage of the built list in the uploaded corpus both among tokens and lemmas if it is required. As mentioned in Chapter 2, a modified lexeme approach is used in this study. Hence, the SK nouns in predicative forms are counted together with the SK noun forms.

As a result, the program can process a corpus containing 1 million tokens along with sub-corpora to give an output table with the word, range, frequency, and collocations in approximately 20-30 sec. If the program also calculates coverage measurements, the processing time of 1 million tokens may increase up to 1 minute (though, the processing time may vary depending on the computer processing power[17]). This short processing time may significantly speed up the process of building word lists for the Korean language and let researchers concentrate more on other important objectives. Furthermore, the adjustable list building settings allows comparing the

---

[17] For the following study the computer with AMD Ryzen™ 7 3700X was used. It has eight CPU cores, sixteen threads and 3.6GHz base clock.

results with different parameters and choose the settings that provide the shortest list and the best coverage.

The flexibility of the *KWordList* makes it useful not only to KAP researchers but also for teachers and students. Teachers can use this program to create small word lists for reading and listening in-class activities. At the same time, students may use this program to create word lists for their own purposes, i.e., lists of words for extracurricular readings.

Moreover, an important feature of the *KWordList* is that it can filter out different levels of vocabulary and not all of it. Thus, it can build wordlists for students of different levels of proficiency in Korean, which also expands the scope of the program. In order to provide a guideline for future users of the program, the detailed description and recommendations were added to the *GitHub* with the script itself (https://github.com/HelgaKr/KWordList/blob/master/README.md).

Nevertheless, it is important to mention the limitation that the program has. When it excludes the words from the KLL or KAVL to build a list, it does not take into account the meaning of the word in a particular case. Therefore, a situation where a word in the researcher corpus has a different meaning from the word that was added to KLL and KAVL lists is possible. Hence, it is the task of the researcher to verify if the meaning of the excluded word coincides with the meaning of the word in the corpus. The possible solutions to make this task easier and less time-consuming may be advised here. Another beneficial feature of the KLL and KAVL is that both contain the number of the word meanings as it is stated in the Standard Korean Language Dictionary. The *KWordList*, in turn, has an optional function to save all the excluded words from each list in separate .txt files. Consequently, after the first trial of compiling the word list, researchers can look through the words that were excluded, find if some of these words are presented by a certain meaning in the KLL or KAVL, and check if the excluded word has the same meaning. In the case

of this study, only a few instances were found, however, it is still very important to doublecheck this information so not to exclude the word important for the list unintentionally.

Moreover, it is also important to check the context in which words from the built list occurred in the corpus. Again, it is necessary to prevent the cases when the same word with different meanings is included in the list as one entry. Therefore, for the lists developed on the basis of a large corpus, it is essential to check each entry carefully. Mainly, it is the case of the SK nouns that have the same written representation but different Chinese origin words (example may be found in Section 4.1).

Another important aspect to keep in mind while developing lists with the program is the possible cases of the different spelling of borrowed words. In the case of this study, two different spelling versions of the English word "algorithm" were identified: *alkolicum* and *alkolitum*. As mentioned in Chapter 4.1, both occurred frequently, and for the list, the one entry for both versions was done, and the frequency of their occurrence was summed. Unfortunately, in the course of the current study, a possible computational solution for this problem was not proposed. Thus, although in Korean such cases are very rare, it is recommended to check if foreign language borrowings that occurred in the corpus have two or more spelling variants in Korean.

To conclude, while the program can process the entire corpus in minutes, it still has its limitations, and thus inspection of the results by the researcher is still necessary. The *KWordList* is planned to be improved considering its limitations and to be reworked as an online version in the future.

5.5   Conclusion

The current study presented the first word list of Korean academic vocabulary for engineering fields, which is divided into two levels of Korean language proficiency (intermediate

and advanced). Information regarding word list lemma collocations was also collected. Additionally, the English translation and origin words were added to the final versions of the lists.

The coverage measurements showed that KEAVL covers a substantial part of the tokens in CKEAT. Moreover, the KEAVL has better coverage than the KAVL in the corpus, which makes it a good alternative to KAVL for engineering students. KEAVL provides them with an academic vocabulary set that is widely used across all main engineering fields. Moreover, it was also shown that KEAVL, in combination with KLL and PVL, boosts the coverage of words in the corpus, thus it is more likely to help learners understand the majority of the lemmas in academic engineering texts. Furthermore, in combination with technical vocabulary lists for their sub-fields, it may give students a nearly complete reference source for the reading and listening comprehension for engineering academic vocabulary in Korean.

Additionally, the present study has developed KWordList, a tool that automatically creates Korean word lists based on the corpora data provided by other researchers or teachers. It has many different options that may be customized according to the purposes of the study. Therefore, it may contribute to further development of the KAP field and the word list creation process in particular.

References

ABBYY Production LLC. (2017). ABBYY FineReader.

An, J., & Seon, U. (2013). *Korean grammar in use. Advanced*. Darakwon.

Anderson, J. (1980). The lexical difficulties of English medical discourse for Egyptian students. *English for Specific Purposes*, *37*(4).

Baker, M. (1988). Sub-technical Vocabulary and the ESP Teacher: An Analysis of Some Rhetorical Items in Medical Journal Articles. *Reading in a Foreign Language*, *4*(2), 91–105.

Bauer, L., & Nation, I. S. P. (1993). Word Families. *International Journal of Lexicography*, *6*(4), 253–279. https://doi.org/10.1093/ijl/6.4.253

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. London: Longman.

Bird, S., & Loper, E. (2009). *Natural Language Processing with Python. Natural Language Processing with Python*. O'Reilly Media Inc.

Brezina, V. (2018). *Statistics in corpus linguistics : a practical guide*. Cambridge: Cambridge University Press.

Brezina, V., & Gablasova, D. (2015). Is There a Core General Vocabulary? Introducing the New General Service List. *Applied Linguistics*, *36*(1), 1–22.

Browne, C. (2014). A new general service list: The better mousetrap we've been looking for. *Vocabulary Learning and Instuction*, *3*(2), 1–10.

Browne, C., Culligan, B., & Phillips, J. (2013). New General Service List.

Campion, M., & Elley, W. (1971). *An academic vocabulary list*. Wellington: New Zealand Council for Educational Research.

Carter, R. (1998). *Vocabulary. Applied Linguisitc Perspectives.* Routledge.

https://doi.org/10.4324/9780203070239

Choe, E. (2016). *Hangmun mokjeong hangugeo hakseupjareul wihan jayeongwahakgyeyeol jeongong giboneohwi seonjeong [A Study on Selection of Fundamental Vocabularies of Natural Science for Learners of Korean for Academic Purposes].* Ewha Woman's University, Graduate School of Education.

Chung, T. M., & Nation, I. S. P. (2004). Identifying technical vocabulary. *System*, *32*(2), 251–263. https://doi.org/10.1016/J.SYSTEM.2003.11.008

Cohen, A., Glasman, H., Rosenbaum-Cohen, P. R., Ferrara, J., & Fine, J. (1979). Reading English for Specialized Purposes: Discourse Analysis and the Use of Student Informants. *TESOL Quarterly*, *13*(4), 551. https://doi.org/10.2307/3586449

Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, *34*(2), 213.

Coxhead, A. (2002). The Academic Word List: A Corpus-based Word List for Academic Purposes. *Teaching and Learning by Doing Corpus Analysis*, 72–89.

Coxhead, A. (2011). The Academic Word List 10 Years On: Research and Teaching Implications. *TESOL Quarterly*, *45*(2), 355–362.

Coxhead, A., & Nation, I. S. P. (2001). The specialized vocabulary of English for academic purposes. In J. Flowerdew & M. Peacock (Eds.), *Research perspectives on English for academic purposes* (pp. 252–267). Cambridge: Cambridge University Press.

Engels, L. K. (1968). The fallacy of word-counts. *IRAL - International Review of Applied Linguistics in Language Teaching*, *6*(1–4), 213–232. https://doi.org/10.1515/iral.1968.6.1-4.213

Farrell, P. (1990). Vocabulary in ESP: A Lexical Analysis of the English of Electronics and a Study of Semi-Technical Vocabulary. *CLCS Occasional Paper*, *25*.

Gang, H. H. (2012). *Hangugeo gyoyung eohwi naeyong gaebal (1 dangye) [Research on Korean*

Language Education Vocabulary Content Development (Stage 1)]. Seoul: National Institute of Korean Language.

Gardner, D., & Davies, M. (2013). A New Academic Vocabulary List. *Applied Linguistics*, *35*(3), 305–327. https://doi.org/10.1093/applin/amt015

Gries, S. T. (2008). Dispersions and adjusted frequencies in corpora. *International Journal of Corpus Linguistics*, *13*(4), 403–437. https://doi.org/10.1075/ijcl.13.4.02gri

Han, C. H., & Palmer, M. (2005). A morphological tagger for Korean: Statistical tagging combined with corpus-based morphological rule application. *Machine Translation*, *18*(4), 275–297. https://doi.org/10.1007/s10590-004-7693-4

Han, S. (2010). Hangmunmokjeong hangugeo gyoyukgwajeong seolgyeui silje (Daehang ipang jeon hangugeo gyoyukgwajeongeul jungsimeuro) [Course Design for Teaching Korean for Academic Purpose]. *Journal of Korean Language Education*, *21*(1), 225–248.

Heatley, A., & Nation, I. S. P. (1996). Range [Computer software]. Wellington: Victoria University of Wellington. Retrieved from https://www.victoria.ac.nz/lals/about/staff/paul-nation

Hyland, K. (2001). Bringing in the Reader: Addressee Features in Academic Articles. *Written Communication*, *18*(4), 549–574. https://doi.org/10.1177/0741088301018004005

Hyland, K., & Tse, P. (2007). Is There an "Academic Vocabulary"? *TESOL Quarterly*, *41*(2), 235–253.

Im, C. (2002). Chogeup hangugeo gyoyug yong eohwi seonjeong yeongu [A Study on the Vocabulary Selection for Beginner Korean Education]. *Korean Language Education*, *30*.

Im, J. (1991). Gugeoui gichoeohwie daehan yeongu [Study on Basic Vocabulary of Korean Language]. *Korean Language Education*, *23*(1), 87–131.

Jeong, J. (2018). Hangugeoneungnyeoksiheomui sagodogueo sayong yangsang yeongu

(TOPIKkwa EJUui sagodogueo gyeryangjeong bunseong bigyoreul jungsimeuro) [A Study on Academic words usage in Korean Proficiency Test (quantitative comparative analysis of academic vocabulary usage in TOPIK and EJU)]. *Korean Language Education*, *160*, 267–301.

Jo, H. Y. (2000). *Hangugeo eohwi gyoyung yeongu [A study of Korean Language Education].* Bagijeong.

Jo, N. (2002). *Hyeondae gugeo sayong bindo josa - hangugeo hanseubyong eohwi seongyeongeul wihan gicho josa [Study on a Korean language word usage frequency - for a study on vocabulary selection for Korean learners].* Seoul: National Institute of Korean Language.

Jo, N. (2003). *Hangugeo hakseubyong eohwi seonjeong gyeolgwa bogoseo [Report on the results of vocabulary selection for learners of Korean].* Seoul: National Institute of Korean Language.

KAIST Semantic Web Research Center. (n.d.). HanNanum. Retrieved from http://semanticweb.kaist.ac.kr/home/index.php/HanNanum

Kang, J. (2007). *Hangmun mokjeong hangugeo hakseupjareul wihan sanggyeonggyeyeol hangmunyong giboneohwi seonjeong [Selection of Basic Academic Vocabulary of Business and Economics for Learners of Korean for Academic Purposes].* Ewha Woman's University, Graduate School of Education.

Kim, E., Kweon, S., & Kim, J. (2017). Korean engineering students' perceptions of English-medium instruction (EMI) and L1 use in EMI classes. *Journal of Multilingual and Multicultural Development*, *38*(2), 130–145. https://doi.org/10.1080/01434632.2016.1177061

Kim, I. (2003). Hangmun mokjeogeul wihan hangugeo yogu bunseong mitgyosuyomong gaebalgimingyu [Needs analysis and syllabus development of Korean for academic purposes]. *Journal of Korean Language Education*, *14*(3).

Kim, N. (2005). *Hangmun mokjeong hangugeo eohwi gyoyung bangan yeongu: inmungyeyeol hakseupjareul jungsimeuro [A study on a Korean Academic Vocabulary Teaching Method: Focused on the students in Humanities Studies]*. Kyung Hee University.

Kim, S. Y. (2015). *Hangmun mokjeong hangugeo hakseupjareul wihan naeyong jungsim hangugeo eohwi gyoyung yeongu : jadongchagonghang jeongongeul jungsimeuro [A study of content-based Korean vocabulary education for the learners on academic purpose : focusing on automobile engineering]*. Kyung Hee Cyber University.

Korean National Institute for International Education. (2018). *2018 oegugin yuhaksaeng hanguksaenghwal jeongchakjiwon ususaryejip [A Case Study on Support of International Students in Korea]*. Seoul.

Kudo, T. (2013). Mecab-ko. Retrieved from https://bitbucket.org/eunjeon/mecab-ko/src/master/AUTHORS

Lam, J. K. (2000). *A study of semi-technical vocabulary in computer science texts with special reference to ESP teaching and lexicography*. University of Exeter.

Laurence, A. (2018). AntConc. Tokyo, Japan: Waseda University. Retrieved from http://www.laurenceanthony.net/software.html

Lee, D., Yeon, J., Hwang, I., & Lee, S. (n.d.). Kkokkoma: hangeul hyeongtaeso bunseokgi [KKMA: Kool Korean Morpheme Analyzer]. Retrieved from http://kkma.snu.ac.kr/

Lee, H. (2004). Hangmun mokjeong hangugeo gyogwagwajeong seolgye yeongu [Research on Course Design of Korean for Academic Purposes]. *Journal of Korean Language Education*, *15*(1).

Lee, H. S. (2014). *Hangmunmokjeong hangugeo hakseupjareul wihan jeongong giboneohwi seonjeong min bunseong : gigyegonghang jeongong gibon gyojae bunseogeul jungsimeuro [A study on the selection of basic technical words for Korean language learners majoring in*

mechanical engineering : focusing on basic textbooks for mechanical engineering majors]. Yonsei University.

Lee, S. (2018). Hangmun mokjeong hangugeo gyoyung chegyehwareul wihan modyulsing gyoyukgwajeong yeongu. [A Study on the Modular Curriculum for the Systematic Korean Language Education for Academic Purpose]. *Korean Language Education and Pedagogy (Pai Chai University)*, *8*.

Lei, L., & Liu, D. (2016). A new medical academic word list: A corpus-based study with enhanced methodology. *Journal of English for Academic Purposes*, *22*, 42–53.

Liu, J., & Han, L. (2015). A corpus-based environmental academic word list building and its validity test. *English for Specific Purposes*, *39*, 1–11. https://doi.org/10.1016/j.esp.2015.03.001

Looney, D., & Lusin, N. (2018). *Enrollments in Languages Other Than English in United States Institutions of Higher Education, Summer 2016 and Fall 2016: Preliminary Report*.

Lynn, R. W. (1973). Preparing word lists: a suggested method. *RELC Journal*, *11*(2), 25–32.

Martínez, I. A., Beck, S. C., & Panza, C. B. (2009). Academic vocabulary in agriculture research articles: A corpus-based study. *English for Specific Purposes*, *28*(3), 183–198.

Mudraya, O. (2006). Engineering English: A lexical frequency instructional model. *English for Specific Purposes*, *25*(2), 235–256.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

Nation, I. S. P. (2013). *Learning Vocabulary in Another Language* (Second Edi). Cambridge: Cambridge University Press.

Nation, I. S. P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In M. McCarthy & N. Schmitt (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6–19). Cambridge University Press.

Nergui, M. (2015). *Seomyugonghang jeongong hangugeo hakseupjareul wihan jeongong giboneohwi seonjeong yeongu [A study on the Selection of Basic Technical Words for Korean Language Learners Majoring in Textile Engineering]*. Yonsei University.

Oh, K. A., & Kim, Y. M. (2012). The Effectiveness of Code-switching and Identity Reconstruction. *English Language and Literature Teaching*, *18*(4), 115–141.

Park, E. (2006). *ET-house Neungnyul hanyeong sajeon [ET-house Korean-English dictionary]*. Neungnyul sajeon pyeonchan yeonguso.

Park, E. L., & Cho, S. (2014). KoNLPy: Korean natural language processing in Python. In *Proceedings of the 26th Annual Conference on Human & Cognitive Language Technology*. Chuncheon.

Park, J. (2018). *Gonghang jeongong hangugeo hakseupjareul wihan byeongjon eoneo gyosubeom giban eohwigyoyung yeongu. Mullihang giboneohwireul jungsimeuro [A study on the vocabulary education based on the adjunct language instruction for Korean language learners majoring in engineering (the case of physics vocabulary)]*. Sookmyung Women's University.

Park, L. (2018). KoNLPy Documentation, Release 0.5.2.

Park, S. H. (2017). *Jeongijeonjagonghang hangmun mokjeong hangugeo hakseupjareul wihan jeongong giboneohwi seonjeong yeongu [A Study on The Selection of Basic Technical Words for Korean Language Learners Majoring in Electrical and Electronics Engineering]*. Ulsan University.

Praninskas, J. (1972). *American University Word List*. London: Longman.

Richards, J. C. (1974). Word Lists: Problems and Prospects. *RELC Journal*, *5*(2), 69–84. https://doi.org/10.1177/003368827400500207

Ryu, W. H. (n.d.). Open Korean Text Processor (Okt). Retrieved from https://github.com/open-

korean-text/open-korean-text

Santos, T. (1988). Professors' Reactions to the Academic Writing of Nonnative-Speaking Students. *TESOL Quarterly*, *22*(1), 69. https://doi.org/10.2307/3587062

Shin, M. (2003). Jijeoksagocheorieoui teukseonggwa geu gugeogyoyukjeong piryoseonge daehan gocha [A Study on Academic Vocabulary's Character and Necessity of Academic Vocabulary in Korean Education]. *The Korean Language and Literature*, *81*, 25–48.

Shin, M. (2004). *Gugeo sagodogueo gyoyung yeongu [A study on Korean academic vocabulary education]*. Seoul National University.

Shineware. (n.d.). Komoran. Retrieved from https://www.shineware.co.kr/products/komoran/

Sinclair, J. (1991). *Corpus, concordance, collocation.* Oxford University Press.

Sinclair, J. M. (Ed.). (2004). *How to Use Corpora in Language Teaching* (Vol. 12). Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/scl.12

Sohn, H. (1999). *The Korean Language*. Cambridge: Cambridge University Press.

Sutarsyah, C., Nation, P., & Kennedy, G. (1994). How Useful Is Eap Vocabulary for Esp? A Corpus-Based Case Study. *RELC Journal*, *25*(2), 34–50. https://doi.org/10.1177/003368829402500203

Tajino, A., Dalsky, D., & Yosuke, S. (2009). Academic vocabulary reconsidered: an EAP curriculum-design perspective. *Iranian Journal of TEFLL*, *1*(4), 3–21. Retrieved from https://www.sid.ir/en/journal/ViewPaper.aspx?id=201458

Valipouri, L., & Nassaji, H. (2013). A corpus-based study of academic vocabulary in chemistry research articles. *Journal of English for Academic Purposes*, *12*(4), 248–263.

Wang, J., Liang, S., & Ge, G. (2008). Establishment of a Medical Academic Word List. *English for Specific Purposes*, *27*(4), 442–458. https://doi.org/10.1016/J.ESP.2008.05.003

Ward, J. (2009). A basic engineering English word list for less proficient foundation engineering

undergraduates. *English for Specific Purposes*, *28*(3), 170–182.

West, M. (1953). *A general service list of English words*. Longmans, Green.

Won, Y. (2016). Common European Framework of Reference for Language (CEFR) and Test of Proficiency in Korean (TOPIK). *International Journal of Area Studies*, *11*(1), 39–58. https://doi.org/10.1515/ijas-2016-0003

Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge University Press.

Xue, G.-Y., & Nation, I. S. P. (1994). A University Word List. *Language Learning and Communication/Zhongying Yuwen Jiaoxue Xue*, *3*(2), 215–229.

Yang, M.-N. (2015). A nursing academic word list. *English for Specific Purposes*, *37*, 27–38.

Yeo, S. (2018). *Hangmun mokjeong hangugeo hakseupjareul wihan jeongchioegyohang jeongong giboneohwi seonjeong [A Selection of Political Diplomacy Technical Basic Words for Korean Learners with Academic Purposes Authors]*. Ewha Woman's University, Graduate School of Education.

Appendix A.

Selected Korean engineering academic journals.

Sub-field of Engineering and letter code	Journal Title (English) and number code	Journal Title (Korean)	Years	Issuing Organization	Place on *DBpia*
General (GE)	Journal of Korea Academia-Industrial cooperation Society (1)	한국산학기술학회논문지	2005-now	Korea Academia-Industrial cooperation Society	TOP 1
	Fire Science and Engineering (2)	한국화재소방학회논문지	1987-now	Korean Institute of Fire Science and Engineering	TOP 5
Mechanical (ME)	Journal of the Korean Society of Mechanical Technology (3)	한국기계기술학회지	2007-now	Korean Society of Mechanical Technology	N/A
	Auto Journal (4)	오토저널	1979-now	The Korean Society of Automotive Engineers	TOP 5
Aerospace (AE)	Aerospace Engineering and Technology (5)	항공우주기술	2002-2014	Korea Aerospace Research Institute	TOP 5
	Journal of the Korean Society for Aeronautical & Space Sciences (6)	한국항공우주학회지	1973-now	The Korean Society for Aeronautical & Space Sciences	
Computer (CS)	Journal of The Korea Society of Computer and Information (7)	한국컴퓨터정보학회논문지	2004-now	Computer and Information	TOP 5
	Journal of Digital Contents Society (8)	디지털콘텐츠학회논문지	2004-now	Digital Contents Society	
Chemical/Biological (CB)	KSBB Journal (9)	N/A	1987-now	Korean Society of Biotechnology and Bioengineering	TOP 5
	Polymer Science and Technology (10)	고분자 과학과 기술	1990-now	The Polymer Society of Korea	
Electrical/Control (EL)	Journal of Korean Institute of Electrical Engineers (11)	전기의세계	1948-now	The Korean Institute of Electrical Engineers	TOP 5
	The transactions of The Korean Institute of Electrical Engineers (12)	전기학회논문지	1982-now	The Korean Institute of Electrical Engineers	
Civil/Environmental (CE)	Journal of Korean Society of Environmental Engineers (13)	대한환경공학회지	1979-now	Korean Society of Environmental Engineers	TOP 5

	Magazine of the Korea Concrete Institute (14)	콘크리트학회지	1989-now	Korea Concrete Institute	
Architectural (AR)	Journal of the Architectural Institute of Korea (Structure & construction) (15)	대한건축학회논문집 구조계	1998-	Architectural Institute of Korea	TOP 5
	Journal of the Korean Institute of Rural Architecture (16)	한국농촌건축학회논문집	1999-	Korean Institute of Rural Architecture	N/A
Industrial (IE)	Journal of the Society of Korea Industrial and Systems Engineering (17)	산업경영시스템학회지	2000	The Society of Korea Industrial and Systems Engineering	N/A
	Journal of Korean Institute of Industrial Engineers (18)	대한산업공학회지	1975-now	Korean Institute of Industrial Engineers	TOP 5
Marine/Ocean (MO)	Journal of the Society of Naval Architects of Korea (19)	대한조선학회논문집	1991	The Society of Naval Architects of Korea	TOP 1
	Journal of Ocean Engineering and Technology (20)	한국해양공학회지	1987	The Korean Society of Ocean Engineers	TOP 5
Resources/ Materials Science (RM)	Journal of the Korean Society of Mineral and Energy Resources Engineers (21)	한국자원공학회지	2002-now	The Korean Society of Mineral and Energy Resources Engineers	N/A
	Journal of Korean Institute of Gas (22)	한국가스학회지	1997	Korean Institute of Gas	TOP 5
Electronics/ Communications (EC)	Journal of the Korea Institute of Information and Communication Engineering (23)	한국정보통신학회논문지	1997-now	Korea Institute of Information and Communication Engineering	TOP 1
	The Magazine of the IEIE (24)	전자공학회지	1963-now	The Institute of Electronics and Information Engineers	TOP 5

Appendix B.

Research articles selected for the corpus.

Code	Journal	Title in English	Year of publication
GE001	1	Mediating Effect of Career Preparation Behavior on Nursing Professionalism Related Factors of Nursing Students	2019
GE002	1	Implementation of abnormal behavior detection system based packet analysis for industrial control system security	2018
GE003	1	Analysis of Fish Blocking Effect using Illuminance Difference	2017
GE004	1	Assessment of Elderly's Isokinetic Muscle Function, Flexibility and Balance in a Region of Seoul	2016
GE005	1	Thermal Design of a Cooling Coil for Building Air Conditioning	2015
GE006	1	Study of the Factors Related to the Labor Market Transition of Job Injured Workers	2014
GE007	1	Disk Sector Antenna fed by CPW for UWB Communications	2009
GE008	1	The Modification of Serial Cadastral Map and Its Applications to Notification of Topographical Maps	2011
GE009	1	A Study on the Formation Plan of Green Cluster by Sectoral Type for the Enhancement of Regional Competitiveness in Green Industry	2013
GE010	1	A case study on troubles analysis and diagnoses of passenger car's engine based on OBD	2006
GE011	1	Identification of Thermal Flow Boundary Conditions for Three-way Catalytic Converter Using Optimization Techniques	2010
GE012	1	A Study on Evaluation System of River Levee Safety Map to Improve Maintenance Efficiency and Disaster Responsiveness	2018
GE013	1	The Impact of Reviewer Professionalism and Expectation on the Intention to Watch: A Comparison between Commercial and Art Films	2012
GE014	1	Investigation of the BSR Noise characteristics in Seat Cushion-frame with respect to Vibration Durability Test Using Multi-simulator	2014
GE015	1	Factors Associated with Patient Safety Care Activity among Nurses in Small-Medium Sized General Hospitals	2019
GE016	2	A Study on Fire Features of Double-Skin Facade Structure by Using Fire Simulation (FDS)	2014
GE017	2	An Experimental Study on Structural Behaviour of Asymmetric H Beam Slimfloor under Load Condition in Fire	2011
GE018	2	Application Study of Design Fire Curves for Liquid Pool Fires in a Compartment	2017
GE019	2	Analysis of the Situation of the Volunteer Fire Brigade in Japan	2011
GE020	2	Fire Alarm Sound Transmission in Apartment Units	2018
GE021	2	A Study on the Optimum Disaster Mitigation Activity to Establish the Early Counter System on the Industrial Accidents	2010
GE022	2	Changes in Fire Characteristics according to the Distance Between the Fire Source and Sidewall in a Reduced-Scale Compartment	2019
GE023	2	Toxicity Evaluation of the Combustion Products from Synthetic Wood as Internal Finish	2016
GE024	2	Improvement Proposal for the Fire Suppression Systems of Open Parking Lots	2016
GE025	2	A Study on the Electrical Fire Risk of Terminal Block Due to Single and Composite Cause	2015
GE026	2	A Study on the Effects of Various Disk Shape of Hydrant on the Pressure Drop	2013
GE027	2	A Study on Improving Role of Firefighting Service Toward an Aging Society	2012

Code	Journal	Title in English	Year of publication
GE028	2	A Study on the Reliability Analysis and Risk Assessment of Liquefied Natural Gas Supply Utilities	2003
GE029	2	A Study on the Application of the Electronic Approval System for the Fire Officer	2006
GE030	2	Development of Road-Map on Evaluation of Fire Resistance Performance	2009
ME004	3	Dynamic perfomance model analysis for six-wheels drive vehicles	2019
ME005	3	Numerical Analysis on Heat Transfer Characteristics of Ventilated Disc Brake with Holes	2018
ME006	3	A Study on Heat Transfer Characteristics according to Block Shapes in a Horizontal PCB Channel	2017
ME007	3	A Study on Velocity Distribution Characteristics of Exit Section and Downstream Straight Length in Square Curved Duct According to Angle Shape	2016
ME008	3	Study of Strawberry High-Density Cultivating System using Movable Bed	2016
ME009	3	Nonlinear Control of Pneumatic Cylinder Actuators with Random Friction Nature	2015
ME010	3	A Numerical Study on the Heat Transfer of a spiral type EGR Cooler for Diesel Engine	2014
ME011	3	Environment Noise Influence Evaluation for remove of Military Training Center	2013
ME012	3	Development of Pressure Measurement Monitoring System and Its Performance Evaluation in Ullage Space of Cargo Tank	2010
ME013	3	Study on characteristics of cold temperature separation for cooling of charging air in a vortex tube	2012
ME014	3	A study on motion analysis of 1-D.O.F MEMS manipulator	2009
ME015	3	Evaluation of stress Intensity Factor using Boundary Element Alternating Method	2008
ME016	3	A Study on Engine Performance of the Ignition Spark Timing Conversion for LPG/Gasoline Bi-fuel Vehicle	2011
ME017	3	Comparison of Shock Isolation Characteristics using Rubber Washer	2017
ME018	3	The Suggestion of Finite Element Modeling Method for Structural Design of Automotive Body	2019
ME019	3	Vehicle Compatibility in Car-to-Car Frontal Offset Crash	2001
ME020	4	Navigating of EU Environmental Regulation for Tire	2009
ME021	4	Prospect od DME Production Technology and Utilization	2003
ME022	4	Trends in Telematics Based or Networks for Vehicle	2005
ME023	4	The legislative Process of Special Act on Seoul Metropolitan Air Quality Improvement and Problems to be Solved	2004
ME024	4	Development of Common Rail Injection System for HSDI Diesel Engine for Passenger Car	2001
ME025	4	Study on the application of ergonomics to AVHS	2000
ME026	4	Development Trend of Car Seat Technology	2009
ME027	4	DE-NOx technology using plasma / catalyst complex system	2000
ME028	4	The Technology of Automative Suspension Systems	2009
ME029	4	Technology Trends of Airbag Control System (Smart Airbag Control Technology)	2000
ME030	4	Trends of LPG Automative	2002
ME031	4	A Study on Development of Competency-Based Curriculum with Automative-Repair Occupational Standards	2004
ME032	4	Atomization Characteristics in the CDI Engine	2000
ME033	4	Analysis of SUV Rollover due to Tire Characteristic Factors	2009
AE004	5	TVC Actuation Tests and Analyses for Real-Sized Kick Motor Assembly of KSLV-I	2007
AE005	5	Attitude SCAS Design for 40% Scaled Smart UAV	2007
AE006	5	Development of the SUAV Drive System - Design and Analysis	2008

Code	Journal	Title in English	Year of publication
AE007	5	Introduction and Application of Worst Case Analysis in Space Environment	2006
AE008	5	Worst Case Analysis for General Conventional Linear Regulator	2007
AE009	5	Study on the Aileron Reversal Characteristics of CUS Composite Aircraft Wings	2009
AE010	5	FDTD Analysis of the Mutual Coupling Between Closely Placed IFAs	2010
AE011	5	Transportation of KSLV-I FM unit 2 VEB Assembly	2011
AE012	5	The Result in Quality Management Activity of Propellant and Compressed Gases during the Operation of KSLV-I	2010
AE013	5	Attitude analysis induced by the disturbances on COMS using the received telemetries during normal mode	2012
AE014	5	Development of Operator Training System Using COMS Simulator for Provision Against Contingency Situation	2012
AE015	5	Development of Main Wing Structure of Long Endurance Electric Powered UAV	2013
AE016	5	Analyzing Season and Place of Mars Dust Storm	2013
AE017	5	COTS Based Air Data Recording System for SmartUAV	2010
AE018	5	Study on the Thruster Plume Behaviors using Preconditioned Scheme and DSMC Method	2009
AE019	6	Data Quality Analysis of Korean GPS Reference Stations Using Comprehensive Quality Check Algorithm	2013
AE020	6	Design and Analysis of Flexbeam in SNUF Blade Equipped with Active Trailing-Edge Flap for Helicopter Vibratory Load Reduction	2018
AE021	6	Unsteady Three-Dimensional Analysis of Transverse Fuel Injection into a Supersonic Crossflow using Detached Eddy Simulation Part I : Non-Reacting Flowfield	2009
AE022	6	Numerical Study about the Effect of Continuous Blowing On Aerodynamic Characteristics of NACA 0015 Airfoil	2006
AE023	6	A Study on the Deep Neural Network based Recognition Model for Space Debris Vision Tracking System	2017
AE024	6	System Development of SCSky CanSat With Smart Phone and Wide Swath Scan Camera Mechanism	2017
AE025	6	Autonomous Formation Flight Tests of Multiple UAVs	2010
AE026	6	A study on the material mechanical properties and the flexural wrinkling of foam-filled sandwich beams	2003
AE027	6	A study of structural response of pipes due to internal gaseous detonation of hydrogen - and hydrogen - air mixtures	2008
AE028	6	Power System Optimization for Electric Hybrid Unmanned Drone	2019
AE029	6	Characteristics and Key Parameters of Dual Bell Nozzles of the DLR, Germany	2015
AE030	6	Optimum design analysis of ICP(Inductively Coupled Plasma) torch for high enthalpy thermal plasma flow	2012
AE031	6	Analysis of Relations between Ice Accretion Shapes and Ambient Conditions by Employing Self-Organization Maps and Analysis of Variance	2011
AE032	6	Development and Validations of the Aerodynamic Analysis Program of Multi-Rotors by Using a Free-Wake Method	2007
AE033	6	Aerodynamic Shape Optimization of Helicopter Rotor Blades in Hover Using a Continuous Adjoint Method on Unstructured Meshes	2005
CS004	7	Transactions Ordering based Secure Concurrency Control Scheme	2005
CS005	7	Ultra-light Mutual Authentication Scheme based on Text Steganography Communication	2019
CS006	7	The design methods of Infrared Camera with Continuous zoom	2016
CS007	7	Improved User Anonymity Authentication Scheme using Smart Card for Traceability	2012
CS008	7	A Study on Automation about Painting the Letters to Road Surface	2018
CS009	7	A Handoff Improvement Method for AP Choose Guarantee Network Performance of Mobile Node in Wireless LAN Systems	2010

Code	Journal	Title in English	Year of publication
CS010	7	A study on the Necessity of the Death Penalty in the Information Society -Focused on the misjudgement cases	2014
CS011	7	Person Tracking by Detection of Mobile Robot using RGB-D Cameras	2017
CS012	7	Effective Automatic Foreground Motion Detection Using the Statistic Information of Background	2015
CS013	7	A method to compute the packet size and the way to transmit for the efficient VoIP using the MIL-STD-188-220C Radio	2008
CS014	7		2017
CS015	7	Design of High-Speed Parallel Multiplier with All Coefficients 1's of Primitive Polynomial over Finite Fields GF(2m)	2013
CS016	7	Type Classification of Korean Characters Considering Relative Type Size	2006
CS017	7	Fuzzy Theory and Bayesian Update-Based Traffic Prediction and Optimal Path Planning for Car Navigation System using Historical Driving Information	2009
CS018	7	Case Study on Global Software Education in Schools	2019
CS019	8	A study on factors Affecting Performance Assessment of Small and Medium Business R&D Project	2018
CS020	8	A Plan to Maximizing the Visual Immersion of 3D Media Art	2015
CS021	8	Analysis of long exposure noise and high sensitivity noise pattern of digital camera	2019
CS022	8	A Study on Methodology for Protection of Malicious Traffic in Groupware Network System	2007
CS023	8	Analyzing the market of corporate e-learning	2009
CS024	8	Design and Implementation of User Feedback Block Editor for Dynamic E-Book	2017
CS025	8	Sematic search of documents using ontology	2014
CS026	8	A Study for a Method of Designing of Security Domain Infrastructure and Its Efficiency Measuring	2010
CS027	8	An Architecture Analysis and Current Status of Games on Demand	2007
CS028	8	Multimedia Contents Dissemination using Mobile Communication and Opportunistic Networks	2013
CS029	8	A Study of Library and IT Services Improvement for enchancing the Domestic Reading Disabled People's Information Accessibility	2011
CS030	8	The impact of emotional storytelling on brand attitude- compare to humor advertising	2014
CS031	8	Following media development, a Study about the convergence of comics and multimedia	2012
CS032	8	Design and Development of e-Learning Contents for the NCS Vacational Core Competencies: Focusing on Interpersonal Competency	2016
CS033	8	A Study for Electronic Surveillance (RFID,CCTV,Electronic Resident Card) in Augmented Reality Environment	2006
CB004	9	Recent Advances in Tyrosinase Research as An Industrial Enzyme	2014
CB005	9	Antibacterial Activities of Fermented Sayuksan Ingredient Extracts for Multidrug-resistant Strainsfor Multidrug-resistant Strains	2014
CB006	9	Study of Effectiveness of Antimicrobial on Restraining Formation of Biofilms on the Surface of Aluminum	2015
CB007	9	Optimization of Cellulase Production from Paenibacillus jamilae BRC 15-1	2015
CB008	9	Live Cell Detection of Monoclonal Antibody Light and Heavy Chain mRNAs using Molecular Beacons	2016
CB009	9	Development and Evaluation of Gastro Retentive Floating Matrix Tablet Containing Valsartan Solid Dispersion	2016
CB010	9	Antioxidant Activity of the Extracts Derived from Korean Native Acer mono Max.	2017
CB011	9	Separation and Purification of Antimicrobial Substance from Syzygium aromaticum Merrill et Perry for Treatment of Microbial Vaginosis	2018

Code	Journal	Title in English	Year of publication
CB012	9	Antibacterial and Proteomic Effects of Legionella pneumophila JK-3 Exposed to Green Tea Catechin, Epigallocatechin Gallate (EGCG)	2019
CB013	9	Inhibition of Melanogenesis by Domestic Bamboo Leaves (Sasa coreana Nakai) Extract in B16F10 Melanoma Cells	2018
CB014	9	Nanomaterials with Peroxidase-like Activity for Biosensor Applications	2019
CB015	9	Identification and Characterization of Hemolytic Bacillus cereus Isolated from Commercial Ssam-jang	2017
CB016	9	Probiotic Properties and Inhibitory Activity of Lactic Acid Bacteria Isolated from Vaginal Microbiota of Korean Women against Gardnerella vaginalis and Candida albicans	2019
CB017	9	Inhibition of Melanogenesis by Domestic Bamboo Leaves (Sasa coreana Nakai) Extract in B16F10 Melanoma Cells	2018
CB018	9	Characterization and Inhibitory Activity of Lactobacillus plantarum MG989 and Lactobacillus fermentum MG901 Isolated from Vaginal Microbiota of Korean Women against Gardnerella vaginalis and Candida albicans	2016
CB019	10	Biosensing Platforms Based on Micro and Nanoscale Structures	2015
CB020	10	Colorless and Transparent Polyimide Films for Flexible Displays	2012
CB021	10	Self-Healing Polymers Based on Covalent Bond Formation	2014
CB022	10	Enhanced Power Conversion Efficiency in Organic Photovoltaic Devices with Plasmonic Effect	2015
CB023	10	Printed Electronics: High Resolution Printing Technology	2007
CB024	10	Solid=State Dye-Sensitized Solar Cells Based on p-Type Semiconductors	2010
CB025	10	Recent Patent Trends in Polymer Nanotechnology	2005
CB026	10	Enhancing Viability of Transplanted Islets by Using Biometric Polymer to Cure Type I Diabetes Mellitus	2011
CB027	10	Synthesis and Characterization of Materials Based on Ladder-Type Phenyless	2009
CB028	10	Directed Self-Assembly and Block Copolymers	2008
CB029	10	Technology Trends of Organic Light Field Effect Transistors	2006
CB030	10	Current Trends in Thermally Conductive Polymer Composites	2013
CB031	10	Development and Application of Nanoclay Polymer Nanocomposite	2007
CB032	10	Polymeric Materials for Lithium-Ion Batteries (Separators and Binders)	2013
CB033	10	Dendrimers as Delivery Carriers of Bioactive Agents	2004
EL004	11	Carbon Material Industry Trends and Directions for Energy Storage	2017
EL005	11	Overseas Deployment Case of Electric Vehicle Charging Demand Smart Monitoring System	2019
EL006	11	3D Digital Breast Cancer Diagnosis Technology Using Fusion Image	2017
EL007	11	Trend Analysis of Semantic Segmentation Technology Using Deep Learning	2018
EL008	11	Technical Trends in Electrical Storage System (EES) Application by Fields	2016
EL009	11	Home Health Care System	2004
EL010	11	Issues related to introduced environment, technology status, distribution policy, and economic performance of electric vehicles	2010
EL011	11	Research Trends in Wind Power and HVDC	2015
EL012	11	Technical guidelines for lightning protection according to IEC 62305	2014
EL013	11	Impact of Large Wind Farms on Smart Grids	2013
EL014	11	Power System Status and Major Improvements of APR1400	2011

Code	Journal	Title in English	Year of publication
EL015	11	Development Direction of Island Railroad Train Control System and Future Transportation	2008
EL016	11	Element Technologies of Autonomous Family Machine	2002
EL017	11	Insulation diagnosis technology of low voltage induction motor by partial discharge measurement	2006
EL018	11	Current Status and Challenges of Message Text Training	2016
EL019	12	Modeling of BLDC Motor Driving System for Platform Screen Door Control applied Fuel Cell Power Generation System	2017
EL020	12	Improved Model Predictive Control Method for Cascaded H-Bridge Multilevel Inverters	2018
EL021	12	Estimation of Insulation Life of PAI/Nano Silica Hybrid Coil by Accelerated Thermal Stress	2019
EL022	12	Wireless Power Transfer System Based on Semi-random Magnetic Flux	2017
EL023	12	A Study on the Efficiency Evaluation Standard and Regulation for Electric Motor Systems	2016
EL024	12	A Study on Deterioration Evaluation Method by Condition Monitoring and Diagnosis for Aging Oil-immersed Power Transformers	2014
EL025	12	A Study of Automatic Multi-Target Detection and Tracking Algorithm using Highest Probability Data Association in a Cluttered Environment	2007
EL026	12	Power Balancing Strategy in the Microgrid During Transient	2010
EL027	12	Design of Gas Identification System with Hierarchical Rule base using Genetic Algorithms and Rough Sets	2012
EL028	12	Analysis for the Ferroresonance on the Transformer by Overvoltage and Prevention Measures	2015
EL029	12	Analysis of Sequence Impedances of 345kV Cable Transmission Systems	2013
EL030	12	A Study on Static Situation Awareness System with the Aid of Optimized Polynomial Radial Basis Function Neural Networks	2011
EL031	12	The Optimal Operation of Distributed Generation Possessed by Community Energy System Considering Low-Carbon Paradigm	2009
EL032	12	Study on the Estimation of Seasonal Ambient Current for the Application of Ambient Adjusted Line Rating(AAR) in Overhead Transmission Lines Using Risk Tolerance(RT) Method	2017
EL033	12	A Study on Reliability Characteristic Curve of Transmission & Substation System considering Device Fault's Uncertainty	2008
CE004	13	A numerical Study for Improvement of Indoor Air Quality of Apartment House	2009
CE005	13	Enhanced Bio-hydrogen Production from pretreated microalgal waste	2019
CE006	13	Unit Mass Estimation and Analysis from Fiber Dyeing and Finishing Facility Nearby Nakdong River Basin	2009
CE007	13	Multidimensional Dynamic Water Quality Modeling of Organic Matter and Trophic State in the Han River System	2013
CE008	13	Preparation of Birnessite (δ-MnO2) from Acid Leaching Solution of Spent Alkaline Manganese Batteries and Removals of 1-naphthol	2016
CE009	13	A feasibility Study on the Movement Control of Particulate Enforced by Electrodynamic Force	2018
CE010	13	Characteristics of Water Quality Change of Urban River according to Installation of Interceptors and Wastewater Treatment Plant	2014
CE011	13	Analysis of Energy Savings and CO 2 Emission Reductions via Application of Smart Grid System	2017
CE012	13	A Study on Operating Condition of Test-Bed Plant using Membrane filtration of D Water Treatment Plant in Gwang-Ju	2017
CE013	13	Assessment of Strategy and Achievements of Eco Industrial Park (EIP) Initiative in Korea	2014
CE014	13	A Study of Burcucumber Biochars to Remediate Soil Pb Considering GWP (Global Warming Potential)	2015

Code	Journal	Title in English	Year of publication
CE015	13	Photocatalytic Oxidation of Arsenite Using Goethite and UV LED	2017
CE016	13	Applicability of Theoretical Adsorption Models for Studies on Adsorption Properties of Adsorbents(II)	2011
CE017	13	Study on the Quantitative Analysis of Styrene Oligomers Originated from Styrene-based Plastic Polymer Materials	2019
CE018	13	Field Application of Waterworks Automatic Meter Reading and Analysis of Household Water Use	2012
CE019	14	Behavior and Design of Hybrid Coupled Shear Wall Composed of Concrete Wall and Steel Beam	2005
CE020	14	Quality Control of Mass and Hot Weather Concrete in the Middle East Countries	2008
CE021	14	Development of High-Strength in 600, 700, 800MPa Class of Yield Strength and Seismic Resistant Steel Deformed Bar	2010
CE022	14	Shear Capacity of RC Beams and Seismic Performance of RC Buildings and Bridge Piers with High-Strength Steel Reinforcement	2010
CE023	14	Volumetric Stability and its Significance in Concrete	2005
CE024	14	Shrinkage Properties Evaluation of High Strength Concrete for Control of Column Shortening	2014
CE025	14	Advanced Technology of Dam Concrete	2004
CE026	14	Mix Proportioning and Constructability of Lightweight Aggregate Concrete	2011
CE027	14	Design of UHPC(Ultra High Performance Concrete) Girder Cable Stayed Footbridge	2011
CE028	14	Seismic Performance Assessment of RC Bridge, Columns	2007
CE029	14	Concrete Repairing by Calcium Carbonate-Precipitating Bacteria	2016
CE030	14	Investigation of Material Strength and Member Detail on Existing Concrete Building	2015
CE031	14	Structural Capacity and Constructability of Composite Concrete Members(TSC beam & PSRC column)	2016
CE032	14	Introduction of Decommissioning and Dismantling Technology for Nuclear Power Plant Structure Based on Domestic and International Practices	2016
CE033	14	Anchorage Behavior of D22 SD600 Reinforcing Bars in Compression Terminated in Exterior Beam-column Joint of 60 and 90 Mpa Concrete	2019
AR004	15	Earthquake-Resistant Capacity of Reinforced Concrete Beam-Column Joints with Steel Fiber	2009
AR005	15	An Experimental Study on the Estimation of Compressive Strength of Ultra-High Strength Concrete Using Warm Water Curing Method	2011
AR006	15	A Root Cause Analysis of Problems in the Use of Environmental Information for Construction Projects: Focusing on Environmental Laws and Regulations	2014
AR007	15	Exact Solution Computation Model for Time-Cost Tradeoff Analysis	2015
AR008	15	A Study on Development of New Assessment Method for the Energy Performance in G-SEED using Dynamic Building Energy Simulation Programs	2018
AR009	15	Estimation and Feature of Greenhouse Gas Emission in Building Sector by National Energy Statistic	2019
AR010	15	Evaluation of the Effect of a Window Protective Device for Securing Safety of Windows in Apartments against Strong Wind	2016
AR011	15	Analysis of Risks Associated with Climate Change in Public Agencies' Buildings	2017
AR012	15	Plan for Introduction and Administration of Performance Warranty Contracting (Based on the Case of Highway Pavement)	2010
AR013	15	A case study of feasibility analysis and decision making method for the construction project development	2003
AR014	15	Experimental Study on the Vibrational Serviceability and Characteristics of Floating Floor Structures	2005
AR015	15	Study on the Relationship between Moisture Content and Length Change of PCM	2007

Code	Journal	Title in English	Year of publication
AR016	15	An Analysis of Effects through Improved Insulation Perfomance for High-Density Residential Area on West High East Low Type	2018
AR017	15	A Study on the System Development for Optimum Method Section in Demolition Works	2008
AR018	15	Effect of Emulsified Refine Cooking Oil on the Engineering Properties of High Volume Admixture Concrete	2013
AR019	16	A Study on the Creating and Prosperity Process of the Siheyuan in China	2005
AR020	16	A Case Study on the Actual Condition and Composition Method of Environment-Friendly Architecture of the School Facilities -Focused on Ecological Environment	2009
AR021	16	A Study on Cheongju-eup Townscape in the Late 1930s by Modeling the Restoration Image	2019
AR022	16	A Study on Current Situation of Rural Community Facilities Applied Universal Design	2015
AR023	16	A Study on the Landscape Color Analysis of the Energy Industry in Jeju Rural Area	2017
AR024	16	A Research on the Purpose of Use and Selection Factor of Walking Trails	2018
AR025	16	A Study on Urban Inhabitants' Consciousness for Urban-Countryside Complex Type Villages Development	2012
AR026	16	A Study on Improvements and Types of Accidents of Kindergarten in Cheongju, Chung-buk Area	2014
AR027	16	A Study on Manufacturing Norms of Wollyang(月梁) Head in the 《Yeongjobeosig(營造法式)》 Song Dynasty	2019
AR028	16	A Study on the Space Composition and the Visitor Circulation of the Theme Museum in Jeju	2018
AR029	16		2005
AR030	16	The Space Plan and Design Characteristic for the Gungjip in Namyangju	2010
AR031	16	Research on the Direction of Forming Rural and Fishing Village Type Community Living Home through Regeneration	2013
AR032	16	Workshop Method Adaptation of SI Theory for Applying Closed Schools	2011
AR033	16	A Study on the Investigation of Color about fishing village and harbor in Gangwon East coast	2006
IE004	17	Generalized Single Manufacturer and Multiple Retailers Supply Chain Model in JIT Purchasing	2011
IE005	17	Scheduling Algorithms for Minimizing Total Weighted Flowtime in Photolithography Workstation of FAB	2012
IE006	17	A New Metric for Evaluation of Forecasting Methods : Weighted Absolute and Cumulative Forecast Error	2015
IE007	17	Analysis of Reward and Royalty Programs Affecting Customer Satisfaction and Recommendations in the Purchase Process in Luxury Goods	2018
IE008	17	On the Exact Cycle Time of Failure Prone Multiserver Queueing Model Operating in Low Loading	2016
IE009	17	Comparisons of Airline Service Quality Using Social Network Analysis	2019
IE010	17	The Effect of Smartphone Purchasing Determinants on Repurchase Intention	2017
IE011	17	Crane Scheduling Considering Tenant Service Time in a Rail-Road Transshipment Yard : Case of the Uiwang ICD	2018
IE012	17	Module Communzation for Product Platform Design Using Clustering Analysis	2014
IE013	17	Effects of Decision Making Style and Working Conditions on Organizational Effectiveness : Focused on R&D Employee of Small Business	2013
IE014	17	Determination of Quality Cost Policy under Multiple Assignable Causes	2003
IE015	17	A Study on Conversion of Supply Chain Organization Based on Theory of Sense and Response : Focus on Marine Corps's supply chain	2007

Code	Journal	Title in English	Year of publication
IE016	17	A Empirical Study on the Moderate Effect of Appraisal Instrumentality in the Effect of the Appraisal Justice on the Organization Commitment : Focused on the University Employees	2009
IE017	17	Efficient Method of Fixing the Setting (4M standardization) in Melting Furnace Operation	2007
IE018	17	Algorithms for Fire Sequencing Problem in Unplanned Artillery Attack Operation	2012
IE019	18	A Solution for Sourcing Decisions under Supply Capacity Risk	2016
IE020	18	A Comparative Study on the Trend of Technological Convergence	2013
IE021	18	Efficiency Comparison and Performance Targets for Academic Departments in the Local Private College Using DEA	2013
IE022	18	Tabu Search Heuristic Algorithm for Designing Broadband Convergence Networks	2008
IE023	18	Ubiquitous Computing Technology Based Environmental Monitoring and Diagnosis System : Architecture and Case Study	2010
IE024	18	An Ant Colony Optimization Approach for the Maximum Independent Set Problem	2007
IE025	18	A Market-Based Replacement Cost Approach to Technology Valuation	2015
IE026	18	The Maximal Covering Location Problem with Cost Restrictions	2004
IE027	18	Issues and Efforts for Technology-Humanities Convergence : Empirical Analysis of Korean SMEs	2014
IE028	18	Development and Implementation of an integrated Evaluation System for Continuous Maturity of IS Performance	2003
IE029	18	Study on the Heterogeneous Fleet Vehicle Routing Problem with Customer Restriction	2005
IE030	18	Valuation and Optimal Timing of the Investment in Next Generation Telecommunication Service Using Real Options	2006
IE031	18	A Quality Management Model Contingent to R&D Characteristics	2017
IE032	18	Estimating the Reliability of Virtual Metrology Predictions in Semiconductor Manufacturing : A Novelty Detection-based Approach	2012
IE033	18	A Fact-oriented Ontological Approach to Process Modeling for Knowledge-based Services	2009
MO004	19	Dynamic Response of Polyurethane Foam with Density and Temperature Effects	2019
MO005	19	Fundamental Studies for Ventilated Supercavitation Experiments in New High-speed Cavitation Tunnel	2018
MO006	19	A Study on Quantifying Sailing Safety Considering Maneuverability of a Vessel	2017
MO007	19	An Application of AHP for the Selection of Optimum Product of BWTS for over 10,000 TEU Container Ship	2015
MO008	19	An Approach for Construction of Shipyard Simulation Environment based on Neutral File Format	2016
MO009	19	Research of Design Improvement regarding Foundation Technologies for Floating LNG	2014
MO010	19	Numerical Analysis of Supercavitating Flows of Two-Dimensional Simple Bodies	2013
MO011	19	Laminar Flow Structures Near a Circular Cylinder in between a Free-Surface and a Moving Wall	2012
MO012	19	A Feasibility Study on the RPM and Engine Power Estimation Based on the Combination of AIS and ECMWF Database to Replace the Full-scale Measurement	2017
MO013	19	A Numerical Study of Effects of Body Shape on Cavity and Drag of Underwater Vehicle	2018
MO014	19	Numerical Analysis of Welding Residual Stresses for Ultra-Thick Plate of EH40 Steel Joined by Tandem EGW	2010
MO015	19	Development of a Numerical Method for the Evaluation of Ship Resistance and Self-Propulsion Performances	2011
MO016	19	Computation of Pressure Fields in the Lagrangian Vortex Method	2004
MO017	19	Cavitating-Flow Characteristics around a Horn-Type Rudder	2007

Code	Journal	Title in English	Year of publication
MO018	19	Study on Resistance Performance of Icebreaking Cargo Vessel in Pack Ice Condition according to Variation of Synthetic Ice Thickness and Hull Form Characteristics	2009
MO019	20	Estimation of Hydrodynamic Coefficients for an AUV Using Nonlinear Observers	2006
MO020	20	3-D Dynamic Response Characteristics of Seabed around Composite Breakwater in Relation to Wave-Stucture-Soil Interaction	2016
MO021	20	An Analysis of the Variation in the Settling Properties of Cohesive Sediments before and after Closure of the Saemankeum Seadike	2008
MO022	20	Global Path Planning for Autonomous Underwater Vehicles in Current Field with Obstacles	2012
MO023	20	Improvement of Tidal Circulation in a Closed Bay using Variation of Bottom Roughness	2005
MO024	20	A Study on Towing Characteristics of Barge Considering Wind Force	2015
MO025	20	Numerical Study for Experiment on Wave Pattern of Internal Wave and Surface Wave in Stratified Fluid	2019
MO026	20	Vibration-based Structural Health Monitoring of Caisson-type Breakwaters Damaged on Rubble Mound	2010
MO027	20	Flow Survey around Two-Dimensional Circular Cylinder using PIV Technique	2004
MO028	20	Comparison of Analysis Methods for Designed Spudcan Bearing Capacity and Penetration Behavior for Southwest Sea Soil	2015
MO029	20	Fracture Simulation of Low-Temperature High-Strength Steel (EH36) using User-Subroutine of Commercial Finite Element Code	2014
MO030	20	Optimum Structural Design of Pipe Loops Used in Large Vessels (I)	2007
MO031	20	Study of Dynamic Characteristics of West Coast Saemangeum Sand by Torsional Shear Test	2013
MO032	20	Estimation of Hedrodynamic Coefficients for an AUV Using Nonlinear Observers	2006
MO033	20	Shearing Properties of Waste Tire Powder-Added Lightweight Soil by Direct Shear Test	2009
RM004	21	Development and Field Test of a Distributed Acquisition System for High Efficiency Deep DC Resistivity Surveys	2019
RM005	21	Estimation of Solar Irradiance at Weather Stations in Korea Using Regionally Trained Artificial Neural Network Models	2019
RM006	21	Simulation of Grinding/Classification Circuit in Domestic Gold ore Processing Plant Using Energy-based Grinding Model and Mathematical Classification Model	2018
RM007	21	History Matching of Gas Production Rates Integrated an Artificial Neural Network with Distance-based Candidate Selection	2017
RM008	21	Investigation of Bond's Work Index of Korean Desulfurization Limestone by Ball Mill	2016
RM009	21	Froth Flotation Characteristics of Cassiterite ores with Oleic Acid in Acidic Region	2015
RM010	21	Hydrogeochemical Assessment on Physico-chemical Treatment Process of Coal Mine Drainage	2013
RM011	21	Guidelines for Experimental Design and Criteria to Evaluate the Performance of Rotary Steerable System	2016
RM012	21	Hydromechanical dynamics of hydraulic and natural fractures	2018
RM013	21	Relationship between Surface Roughness and Contact Angle of Pyrite	2014
RM014	21	An Evaluation of the Influence of the Rock Property Change under Saturated-Loading Conditions on Rock Stability	2015
RM015	21	Effects of Well Fluids Characteristics on the Design of Offshore Gas Production System: Case Studies of Western Australian Gas Fields	2013
RM016	21	Classification and Example of Assessment Methodologies for Shale Resources	2017
RM017	21	Estimation of Solar Irradiance at Weather Stations in Korea Using Regionally Trained Artificial Neural Network Models	2019
RM018	21	Optimum Receiver Geometry for Moment Tensor Inversion of Microseismic Data	2014

Code	Journal	Title in English	Year of publication
RM019	22	Risk Assessment of Fire and Explosion of Methane	2005
RM020	22	Major Control Techniques for Chemical and Gas Process Industires	2011
RM021	22	A study on Knocking Characteristics of a 300kW Class CNG Engine for CHP	2008
RM022	22	A Study on Determination of Range of Hazardous Area Caused by the Secondary Grade of Release of Vapor Substances Considering Material Characteristic and Operating Condition	2018
RM023	22	Feasibility Study of Pressure Letdown Energy Recovery from the Natural Gas Pressure Reduction Stations in South Korea	2015
RM024	22	The Development and Introduction of External Corrosion Direct Assessment Measures for Urban Gas Pipelines	2014
RM025	22	Experimental Study on the Performance Characteristics of a Scroll Expander for 1kW-class Organic Rankine Cycle	2015
RM026	22	The necessity of Introducing the In-service Test based on Analysis of Performance Test Result of Pressure Safety Valve	2017
RM027	22	Consequence Analysis of Toxic Gases Generated by Fire of Lithium Ion Batteries in Electric Vehicles	2018
RM028	22	Adsorption Characteristics of Nitrogen monoxide over Dealuminated and Alkali/Alkaline-earth Metal Ion Exchanged Y-Zeolites	2005
RM029	22	Parameter Analysis of the Damage Area and the Financial Loss by the Gas Release Accident at Pressure Vessels	2011
RM030	22	A Study on DNA Degeneration by Comet Assay & Pathological Observation for Mouse Which were Exposed HCN Gases from Fire	2012
RM031	22	Frequence Analysis for City Gas Pipeline	2003
RM032	22	A Study on the Physical and Chemical Characteristics and Hazards Buthanethiol	2013
RM033	22	A Study on the Correlation of MESG and Explosion Pressure	2016
EC004	23	A study on th Mean Variations of EEG for the indirect moxibustion stimulation	2008
EC005	23	Analysis on the EMC evaluating method for applying wireless communications in NPP	2017
EC006	23	Biological Early Warning System for Toxicity Detection	2010
EC007	23	The Implementation of Idle Stop System with the OBD-II Interface in the Automotive Smart Key System	2013
EC008	23	Effects of Self-assessment using Smartphone Video Recording on Essential Fundamental Nursing Skills Education	2018
EC009	23	On the Spectral Efficient Physical-Layer Network Coding Technique Based on Spatial Modulation	2016
EC010	23	Optimal Control of Time and Energy for Mobile Robots Using Genetic Algorithm	2017
EC011	23	Nature and Sources of Business Values in Digital Social Innovation	2019
EC012	23	Performance Improvement of Speaker Recognition Using Enhanced Feature Extraction in Glottal Flow Signals and Multiple Feature Parameter Combination	2015
EC013	23	Fimplementation of RF Controller based on Digital System for TRS Repeater	2005
EC014	23	The effects on Social Welfare of Asymmetric Regulation in Mobile Telecommunications Termination Access Charge	2009
EC015	23	Internet Network Pricing under the Change of Internet Traffic Patterns	2014
EC016	23	Construction of Complemented Hybrid Group Cellular Automata with Maximum Equal Lengths	2006
EC017	23	Design of Shipboard integrated network platform for Digital-ship	2005
EC018	23	Optimization of Mobile Robot Predictive Controllers Under General Constraints	2018
EC020	24	Present and Future of IoT-based Smart Healthcare Industry	2016
EC021	24	Status and Prospect of Medical Information Security	2010
EC022	24	Neuromorphic Architecture and CAD Research Trends (Reverse-Engineering of the Brain)	2012

Code	Journal	Title in English	Year of publication
EC023	24	Trend of 3D image reconstruction method using low resolution depth information	2011
EC024	24	Development and Prospect of Computerized Tomography	2017
EC025	24	Information and Communication Technology (ICT) Standardization Policy	2009
EC026	24	Methods of system component quality improvement through enterprise architecture	2008
EC027	24	Advances in Terrestrial DTV Broadcasting Technology	2006
EC028	24	Auditory Algorithm Research Platform	2016
EC029	24	Trends in Information Security Technology in the Financial Sector	2013
EC030	24	SoC Design Technology Trend for Wearable Device UIUX	2014
EC031	24	Linearization technology of high output power damper	2002
EC032	24	Graphene device technology	2015
EC033	24	Necessity and Technology Development Status of Domestic Electric Vehicles	2015
EC034	24	Voice interface technology view and outlook	2017

Appendix C.

Textbooks selected for the corpus

Sub-field	Title	Authors	Year
ME	*Ilban gigye gonghak* [General Mechanical Engineering]	Kim Gi-mun, Kim Dong-gi, Park Il-ju, Im Hong-seop, Lee Hwa-sun	2016
AE	*Hanggonguju gonghak* [Aerospace Engineering]	Song Yun-seop, Kang Seung-hui	2014
CS	*Keompyuteo gaeron* [Introduction to computers]	Yu Chi-hyeong, Kim Do-yeon, Yu Han-na	2017
CB	*Saengmulhwahakgonghak* [Biological and Chemical Engineering]	Jang Ho-nam, Seo Jin-ho	2006
EL	*Gicho jeongijeonja gonghak* [Basics of Electrical and Electronics Engineering]	Jin Sang-ho	2018
CE	*Jayeongwa munmyeongui johwa tomokgonghak* [Civil Engineering (Harmony of Nature and Civilization)]	*Daehantomokakoe chulpanwiwonhoe* [Korean Society of Civil Engineers Publishing Committee]	2018
AR	*Geonchukwangyeong gonghang geron* [Architectural Environmental Engineering]	Kim Hak-cheol, Park Jae-hun, Seo Myeong-gseok, An Tae-gyeong, Yang Gwan-mok, Yoon Seok-ho, Lee Yeong-do, Lee Yun-hui, Lee Jongho, Jung Hwan-mok, Choi Yeong-cheol, Han Yangsu	2003
IE	*Seumateu sesangeul yeoneun saneopgonghak* [Industrial engineering opening a smart world]	*Daehan saneopgonghakoe* [Korean Society of Industrial Engineers]	2016
MO	*Seonbakaeyanggonghang gaeron* [Introduction to Marine Engineering]	Lee Seung-jun, Yum Deok-jun	2018
RM	*Jawongebalhwangyeonggonghak* [Resource Development and Environmental Engineering]	Lee Chan-gu, Kim Jin, Kim Jae-dong, Jeon Seok-won, Kim Seon-jun, Jung Myeong-chae, Im Gil-jae, Jeong Yeong-uk	2018
EC	*Jeongbotongsingonghak* [Information and Communication Engineering]	Im Seung-ha, Ku Gi-joon	2002

Appendix D.

Mecab tags set

NNG	noun	JC	conjunctive postpositional marker
NNP	proper noun	JX	auxiliary postpositional marker
NNB	bound noun	EP	pre-final ending
NNBC	unit counting noun	EF	sentence-closing ending
NR	numeral	EC	connective ending
NP	pronoun	ETN	nominalizer suffix
VV	verb	ETM	adnominal ending
VA	adjective	XPN	prefix
VX	auxiliary predicate	XSN	noun derivational suffix
VCP	copula	XSV	verbal derivational suffix
VCN	negative copula	XSA	adjectival derivational suffix
MM	determiner	XR	root
MAG	adverb	SF	period, question mark, exclamation mark
MAJ	conjunctive adverb		
IC	exclamation	SE	ellipsis
JKS	subject case marker	SSO	opening bracket
JKC	complement case marker	SSC	closing bracket
JKG	adnominal case marker	SC	delimiter
JKO	object case marker	SY	other symbols
JKB	adverbial case marker	SH	Chinese character
JKV	vocative case marker	SL	foreign word
JKQ	citation marker	SN	number

Appendix E.

Words in the corpus covered by level A vocabulary of KLL

가게 가깝다 가끔 가다 가르치다 가방 가볍다 가수 가슴 가운데 가을 가장 가져오다 가족 가지다 간호사 감기 감사 갑자기 값 강 같다 같이 개 개월 거기 거리 거울 걱정 건강 건물 걷다 걸다 걸어가다 검은색 것 게임 겨울 결혼 경복궁 경주 경찰 경찰관 경찰서 경치 계란 계속 계시다 계절 계획 고기 고등학교 고등학생 고향 곧 곳 공 공부 공원 공책 공항 공휴일 과 과일 과자 괜찮다 교과서 교수 교실 교통 교회 구 구경 구두 구름 군인 권 귀 그 그거 그것 그곳 그동안 그때 그렇다 그릇 그리다 그림 극장 근처 급 기다리다 기분 기숙사 기차 길 길다 김치 깎다 꼭 꽃 꿈 끄다 끝 끝나다 끝내다 나 나가다 나다 나라 나무 나쁘다 나오다 나이 나중 날 날다 날씨 날짜 남녀 남대문 남자 남쪽 남편 남학생 낮 낮다 내년 내다 내려가다 내려오다 내리다 내일 냉장고 너 너무 넓다 넣다 네 넷 넷째 년 노란색 노래 노트 놀다 놀라다 높다 놓다 누구 눈 눈물 뉴스 늦다 다 다니다 다르다 다른 다리 다섯 다섯째 다시 다음 닦다 단어 닫다 달 달다 달러 담배 대답 대학 대학교 대학생 대화 더 덥다 도서관 도시 도착 독일 돈 돌아가다 돌아오다 돕다 동물 동생 동안 동쪽 돼지 되다 두 둘 둘째 뒤 드리다 듣다 들다 들어가다 들어오다 등산 딸 딸기 때 때문 떠나다 또 똑같다 똑바로 뛰다 뜨겁다 라디오 라면 러시아 마리 마시다 마음 마지막 만 만나다 만들다 많다 많이 말 말다 말씀 맑다 맛 맛있다 매우 매일 맥주 맵다 머리 먹다 먼저 멀다 메뉴 며칠 명 몇 모두 모든 모르다 모자 목 목욕 몸 못 무겁다 무슨 무엇 문 문제 묻다 물 물건 물론 뭐 미국 미터 밑 바꾸다 바나나 바다 바람 바로 바지 박물관 밖 반 반갑다 받다 발 발음 밝다 밤 밥 방 방학 배 배우다 백 백화점 버리다 버스 번 번호 벌써 벗다 별 병 병원 보내다 보다 보통 봄 부르다 부모 부부 부산 부엌 부인 북쪽 분 불 불다 비 비누 비디오 비싸다 비행기 빠르다 빨간색 빨리 빵 사 사과 사다 사람 사랑 사무실 사용 사이 사장 사전 사진 사탕 산 산책 살 살다 삼 새 색 색깔 샌드위치 생각 생기다 생일 생활 서다 서로 서울 서울역 서점 서쪽 선물 선생 설명 설탕 세 세수 세탁기 센티미터 셋 셋째 소개 소금 소파 속 손 손가락 손님 쇼핑 수 수건 수업 수영 수영장 숙제 술 쉬다 쉽다 슈퍼마켓 스무 스키 스트레스 스포츠 시 시간 시계 시원 시작 시장 시험 식당 식사 식탁 신다 신문 신발 실례 십 싶다 싸다 싸우다 쓰다 쓰레기 씨 씻다 아기 아내 아니 아니다 아들 아래 아름답다 아마 아무 아버지 아빠 아이 아이스크림 아저씨 아주 아직 아침 아파트 아프다 아홉 안 안경 안다 앉다 않다 알다 앞 야구 약 약국 약속 양복 얘기 어 어깨 어느 어디 어떤 어떻다 어렵다 어른 어린이 어머니 어서 어제 언제 언제나 얼굴 얼마 얼마나 엄마 없다 에어컨 여권 여기 여덟 여러 여러분 여름 여섯 여자 여학생 여행 역 역사 연습 연필 열 열다 열쇠 열심히 영국 영어 영화 옆 예 옛날 오 오늘 오다 오래 오랜만 오르다 오른쪽 오전 오후 올라가다 올해 옷 왜 외국 외국어 외국인 왼쪽 요리 요일 요즈음 요즘 우리 우리나라 우산 우유 운동 운동장 운전 울다 웃다 원 월 위 위험 유명 육 은행 음식 음악 의사 의자 이 이거 이것 이곳 이다 이때 이런 이렇다 이름 이번 이분 이야기 이월 이제 이해 인사 인천 일 일곱 일본 일어나다 일요일 일월 일찍 읽다 잃다 입 입다 있다 잊다 잎 자다 자동차 자리 자장면 자전거 자주 작년 작다 잔 잘 잠 잠깐 잠시 잠자다 잡다 잡지 장 장미 장소 재미 재미있다 저 저기 저녁 적다 전 전화 점심 젓가락 정류장 정말 제일 제주도 조금 조용 졸업 좀 종이 좋다 주 주다 주말 주소 주스 주인 죽다 준비 중국 중요 중학교 중학생 즐겁다 지갑 지금 지내다 지도 지우다 지하 지하철 질문 집 짜다 짧다 쪽 찍다 차 참 창문 찾다 책

책상 처음 천 천천히 첫째 청소 초대 초등학교 초콜릿 추다 축구 축하 출발 춤 춥다 취미 층 치다 치마 친구 친절 칠 칠판 침대 카드 카메라 칼 캐나다 커피 컴퓨터 컵 켜다 코 콜라 크다 크리스마스 키 타다 태어나다 택시 테이블 텔레비전 토요일 팀 파란색 파티 팔 팔다 퍼센트 편지 포도 표 프랑스 피곤 피아노 피우다 필요 하나 하늘 하다 학교 학년 학생 한 한강 한국 한국어 한글 한번 한자 할머니 할아버지 함께 항상 해 핸드폰 햄버거 허리 형 호 호주 호텔 혼자 화 화장실 환자 회사 회의 후 휴지 흰색 힘 힘들다

Appendix F.

Words in the corpus covered by level B vocabulary of KLL

가격 가구 가까이 가까이 가꾸다 가난 가늘다 가능 가득 가득 가로 가루 가리다 가리다 가리키다 가만히 가스 가요 가위 가이드 가정 가져가다 가져다주다 가죽 가지 가지 가짜 각 각각 각각 각국 각자 각자 각종 간 간단히 간장 갈다 갈다 갈색 갈아타다 감 감다 감동 감상 감상 감자 감정 강당 강도 강물 강원도 강제 강조 강하다 갖다 갖다 개구리 개발 개인 거리 거실 거의 거짓 건너가다 건너다 건너편 건축 걸리다 걸음 검다 검사 겁 겉 게다가 겨우 겨울철 견디다 결과 결국 결심 결정 결정 경기 경기도 경기장 경상도 경영 경우 경제 경험 경험 계단 계산 계산 계산기 계속 계약 계획 고개 고객 고교 고급 고려 고르다 고르다 고민 고민 고생 고생 고속 고속도로 고장 고장 고추장 고치다 고통 곧바로 골목 골목길 골프 곰 곱다 곳곳 공간 공기 공동 공무원 공사 공연 공장 공짜 공항버스 과거 과목 과장 과제 과학 과학자 관객 관계 관광 관광객 관광지 관련 관련 관리 관심 관찰 관하다 광고 광주 굉장히 교류 교문 교육 교통사고 교환 구경 구멍 구하다 구하다 국 국가 국기 국내 국립 국물 국민 국수 국어 국제 국회의원 군 군 군대 군데 굳이 굵다 굽다 귀국 귀국 규칙 그냥 그늘 그다지 그대로 그램 그런 그렇게 그려지다 그룹 그리 그림자 그립다 그만 그만 그만큼 그중 그치다 그해 근로자 근무 근무 글 글씨 글자 금 금년 금방 금지 금지 급하다 기간 기계 기대 기대 기도 기록 기르다 기름 기본 기쁘다 기쁨 기사 기사 기술 기억 기억 기억나다 기온 기운 기자 기준 기초 기침 기타 기회 긴장 긴장 길가 길거리 길이 김 깊다 깊이 깊이 까닭 깜짝 깨끗이 깨다 깨다 깨지다 꺼내다 꺼지다 껌 껍질 꼭대기 꽃잎 꽤 꾸다 꾸미다 꾸준히 꿈꾸다 끊다 끌다 끓다 끓이다 끼다 끼다 나누다 나다 나머지 나뭇가지 나뭇잎 나비 나서다 나타나다 나타내다 나흘 낙엽 낚시 날개 날아가다 낡다 남 남 남기다 남다 남산 남성 낫다 낫다 낯설다 낳다 내내 내놓다 내려놓다 내밀다 내용 내지 냄비 냄새 너머 너무나 넓다 넘다 넘어가다 넘어서다 넘어지다 넘치다 녀석 노동자 노력 노력 노인 녹색 녹음 녹음 녹차 논문 논의 놀랍다 놀이 놀이터 농담 농사 농업 농촌 높이 높이 높이다 놓이다 놓치다 누르다 눈앞 눕다 뉴욕 느끼다 느낌 느리다 늘 늘다 늘리다 늘어나다 능력 다가가다 다가오다 다수 다양 다치다 다행히 단 단맛 단순 단순히 단점 단지 단지 단체 닫히다 달걀 달다 달다 달라지다 달리다 달리다 달빛 닮다 담그다 담다 담당 담요 답장 당근 당기다 당신 당연 당연히 당장 닿다 대 대 대구 대기업 대단 대단히 대도시 대문 대부분 대신 대전 대중 대중교통 대통령 대표 대하다 대학원 대한민국 대화 대회 더위 덕분 던지다 덜 덮다 데 도구 도로 도로 도시락 도움 도자기 도중 독서 돌 돌다 돌려주다 돌리다 돌아서다 동기 동네 동시 동아리 동양 동전 동화 되게 된장 두껍다 두다 두다 두드리다 두부 두세 두통 둥글다 뒤집다 뒤쪽 드라마 들 들다 들려오다 들려주다 들리다 들어서다 들여다보다 듯이 등 등 등 등록 등록금 디자이너 디자인 따다 따라가다 따라오다 따로 따르다 따르다 딱 딴 땀 땅 때때로 떠오르다 떠올리다 떨다 떨리다 떨어뜨리다 떨어지다 떼다 똑같이 뚜껑 뛰어들다 뜨다 뜨다 뜻 라이터 런던 레스토랑 렌즈 리듬 마늘 마련 마루 마을 마음대로 마주 마중 마찬가지 마치 마치다 마침 마침내 막 막 막다 막히다 만 만남 만약 만일 만족 만족 만큼 만화 말 말 맞다 맞다 맞다 맞은편 맞추다 맡기다 맡다 매년 매다 매력 맨 머리카락 머무르다 먹이다 먼지 멀리 멈추다 멋 멋있다 메다 메모 메시지 메일 명령 명절 명함 몇몇 모기 모니터 모델 모래 모레 모습 모시다 모양 모양 모으다 모이다 모임 모자라다 목소리 목욕탕 목적 목표 몰다 몰래 몸무게 몸살 몹시 무 무게 무늬 무더위 무렵 무릎 무시 무어 무역 무용 무조건 묶다 문장 문제점 문학 문화 묻다 물고기 물다 물론 물속 물음 미디어 미래 미리 미소 미술 미술관 미인 미치다 미치다 민족 믿다 믿음 밀가루 밀다 밀리미터 밉다 바깥 바뀌다 바늘 바닥 바닷가 바닷물 바라다 바라보다 바르다 바르다 바보 바위 박사 박스 반대 반대 반드시 반복 반장 반지 반하다 받아들이다 받침 발견 발견 발달 발달 발등 발목 발생 발생 발음 발전 발전 발표 발표 밝다 밝히다 밟다 밤 밤낮 밤새다 밥솥 방문 방문 방문 방법 방송 방송국 방향 밭 배경 배구 배달 배우 버리다 버섯 버터 버튼 번역 벌다 벌레 벌리다 벌어지다 벌이다 법 법 벨트 벽 변하다 변호사 변화 변화 별 별로 병들다 보고 보고서 보관 보람 보이다 보이다 보통 보험 보호 복도 복사 복사 복습 복습 본래 봉지 봉투 뵙다 부 부근 부동산 부드럽다 부딪치다 부르다 부분 부자 부작용 부장 부족 부족 부족 부지런 부처 부탁 부탁 북 북한 분명 분명히 분석 분위기 불교 불꽃 불리다 불만 불빛 불안 불안 불편 불행 붉다 붓다 붙다 붙이다 비교 비교 비교적 비닐 비밀 비서 비용 비타민 비판 빌다 빌딩 빌리다 빗 빗물 빛 빠지다 빠지다 빨다 빨래 빼다 뻗다 뼈 뽑다 뿌리 뿌리다 뿐 사계절 사고 사과 사라지다 사무 사물 사실 사실 사업 사용 사용자 사원 사이좋다 사자 사회 산소 살 살리다 살짝 살펴보다 삶다 상 상대 상대방 상상 상상 상자 상처 상추 상품 상하다 새끼 새로 새벽 새우 새해 생각나다 생겨나다 생신 생활 서너 서랍 서로 서류 서비스 서양 서투르다 석유 섞다 섞이다 선배 선수 선택 선택 선풍기 설거지 섬 성 성격 성공 성공 성별 성인 성적 성함 세계 세금 세기 세다 세로 세상 세우다 세탁 세탁소 센터 소 소개 소나기 소나무 소년 소리 소문 소비자 소설 소스 소식 소포 속도 속옷 속하다 손목 손바닥 손잡이 송아지 송이 쇼 수고 수년 수돗물 수많다 수상

수술 수입 수입 수입 수출 수출 수표 수학 수학 수화기 숙소 숙이다 순간 순서 순수 숨 숨다 숫자 숲 쉬다 스스로 스스로 스케줄 스타 스타일 슬픔 습관 승용차 시골 시끄럽다 시내 시내버스 시대 시도 시리즈 시민 시설 시외 시인 시절 시청 시청자 시키다 식 식구 식다 식물 식빵 식품 신 신고 신문지 신부 신선 신용 신입생 신청 신체 신호 신호등 싣다 실내 실력 실수 실수 실제 실패 실패 심각 심다 심리 심부름 심하다 싸다 싸움 쌀 쌓다 쌓이다 썩다 쏘다 쏟다 쏟아지다 쓰다 쓰이다 씨 아까 아까 아니 아래쪽 아래층 아마도 아무것 아무래도 아무리 아시아 아이 아프리카 악기 안내 안방 안전 안전 안쪽 안타깝다 알 알리다 알맞다 알아듣다 앓다 암 앞뒤 앞서다 앞쪽 애 애인 액세서리 야 야외 야채 약 약간 약간 약하다 얇다 양 양념 양배추 양보 양쪽 양파 얕다 얘 어둠 어둡다 어려움 어리다 어리다 어린아이 어울리다 어쩌면 억 얹다 얻다 얼다 얼른 얼음 업무 없애다 없이 에너지 엔 엘리베이터 여 여관 여대생 여름철 여성 여유 역시 역할 연결 연구 연구 연구소 연구자 연기 연기 연락 연락처 연말 연세 연하다 연휴 열 열리다 열리다 열차 염려 영상 영원히 영하 옆방 옆집 예 예금 예매 예상 예술 예술가 예약 예전 예정 옛 오 오가다 오늘날 오래전 오랜 오랫동안 오른발 오른손 오리 오븐 오직 옥수수 온 온도 온통 올라오다 올려놓다 올림픽 옮기다 옳다 와 와인 완벽 완전히 왕 외 외교 외교관 외롭다 왼발 왼손 요구 요금 요새 요청 욕실 욕심 용돈 우선 우수 우승 우승 우연히 운 운전기사 운전자 울리다 울산 움직이다 원래 원숭이 원하다 월급 월드컵 위반 위아래 위쪽 위층 위치 위치 위하다 유교 유난히 유럽 유리 유리창 유명 유치원 유학 유학 유행 유행 음 음료 음료수 응 의견 의미 의미 의심 의하다 이기다 이날 이동 이러 이렇게 이루다 이루어지다 이르다 이르다 이리 이마 이모 이미 이미지 이불 이사 이사 이상 이상 이상 이상 이성 이외 이용 이용 이웃 이웃집 이유 이익 이전 이틀 이하 이해 이혼 이혼 이후 익다 인간 인구 인기 인분 인삼 인상 인상 인생 인원 인제 인터넷 인터뷰 일기 일기 일반 일부 일상 일상생활 일으키다 일정 일회용 일회용품 임금 임시 임신 입구 입술 입원 입원 입학 입학 자 자가용 자격 자기 자기 자꾸 자동 자라다 자랑 자료 자르다 자세히 자식 자신 자신 자연 자유 자체 자판기 작가 작아지다 잘못 잘못 잘못 잠들다 잠시 잡히다 장 장갑 장난감 장래 장마 장사 장점 장학금 재료 재산 재채기 저 저리 저자 저축 적 적 적극 적당 적당히 적어도 적어지다 적용 전공 전국 전기 전날 전달 전문 전문가 전부 전부 전자 전쟁 전철 전체 전통 전하다 전혀 전화기 절 절 절대 절대로 절반 절약 젊다 점 점 점수 점원 점점 점차 접시 젓다 정 정거장 정답 정도 정리 정리 정말 정말 정말로 정문 정보 정상 정식 정신 정원 정치 정하다 정확 정확히 젖다 제공 제대로 제목 제발 제법 제출 제품 제한 조각 조건 조사 조사 조상 조심 조용히 졸업생 좁다 종 종교 종류 종업원 종일 종합 좌석 주로 주머니 주먹 주문 주변 주부 주사 주요 주위 주인공 주장 주제 주차 주차 주차장 주택 죽 죽다 죽음 죽이다 준비물 줄 줄 줄기 줄다 줄무늬 줄이다 중 중간 중심 중요 쥐 쥐다 즉시 즐거움 즐기다 증상 증세 지 지구 지구 지나가다 지나다 지나치다 지난해 지다 지다 지다 지르다 지방 지붕 지역 지점 지키다 지하도 직업 직원 직장 직접 직접 진짜 진짜 진출 진하다 진행 질 질서 짐 집다 집안 집중 짓 짓다 짙다 짜증 짧아지다 쪽 차갑다 차다 차다 차례 차이 착하다 참 참가 참다 참석 창고 창밖 찾아가다 찾아내다 찾아보다 찾아오다 채널 채소 채우다 책임 책임자 챔피언 챙기다 철 첫 청년 청소기 청소년 체육 체중 초 초 초록색 초보 촬영 최고 최근 최대 최선 최소한 최초 추억 추위 축구공 축제 축하 출구 출근 출근 출발 출연 출입 출입문 출장 충격 충분 충분히 취소 취소 취직 취하다 취하다 치과 치료 치료 치우다 치즈 친절 친척 침실 카운터 카페 캠퍼스 커지다 커튼 코끼리 코너 콘서트 콩 크기 키우다 킬로 킬로그램 킬로미터 탑 태도 태양 태우다 태우다 태풍 터 터널 터미널 턱 털 털다 테스트 테이프 토끼 토론 토마토 톤 통 통 통신 통일 통장 통하다 퇴근 퇴근 특별 특별 특별히 특징 특히 틀다 틀리다 틀림없다 티셔츠 파 파도 파랗다 파리 파리 파일 판단 판매 판매 팔리다 패션 팩 팬 팬 펴다 편 편리 편안 편하다 펼쳐지다 평생 평소 평일 평화 포기 포도주 포스터 포장 포함 폭 표시 표정 표현 표현 푹 풀 풀 풀다 풀리다 풍경 프로 프로 프로그램 플라스틱 피 피다 피로 피로 피부 피시 필름 하루 하얀색 하얗다 하하 학기 학습 학원 한꺼번에 한동안 한두 한숨 한쪽 한참 한편 한편 할인 함부로 합격 합치다 해결 해결 해마다 해석 해외 해외여행 햇볕 햇빛 햇살 행동 행동 행복 행복 행사 향기 향수 향하다 허용 헌 혀 현관 현금 현대 현재 현재 현지 형제 호선 호수 홈페이지 화가 화나다 화분 화장품 확대 확실히 확인 확인 환경 환영 활동 활동 활용 회복 회색 회원 회장 회화 효과 후배 훨씬 휴가 흐르다 흐리다 흔들다 흔히 흘러가다 흘러나오다 흘리다 흙 흥미 희망

Appendix G.

Words in the corpus covered by level C vocabulary of KLL

가구 가난 가능 가라앉다 가령 가로등 가로막다 가로수 가르다 가만 가뭄 가사 가상 가입 가입자 가장 가정 가치 가치관 가톨릭 가하다 각기 간 간격 간부 간섭 간신히 간접 간판 간호 간혹 같이다 갈등 갈수록 감 감각 감다 감독 감소 감수성 감싸다 감옥 감추다 감히 갑 갑작스럽다 값싸다 강남 강도 강력 강변 강사 강수량 강요 강의 강조 강화 갖가지 갖추다 같이 갚다 개개인 개국 개다 개미 개방 개별 개선 개성 거꾸로 거대 거두다 거듭 거리 거부 거액 거절 거치다 거칠다 거품 건 건너 건네다 건설 건전 건조 걷기 걷다 걸치다 검사 검토 게 게시판 겨자 겪다 견해 결론 결석 결심 결코 경계 경고 경기 경력 경비 경영 경쟁 경쟁력 경제력 경제학 경향 계곡 계좌 계층 고개 고소 고요 고작 고전 고집 고함 곡 곡식 곤란 곧다 곧이어 곧장 골 골고루 골짜기 공개 공격 공공 공군 공급 공기 공식 공업 공연 공주 공중 공통 공통점 공포 공해 과 과거 과연 과외 과정 관계없이 관계자 관념 관람 관람객 관리 관습 관심사 관점 관찰 광경 광장 괴롭히다 교내 교대 교사 교시 교양 교외 교육비 교장 교재 교체 교환 교훈 구 구르다 구별 구분 구석 구석구석 구성 구속 구역 구입 구조 구청 국내외 국사 국산 국적 국회 군 군사 굳다 굳어지다 굽히다 권리 권위 귀가 귀신 귀중 귀하다 규모 규정 균형 귤 그간 그나마 그녀 그대 그래픽 그러다 그루 그야말로 그저 그토록 극 극복 극히 근거 근교 근래 근로 근본 근원 근육 긁다 금 금메달 금세 금액 급격히 급속히 급증 급히 긋다 기 기관 기구 기기 기념 기능 기다 기대다 기도 기둥 기록 기법 기성 기술 기술자 기업 기업인 기여 기울다 기울이다 기원 기원전 기적 기초 기타 기호 기혼 기획 기후 긴급 긴장감 김 깊숙이 까다 깔다 깨닫다 깨달음 깨뜨리다 깨어나다 깨우다 꺾다 꼬리 꼬마 꼭 꼴 꼽히다 꽂다 꽉 꾸다 꾸리다 끈 끊기다 끊임없다 끊임없이 끌리다 끌어당기다 끝없이 끼 끼어들다 끼우다 나뉘다 나란히 나르다 나름 나아가다 나침반 난방 날리다 남 남부 남북 납득 낭비 낮추다 낱말 내 내다보다 내려지다 내보내다 내부 내외 내용물 내적 내주다 내후년 냉동 냉방 널리 넘기다 넘어오다 노동 노선 녹다 녹이다 녹화 논 논리 논쟁 논하다 놀리다 놈 농민 농산물 농장 놔두다 뇌 눈길 눈동자 눈부시다 늘어놓다 늘어지다 다가서다 다듬다 다루다 다소 다지다 다짐 다투다 다툼 다행 단 단계 단독 단순 단위 단편 단풍 달다 달래다 달리 담 담당 담당자 답 답변 당당 당분간 당시 당하다 대 대가 대강 대개 대기 대다 대다수 대략 대량 대로 대륙 대비 대사 대상자 대신 대응 대입 대접 대책 대처 대체 대체로 대출 대표 대형 댐 더더욱 더불다 더욱 더욱더 덕 덜다 덩어리 덮이다 데우다 데이트 도 도달 도대체 도덕 도마 도망 도심 도입 도장 도저히 도전 독감 독립 독특 독하다 돌아다니다 동 동그라미 동그랗다 동기 동료 동부 동서 동서남북 동의 동일 동작 동행 동화책 되돌리다 되돌아가다 되돌아보다 되돌아오다 되찾다 되풀이 두께 두뇌 두드러지다 두려움 두르다 두어 둘러싸다 둘러싸이다 둥지 뒤늦다 뒤따르다 뒤지다 뒤편 드디어 드러나다 드물다 들이다 듯 등등 등록 등록증 등장 디스크 따로따로 딱 땅속 때 때로 때리다 떼 뚫다 뛰놀다 뛰어나다 뛰어넘다 뜨다 뜯다 뜰 띄다 띄우다 라운드 라인 라켓 레이저 레저 로봇 로터리 리 리그 리터 마당 마디 마라톤 마련 마무리 마음껏 마이크 마주치다 마찰 마크 막상 만 만점 만족 만하다 말기 말다 말리다 맘 맛보다 망설이다 망원경 망하다 맡다 매 매달 매달다 매달리다 매번 매장 매주 매체 맺다 머리말 머물다 먹다 먹이 멋지다 메우다 면 면담 면적 면접 면하다 명단 명령어 명예 명의 명칭 모 모금 모범 모색 모여들다 모집 모처럼 모퉁이 목록 목숨 몰려들다 몸짓 몸통 못 못지않다 묘사 무 무기 무너지다 무대 무덤 무려 무료 무리 무의미 무지개 묵다 묶이다 문구 문법 문서 문자 문화재 묻다 묻히다 물가 물결 물리학 물질 물체 뭘 미 미끄러지다 미끄럽다 미니 미루다 미만 미사일 미용실 미움 미처 미혼 민간 민속 민주 민주주의 밀리다 밀접 밑바닥 바 바구니 바깥쪽 바람 바로잡다 바이러스 바퀴 바탕 박 박다 박히다 반대편 반면 반발 반성 반영 반응 반죽 발걸음 발바닥 발전 발톱 발휘 밝혀내다 밤하늘 방면 방바닥 방송 방송사 방식 방안 방울 방지 방해 배 배다 배치 백색 백성 뱃사람 버티다 번개 번역 번지 벌 벌금 벌어지다 범위 범인 범죄 법률 법원 법칙 베다 벤치 벼 변경 변동 변신 별도 별명 병실 보고 보관 보너스 보도 보라색 보름 보리 보상 보수 보안 보완 보장 보전 보조 보존 보충 보호 복 본 본부 본사 본성 본인 본질 볼 봉사 부 부담 부대 부딪히다 부러지다 부문 부상 부서 부서지다 부위 부인 부재 부정 부지런히 부채 부품 부피 부회장 북부 분량 분리 분명 분석 분야 분주 분포 분홍색 불과 불러일으키다 불리 불리다 불법 불어오다 불완전 불이익 불편 불평 불행 불확실 붓다 붙들다 붙잡다 브랜드 비 비극 비난 비다 비둘기 비로소 비만 비명 비바람 비비다 비상 비율 비중 비추다 비치다 비판 비하다 비행 비행장 빗방울 빛깔 빛나다 빠뜨리다 빠져나가다 빨다 빼놓다 빼앗기다 빼앗다 뺏다 뻔하다 사건 사고 사과 사기 사냥 사들이다 사례 사립 사망 사무소 사방 사상 사생활 사설 사소 사실상 사업가 사업자 사전 사정 사춘기 사회생활 사회주의 사회학 산업 살림 살아가다 살아남다 살아오다 살인 살피다 삶 삼국 삼다 상 상관 상관없다 상관없이 상담 상당 상당수 상당히 상대편 상류 상반기 상상력 상식 상업 상인 상점 상징 상태 상표 상황 새 새기다 새다 새로이 새롭다 생 생기 생명 생물 생산 생산력 생산자 생활비 생활수준 서구 서명 서민 서부 서서히 서적 석

석사 선 선거 선명 선언 선원 선장 선전 선정 선진 선진국 선호 설득 설립 설문 설사 설치 성 성경 성당 성립 성명 성숙
성실 성장 성적 성질 세 세계관 세다 세대 세미나 세월 세제 세트 셈 소규모 소득 소망 소매 소비 소속 소수 소용 소원
소위 소유 소유자 소음 소재 소중 소중히 소지품 소질 소프트웨어 소형 소홀히 소화 속담 속마음 속이다 손길 손쉽다 손실
손질 손해 솜 솜씨 솟다 쇠 수 수고 수단 수도 수도꼭지 수리 수만 수면 수명 수백 수석 수시로 수십 수업 수없이 수요
수입품 수저 수준 수집 수천 수컷 수행 순수 순식간 순위 순하다 숨기다 숨지다 스위치 스튜디오 습기 승객 승리 승부 승진
시 시각 시기 시나리오 시도 시멘트 시선 시스템 시야 시위 시일 시장 시점 시중 식기 식량 식료품 식생활 식욕 식히다 신
신경 신고 신규 신기 신념 신부 신분 신비 신사 신설 신세 신속 신인 신중 신청 신청서 신화 실 실감 실로 실리다 실습
실시 실은 실장 실정 실제 실제로 실천 실체 실태 실험 실현 심사 심장 심정 심판 싹 쌍 쌍둥이 썩 쓰러지다 쓰이다 씌우다
씨앗 아끼다 아니 아르바이트 아무런 아쉬움 아쉽다 아스팔트 아예 아울러 아이디어 안 안개 안내 안부 안심 안정 안팎
알루미늄 알아내다 알아보다 알코올 암시 암컷 압력 앞날 앞두다 앞바다 앞서 앞장서다 애 애정 애초 액수 앨범 야간 약
약수 약점 약품 양 양국 양력 양보 양식 양주 얘 어기다 어찌 언덕 언론 언어 얼마간 엄격 엄청나다 업다 업종 업체 엇갈리다
엉망 에 엔진 여가 여건 여기다 여럿 여론 여왕 여전 여전히 역 연간 연결 연관 연구실 연구원 연기 연기자 연두색 연락
연령 연상 연속 연예인 연인 연장 연주 연출 연합 열기 열매 열정 염려 영 영남 영상 영양 영업 영역 영웅 영향 영향력
영화관 영화배우 예 예고 예방 예보 예비 예산 예상 예선 예약 예외 예의 예측 예컨대 오락 오로지 오르내리다 오염 오페라
오해 옥상 온갖 온돌 온라인 온종일 올 올리다 올바르다 완성 완전 외과 외다 외로움 외면 외부 요 요구 요리사 요약 요청
욕 용 용기 용도 용어 우려 우정 우주 우편 운명 운반 운영 운행 울리다 움직임 워낙 원 원고 원인 원장 웬 위 위기 위로
위반 위법 위성 위원 위원장 위주 위협 유능 유리 유머 유물 유발 유사 유산 유의 유적 유적지 유지 유형 육군 육상 육체
은 음력 음성 음식물 음식점 음주 응답 의도 의류 의무 의문 의복 의사 의식 의심 의외로 의원 의존 의지 의학 이 이끌다
이내 이념 이달 이대로 이데올로기 이동 이따금 이래 이러다 이롭다 이르다 이민 이성 이슬 이어 이어서 이어지다 이용자
이자 이제야 이중 이해 이해관계 익다 익히다 인 인격 인공 인근 인도 인류 인물 인사 인상 인쇄 인식 인연 인재 인정
인정받다 인제 인종 인체 인하 인하다 일단 일대 일반인 일생 일식 일일이 일자 일자리 일정 일종 일찍이 일체 일치 임금
임무 임신 임신부 입국 입대 입력 입사 입시 입장 입히다 잇다 잇따르다 자 자격증 자극 자녀 자랑 자리 자부심 자살 자세
자신감 자연환경 자연히 자원 자율 자정 자취 자판 작성 작업 작용 작품 잔디 잠그다 잠기다 잠수함 잡아당기다 잡아먹다
장관 장기간 장난 장르 장면 장모 장비 장수 장식 장애인 장인 장차 잦다 재능 재다 재밌다 재빨리 재생 재수 재정 재판
재학 재활용품 저기 저마다 저울 저지르다 적성 적용 적응 적합 전 전개 전공 전구 전기 전달 전망 전문점 전문직 전반
전선 전설 전세 전시 전시장 전시회 전용 전환 전후 절 절대 절약 절차 점검 접근 접다 접촉 접하다 정기 정당 정면 정반대
정부 정비 정상 정성 정신과 정오 정장 정지 정직 젖 제거 제공 제과점 제대 제사 제시 제안 제약 제외 제의 제자 제자리
제작 제출 제한 조 조각 조기 조명 조미료 조절 조정 조직 존재 졸리다 졸음 좁히다 종 종종 종합 좌우 죄 주 주거 주고받다
주년 주름 주문 주민 주방 주식 주어지다 주요 주의 주일 주장 죽 줄거리 줄어들다 중단 중대 중독 중반 중부 중세 중소기업
중순 중식 중심지 중앙 중요시 즉석 증가 증거 증권 증명 지각 지경 지극히 지금껏 지급 지나치다 지능 지니다 지다 지대
지도 지루 지방 지배 지불 지시 지식 지우다 지원 지위 지적 지점 지지 지진 지출 지켜보다 지폐 지혜 직선 직장인 직전
직후 진단 진동 진로 진료 진리 진실 진실로 진지 진찰 진출 진통 진행 질병 짐작 집단 집안일 집어넣다 집중 짚다 짜다 짝
쪽 쫓다 쭉 찌꺼기 찍히다 찢어지다 차 차다 차라리 차량 차별 차선 차이점 차차 차츰 착각 찬성 참 참가 참고 참새 참석
참석자 참여 참으로 참조 창 창구 창작 창조 찾아다니다 채 채점 책임감 책임지다 처리 처벌 처지 척 척하다 천 천둥 천장
천재 철 철도 철저 철저히 철학 철학자 체력 체온 체육관 체하다 체험 초기 초대 초등학생 초반 초상화 초순 초여름 초점
총 총리 최대한 최상 최소 최신 최악 최저 최종 최후 추가 추석 추진 추천 추측 축구장 축소 출국 출발점 출산 출석 출신
출연 출입국 출퇴근 출판 출판사 출현 충고 충돌 취업 취재 취향 치다 치료법 치르다 치아 치우다 침 침묵 침착 칸 캐릭터
캠페인 컨디션 컬러 코드 코스 코트 클럽 키스 타다 타입 타자기 탁 탁자 탄생 탈출 탓 택하다 터 터지다 턱 텅 테러 텍스트
토대 토론 토론자 토하다 통계 통과 통로 통역 통일 통제 통증 통합 통화 투명 투자 투표 튀김 튀다 트럭 트이다 특급 특성
특수 특이 특정 틀 틈 파괴 파다 파악 판 판결 판단 판사 팩스 팩시밀리 퍼지다 페인트 편 편견 편의 평 평가 평균 평상시
폐지 포인트 포크 포함 폭넓다 폭력 표 표면 표시 표준 푸다 품 품다 품목 품질 풍속 프린터 피곤 피하다 피해 피해자 필수
필자 핑계 하 하나님 하나하나 하도 하드웨어 하룻밤 하반기 하천 학과 학급 학력 학번 학부모 학술 학위 학자 학점 한
한가 한결 한계 한눈 한때 한마디 한순간 한식 한정 한창 한층 한하다 함께 합격 합하다 항공 항공기 항구 항의 해 해군
해내다 해답 해당 해롭다 해물 해석 해설 해소 해수욕장 해안 핵 핵심 햄 행사 행운 행위 행하다 향 향상 허가 허락 허용
헤아리다 헬기 현 현관문 현대인 현상 현실 현장 혈액 협력 형 형사 형성 형수 형식 형태 형편 혜택 호기심 호남 호박 호실
호흡 혹시 홀로 홍보 홍수 화면 화살 화장 화장지 화재 화제 화학 확 확립 확보 확신 확장 확정 환영 환율 환하다 활기

활발히 활용 회 회관 회복 회의 회전 횟수 후기 후반 후보 훈련 훔치다 휴식 흉내 흐름 흐리다 흑백 흔들리다 흔적 흔하다 흘러내리다 흥분 희곡 희망 희생 힘겹다 힘쓰다 힘차다

Appendix H.

The Korean Engineering Academic Vocabulary List (intermediate and advanced levels in alphabetical order)

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
1	가상	假想	가상적 가상하다	가상 계측, 가상 경로점	imagination, virtuality; (-적) imaginary; (-하다) to imagine
2	가치	價值		가치 사슬	value, worth
3	가하다 02	加--			to give, to inflict, to deliver, to deal, to apply
4	간격 02	間隔		시간 간격	interval, gap, space, distance
5	간섭	干涉	간섭하다		interference; (-하다) to interfere (in), to meddle (in), to intrude (into/on)
6	간접	間接	간접적	간접 검사	indirectness; (-적) indirect
7	감소 02	減少	감소하다	급격히 감소하다	decrease; (-하다) to decrease, to decline
8	강화 04	强化	강화하다		reinforcement; (-하다) to reinforce, to strengthen
9	갖추다	(native)			to prepare, to get ready
10	개방 03	開放	개방적 개방하다		(-적) open; (-하다) to open, to open up
11	개별	個別	개별적		individual; (-적) individual, being individual
12	개선 01	改善	개선적 개선하다 개선되다	성능 개선, 환경 개선	improvement (in/on/to); (-적) ameliorative; (-하다) to improve, to reform;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
					(-되다) be improved
13	건 03	件			case of, matter
14	건조 04	乾燥	건조하다	건조 수축, 건조 개시	(-하다) dry
15	검토	檢討	검토하다 검토되다		examination, review, consideration; (-하다) to examine, to review, to check; (-되다) be reviewed, be examined
16	경계 04	境界		경계 조건	boundary, border
17	경쟁	競爭	경쟁적 경쟁하다		competition, contest (for); (-적) competitive; (-하다) to compete
18	경쟁력	競爭力			competitiveness
19	경향 02	傾向			tendency, trend
20	계층	階層	계층적	물리 계층	class, stratum (strata); (-적) hierarchical
21	공격	攻擊			attack, assault, aggression; (-하다) to attack
22	공공 02	公共			public, community
23	공급 02	供給		공급 체인, 공급 능력	supply, provision; (-하다) to supply, to provide
24	공식 01	公式			formula
25	공통	共通	공통하다 공통되다		commonness; (-하다) be common; (-되다) be common
26	과정 03	過程		교육 과정, 처리 과정, 설계 과정	process

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
27	관계없이	關係없이			regardless of
28	관점 02	觀點			point of view, viewpoint, standpoint, perspective, angle
29	구분 06	區分			division, separation, classification; (-하다) to divide, to separate; (-되다) be divided, be sorted
30	구성 06	構成		구성 요소, 구성 성분, 시스템 구성	composition, organization, constitution; (-하다) to form, to construct; (-되다) be formed, be organized
31	구속 02	拘束			restriction, restraint; (-하다) to restrict, to restrain
32	구역 04	區域		구역 전기	area, zone
33	구조 08	構造	구조적	외피 구조, 바닥 구조, 구조 부재, 콘크리트 구조, 네트워크 구조, 구조 설계, 구조 해석, 구조 시스템	structure, organization, constitution; (-적) being structural
34	규모	規模		시장 규모	size, scale
35	규정 04	規定			rule, regulation, code; (-하다) to prescribe, to stipulate
36	균형	均衡			balance

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
37	극복 01	克服	극복하다 극복되다		conquest, subjugation; (-하다) overcome, conquer; (-되다) be overcome
38	근거	根據	근거하다		grounds, basis; (-하다) be based on
39	급격히	急激히		급격히 감소, 급격히 증가	rapidly, sharply
40	기관 11	機關		내연 기관, 금융 기관, 훈련 기관, 연구 기관	engine; organization, institution, agency
41	기구 14	器具			apparatus, instrument
42	기기 12	器機		의료 기기	equipment, instrument, appliance
43	기능 03	機能		차단 기능, 제어 기능, 통신 기능	function
44	기둥	(native)		기둥 접합부	pillar, column
45	기법 01	技法		인증 기법, 해석 기법, 예측 기법, 분석 기법	technique
46	기업	企業	기업적		business, enterprise, company; (-적) corporate
47	기여 02	寄與	기여하다		contribution; (-하다) to contribute
48	기후 05	氣候		기후 변화	climate
49	낮추다	(native)			to drop, to reduce, to bring down
50	내 07	內			inside, within, within the scope

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
51	내부 04	內部			the inside, interior, inner part
52	널리	(native)		널리 사용, 널리 이용	wide(ly), extensively, far and wide, generally
53	노동 02	勞動	노동하다	노동 시장	labor, work; (-하다) to work
54	다루다	(native)			handle, deal with
55	다소 01	多少			somewhat, a little, a (little) bit, to a certain degree
56	단계 02	段階	단계적	초기 단계, 설계 단계	stage, phase, step; (-적) phased
57	단독 02	單獨	단독적		independence, singleness, separateness; (-적) independently
58	단위 02	單位		단위 중량, 단위 면적, 능력 단위, 평가 단위, 단위 시간	unit, measure
59	달리	(native)			differently, dissimilarly, in a different way, in some other way
60	당시 02	當時			then, at that time
61	대가 03	大家		대가 표시	authority (on), expert
62	대기 04	大氣		대기 오염, 대기 온도, 대기 환경	the atmosphere, the air
63	대략	大略	대략적		summary, outline; approximate; (-적) rough, approximate

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
64	대량 01	大量		대량 생산	large quantity, enormous volume, large amount
65	대비 09	對備	대비하다		preparation; (-하다) respond, correspond
66	대응 02	對應	대응하다		action, maneuver, correspondence; (-하다) deal with, respond (to), correspond to
67	대입 02	代入	대입하다 대입되다		substitution; (-하다) to substitute; (-되다) be substituted
68	대책 03	對策			measure(s), step(s)
69	대체 03	代替	대체하다 대체되다	대체 원가, 대체 자료	substitution, replacement; (-하다) to substitute, to replace
70	대체로	大體—			generally, mostly, overall
71	대형 03	大形		대형 항공사	large[full] size
72	더불다	(native)			to join in (a thing), to do together, to partake of
73	더욱	(native)			more, increasingly, more and more
74	도달 01	到達	도달하다 도달되다		arrival, reaching, attainment; (-하다) to arrive, to reach; (-되다) to get to
75	도입	導入	도입하다 도입되다		introduction; (-하다) to introduce; (-되다) to be introduced
76	독립	獨立	독립하다 독립되다	독립 변수	independence; (-하다) to be independent;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
					(-되다) to be independent
77	동일 02	同一	동일하다		same (as), identical (to/with), equal (to); (-하다) to be identical
78	두께	(native)		피복 두께	thickness
79	등장 01	登場	등장하다		appearance; (-하다) to appear
80	라인	line			line
81	로봇	robot		모바일 로봇	robot
82	마이크	microphone			microphone
83	면적	面積		단위 면적	(surface) area, extent
84	문서	文書			document, papers
85	물체	物體		전경 물체	object
86	미만 01	未滿			under, below
87	바탕	(native)			foundation (of), basis (of/for); background
88	반면 02/03	反面/半面			while, whereas, on the other hand; half surface
89	반영 01	反映	반영하다 반영되다		reflection; (-하다) to reflect; (-되다) to be reflected
90	방식 01	方式		변조 방식, 전송 방식, 통신 방식, 신호 방식, 생산 방식	way, means, method
91	방안 01	方案		해결 방안	way, measure
92	방지 03	防止	방지하다 방지되다	공회전 방지, 방지 시스템	prevention; (-하다) to prevent; (-되다) to be prevented
93	배치 02	配置	배치하다		arrangement, placement, disposition;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
					(-하다) to arrange, to place
94	범위	範圍		적용 범위	scope, range, scale
95	법칙	法則	법칙적		law, rule; (-적) lawful
96	변경 04	變更	변경하다 변경되다		change; (-하다) to change; (-되다) to be changed
97	변동	變動	변동하다 변동되다		change, fluctuation; (-하다) to change, to fluctuate; (-되다) to be changed, to fluctuate
98	별도 01	別途			a separate way, separate use
99	보상 02	補償	보상적 보상하다 보상되다		compensation; (-적) compensatory; (-하다) to compensate; (-되다) to be compensated
100	보수 10	補修	보수하다	유지 보수	repair; (-하다) to repair, to fix
101	보완	補完	보완적 보완하다 보완되다		supplementation; (-적) supplementary, complementary; (-하다) to supplement; (-되다) to be supplemented
102	보장 01	保障	보장하다 보장되다		guarantee (of/that); (-하다) to guarantee; (-되다) to be guaranteed
103	보조 02	補助	보조적 보조하다	보조 기억	help, assistance; (-적) supplementary; (-하다) to help, to assist;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
104	보존	保存	보존적 보존하다 보존되다		preservation, conservation; (-적) preservational; (-하다) to preserve; (-되다) to be preserved
105	부담 01	負擔	부담하다 부담되다		burden, load; (-하다) to bear; (-되다) to be burdensome
106	부문 06	部門			field, sector
107	부위 04	部位			part, region
108	부재 04	部材		구조 부재	member, part
109	부품	部品			part, component
110	부피	(native)			volume, bulk
111	분리 04	分離		원심 분리, 분리 세균, 분리 공정	separation (from), division; (-하다) to separate, to isolate; (-되다) to be separated
112	분야	分野		응용 분야, 공학 분야, 산업 분야, 기술 분야	area, field, sphere
113	분포	分布		정규 분포, 온도 분포, 압력 분포, 속도 분포	distribution, dispersion
114	비율 02	比率		중첩 비율	ratio, proportion, rate
115	비중 01	比重			importance, weight
116	비하다	比--			compare
117	사건	事件	사건적		incident, affair, event; (-적) episodic

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
118	사례 05	事例		사례 연구, 사례 분석	instance, case
119	사업자	事業者		후발 사업자, 전기 사업자	businessperson, business operator
120	산업	産業	산업적	산업 공생, 녹색 산업, 산업 공학, 산업 현장, 산업 분야, 산업 제어	industry; (-적) industrial
121	삶	(native)			life
122	상관 03	相關	상관적 상관하다 상관되다	상관 관계, 상관 계수	correlation; (-적) correlative; (-하다) to interrelate, to care (about); (-되다) to be interrelated
123	상당 04	相當	상당하다	상당 부분	(-하다) to be considerable, sizeable, respectable
124	상당히	相當히			considerably, quite, rather, fairly
125	상업 02	商業	상업적	상업 영화	commerce, business; (-적) commercial
126	상태 01	狀態		상태 전이, 정상 상태, 상태 변수, 상태 정보	condition, state
127	상황 02	狀況		교통 상황	situation, conditions, circumstances, state of affairs
128	생명	生命			life, bio-
129	생산	生産	생산적	대량 생산,	production;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
			생산하다 생산되다	생산 경영, 생산 공정, 생산 계획, 생산 방식	(-적) productive; (-하다) to produce, to manufacture; (-되다) to be produced
130	선 13	線			line
131	선정 07	選定	선정하다 선정되다	공법 선정	selection, choice; (-하다) to select, to choose; (-되다) to be selected
132	선호 03	選好	선호적 선호하다		preference; (-적) preferential; (-하다) to prefer; (-되다) to be preferred
133	설문 01	說問		설문 조사	survey
134	설치 02	設置	설치하다 설치되다		installation; (-하다) to install, to equip, to fit; (-되다) to be installed
135	성장 01	成長	성장하다 성장되다		growth, development; (-하다) to grow, to develop; (-되다) to grow, to develop
136	성질	性質			properties, nature, quality, character
137	세대 01	世代			generation
138	소비 03	消費	소비하다 소비되다		consumption, consume, spend; (-하다) to consume; (-되다) to be consumed
139	소형 03	小型			small size, pocket-size
140	손실	損失	손실하다 손실되다	압력 손실, 손실 비용	loss; (-하다) to lost;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
					(-되다) to be lost
141	수단 01	手段			means, way, measure, method
142	수리	修理		수리 모형	repair(s), repair; (-하다) to repair; (-되다) to be lost
143	수면 01	水面		자유 수면	surface (of the water)
144	수명 05	壽命			life, life expectancy
145	수요 06	需要		충전 수요	demand
146	수준	水準		수준 성숙, 일정 수준, 서비스 수준	level, standard
147	수집 02	蒐集	수집하다 수집되다	정보 수집	collection; (-하다) to collect, to gather; (-되다) to be collected
148	수행 02	遂行	수행적 수행하다 수행되다	수행 능력	fulfillment, performance; (-적) executive; (-하다) to execute; (-되다) to be executed
149	순위	順位		우선 순위	ranking
150	승객	乘客			passenger
151	시각 05	視覺	시각적		sight, vision; (-적) visual, optical
152	시기 04	時期	시기적	수선 시기	time; (-적) timely
153	시나리오	scenario			scenario
154	시멘트	cement		시멘트 모르타르, 폴리머 시멘트,	cement

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
				시멘트 페이스트	
155	시스템	system		결재 시스템, 모니터링 시스템, 관리 시스템	system
156	시점 02	時點			time
157	신속 03	迅速			quickness, rapidity; (-하다) to be quick
158	실시 03	實施	실시하다 실시되다		implementation, enforcement; (-하다) to execute, to implement; (-되다) to be executed, to be enforced
159	실정 03	實情			(actual/real) state
160	실제로	實際			really, actually, in reality, in effect, in practice
161	실험	實驗		실험 결과, 실험 장치, 실험 데이터, 실험 방법	experiment, test
162	실현	實現	실현하다 실현되다		realization; (-하다) to realize, to materialize; (-되다) to be materialized
163	안정 01	安定	안정적 안정하다		stability, stabilization; (-적) stable; (-하다) be stable
164	알루미늄	aluminum			aluminum

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
165	알아보다	(native)			to investigate, to check, to search, to look into, to inquire
166	앞서	(native)		앞서 언급, 앞서 설명	before, earlier, previously
167	업체	業體		제조 업체	business, (business) enterprise
168	연간 02	年間			during the course of a year, for a year
169	연관 06	聯關	연관하다 연관되다		relation(s) (with/between), connection (with/between/to); (-하다) to connect; (-되다) to be related, to be connected
170	연구원	研究員			researcher
171	연령 02	年齡	연령적		age; (-적) age-concerned
172	연속 01	連續	연속적	연속 지적도	continuity, continuation; (-적) continuously
173	영역 03	領域		주파수 영역, 계산 영역, 시간 영역	field, area, domain
174	영향 04	影響		정의 영향	influence (on), effect (on), impact (on)
175	영향력	影響力			influence, leverage
176	예방 01	豫防	예방적 예방하다 예방되다		prevention; (-적) preventive; (-하다) prevent, stave off; (-되다) to be prevented
177	예비 02	豫備	예비적 예비하다	예비 교사	reserve; (-적) preliminary; (-하다) to reserve, to prepare

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
178	예측	豫測	예측하다 예측되다	예측 제어기, 예측 기법	prediction, forecast; (-하다) to predict; (-되다) to be predicted
179	오염	汚染		대기 오염, 오염 물질, 수질 오염, 토양 오염, 오염 배출, 환경 오염	pollution, contamination
180	온라인	on-line		온라인 서비스	on-line
181	완성 01	完成	완성적 완성하다 완성되다		completion; (-적) completive; (-하다) to complete; (-되다) to be completed
182	외부 02	外部		외부 표면, 외부 환경	the outside
183	요약 02	要約	요약적 요약하다 요약되다		summing-up; (-적) recapitulative; (-하다) to summarize, to sum up; (-되다) to be summarized
184	용기 03	容器			container
185	용도 02	用途			use
186	용어 02	用語			term, terminology
187	우려	憂慮	우려하다 우려되다		concern; (-하다) to be concerned, to worry; (-되다) to be concerned
188	운반 02	運搬	운반하다 운반되다		transportation; (-하다) to transport;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
					(-되다) to be transported
189	운영 03	運營	운영적 운영하다 운영되다	운영 체제	(-적) operational; (-하다) to manage, to run, to operate; (-되다) to be managed, to be operated
190	운행 02	運行	운행하다 운행되다		operation, service (transport); (-하다) to operate, to move (about transport); (-되다) to be operated
191	원인 02	原因		교체 원인, 주요 원인, 이상 원인, 원인 분석, 발생 원인	cause, reason
192	위협	威脅	위협적 위협하다		threat (of), menace (to); (-적) threatening; (-하다) to threat, intimidate;
193	유발 03	誘發	유발적 유발하다 유발되다		induction, induce, cause; (-적) induced; (-하다) to cause, to induce; (-되다) to be caused, to be induced
194	유사 18	類似	유사하다	유사 과제, 매우 유사	similar, alike; (-하다) to be similar
195	유의 04	留意	유의하다		(-하다) to pay attention
196	유지 08	維持	유지하다 유지되다	재고 유지, 유지 보수, 유지 비용, 유지 관리	(-하다) to keep, to maintain; (-되다) to be maintained

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
197	유형 07	類型	유형적	성격 유형, 유형 분류, 결정 유형	type, category, class, pattern; (-적) being categorical
198	육상 02	陸上			(on) land, (on the) ground
199	음성 02	音聲		음성 인식, 음성 신호	voice, audio
200	응답	應答	응답하다	진동 응답, 주파수 응답	answer, response, reply; (-하다) to answer
201	의존	依存	의존적 의존하다 의존되다		dependence (up) on, reliance (up) on; (-적) being dependent, being reliant; (-하다) to depend on, rely on; (-되다) to depend on, rely on
202	이내 05	以內			within, in less than
203	이중 03	二重		이중 외피	duplication, doubleness
204	인공 01	人工	인공적	인공 신경망	artificiality; (-적) artificial
205	인류 01	人類			humankind, man
206	인쇄	印刷	인쇄하다 인쇄되다		printing; (-하다) to print; (-되다) to be printed
207	인식	認識	인식적 인식하다 인식되다	화자 인식, 음성 인식	awareness, realization, cognition; (-적) cognitive; (-하다) to recognize, to be aware of; (-되다) to be perceived
208	인체	人體			human body, human organism

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
209	인하다	因—			to be caused by, to arise from, to result from, to be due to
210	일치 01	一致	일치하다 일치되다		agreement, accord, consensus; (-하다) to accord; (-되다) to be in accord with
211	입력	入力	입력하다 입력되다	입력 전압, 입력 신호, 입력 변수, 입력 자료, 입력 장치, 입력 데이터	entry; (-하다) to enter, to input; (-되다) to be input
212	자극 01	刺戟		자극 대상	stimulation, irritation; (-하다) to stimulate, to irritate; (-되다) to be stimulated
213	자세 02	姿勢		자세 변화	posture, position, attitude
214	자원 04	資源		자원 개발	natural resources
215	자율	自律	자율적	자율 실습	autonomy; (-적) being autonomous
216	작성 01	作成	작성하다 작성되다		(-하다) to make (out), to draw up, to prepare; (-되다) to be written, to be drawn up
217	작업	作業	작업하다 작업되다	작업 시간	work, job, operations; (-하다) to work; (-되다) to work
218	작용 01	作用	작용하다 작용되다	상호 작용	action, effect; (-하다) to act (on), to work (on); (-되다) to work, to function

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
219	장비 04	裝備			equipment, gear, apparatus
220	재생 01	再生	재생하다 재생되다		regeneration, recycling; (-하다) to recycle, to reclaim; (-되다) to be recycled, to be restructured
221	적응 02	適應	적응적 적응하다 적응되다		adaptation; (-적) adaptive; (-하다) to adapt, to adjust; (-되다) to be adapted, to be adjusted
222	전개 02	展開	전개하다 전개되다		development; (-하다) to develop, to unfold; (-되다) to be unfolded
223	전망 03	展望	전망하다 전망되다		prospect, forecast; (-하다) to predict, to foresee, to forecast; (-되다) to be predicted
224	전반 02	全般	전반적		whole, general; (-적) overall
225	전용 04	專用	전용하다 전용되다		exclusive use, private use; (-하다) to use exclusively; (-되다) to be used exclusively
226	전환 03	轉換		방향 전환	change, switch, transition; (-하다) to change, to switch; (-되다) to be changed
227	전후 01	前後			the front and back

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
228	절차 02	節次	절차적		procedure (for), process (of); (-적) procedural
229	접근	接近	접근하다	접근 방법	approach; (-하다) to approach
230	접촉	接觸	접촉하다	접촉 저항	contact, touch; (-하다) to touch; to have[make] contact (with)
231	정부 08	政府	정부적		government, administration; (-적) governmental
232	정지 05	停止	정지하다 정지되다		stop, halt, standstill; (-하다) to stop, to suspend; (-되다) to be suspended
233	제거 03	除去	제거하다 제거되다		removal, elimination; (-하다) to remove, to eliminate, to get rid of; (-되다) to be removed
234	제시 02	提示	제시하다 제시되다		suggestion, presentation; (-하다) to suggest, to propose, to present; (-되다) to be presented
235	제안 02	提案	제안하다 제안되다		suggestion, offer, proposal; (-하다) to suggest, to propose; (-되다) to be suggested
236	제약 01	制約	제약적 제약하다 제약되다	제약 조건	restriction (on), limitation (on/upon), constraint (on); (-적) constrained; (-하다) to restrict, to limit;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
					(-되다) to be limited
237	제외 01	制外	제외하다		exemption; (-하다) to except (from); to make an exception
238	제작 02	製作	제작하다 제작되다		production, manufacturing; (-하다) to make, produce; (-되다) to be produced
239	조절 02	調節	조절적 조절하다 조절되다		control; (-적) controlled; (-하다) to control, to regulate; (-되다) to be controlled
240	조정 09	調整	조정하다 조정되다		adjustment, modification, revision; (-하다) to adjust, to change; (-되다) to be adjusted, to be changed
241	조직	組織	조직적 조직하다 조직되다	조직 몰입, 조직 유효	group, organization, syndicate; (-적) organized; (-하다) to organize; (-되다) to be organized
242	존재	存在	존재적 존재하다		existence; (-적) existential; (-하다) to exist;
243	좌우 01	左右	좌우하다 좌우되다		left and right; influence; (-하다) to influence, to affect; to dominate; (-되다) to be influenced; to be dominated

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
244	주민	住民		지역 주민	(local) resident, inhabitant
245	주의 07	注意	주의하다		care, caution; (-하다) to care, to be careful
246	줄어들다	(native)			to decrease
247	중단 04	中斷	중단하다 중단되다		interruption, suspension (of), halt; (-하다) to halt (-되다) to be halted
248	중앙	中央		중앙 기어, 중앙 처리	center, the middle
249	증가 01	增加	증가하다 증가되다 증가시키다	급격히 증가하다	increase, growth, rise; (-하다) to increase, to grow (-되다) to be increased; (-시키다) to increase, to enhance
250	지니다	(native)			keep, carry
251	지배 01	支配	지배적 지배하다 지배되다	지배 방정식	control, domination; (-적) dominant; (-하다) to control, to dominate; (-되다) to be dominated
252	지식 02	知識			knowledge, understanding
253	지원 02	支援	지원하다	지원 체계	support, backing, aid; (-하다) to support, to back (up)
254	지지 05	支持	지지하다		support, backing; (-하다) to support,
255	직선 01	直線	직선적		straight line; (-적) straightforward
256	진단 02	診斷	진단하다 진단되다		diagnosis, prognosis;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
					(-하다) to diagnose, examine; (-되다) to be diagnosed, to be examined
257	진동 03	振動	진동하다	진동 응답	vibration, vibrate; (-하다) to vibrate
258	집단	集團	집단적		group, body (-적) being collective
259	차별	差別	차별적 차별하다 차별되다		differentiation; (-적) differential; (-하다) to distinguish; (-되다) to be distinguished
260	참고 01	參考	참고적 참고하다 참고되다	참고 문헌	reference; (-적) referential; (-하다) to refer (to), to consult; (-되다) to be referred to
261	참여	參與	참여하다	참여 설계	participation; (-하다) to participate (in)
262	참조 02	參照	참조하다		reference, consultation; (-하다) to refer (to), to consult;
263	처리	處理	처리하다 처리되다	형수 처리, 하수 처리, 폐수 처리, 중앙 처리, 처리 시설, 처리 장치, 신호 처리, 처리 과정, 처리	handling, disposal, processing; (-하다) to handle, to process

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
				공정, 정보 처리, 데이터 처리, 처리 속도, 처리 시간, 처리 방법, 처리 기술	
264	철도	鐵道			railway
265	초기 02	初期		초기 투자, 초기 단계, 초기 위치	beginning, the early part (of)
266	초점 02	焦點			focus, point
267	최소	最小	최소하다		minimum; (-하다) minimize
268	최종	最終	최종적		last, final; (-적) last, final
269	추가 02	追加	추가적 추가하다 추가되다		addition, supplement; (-적) additional; (-하다) to add; (-되다) to be added
270	추진 02	推進	추진하다 추진되다		(-하다) to propel, to push ahead; (-되다) to be propelled, to be driven
271	축소	縮小	축소하다 축소되다		reduction; (-하다) to reduce; (-되다) to be reduced
272	충돌	衝突	충돌하다 충돌되다		collision (with), impact, crash; (-하다) to crash; (-되다) to be collided, to be crashed
273	코드	code			code
274	토대	土臺	토대하다		foundation, substructure;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
					(-하다) be based on;
275	통계 04	統計	통계적 통계하다		statistics; (-적) statistical; (-하다) to compile statistics;
276	통과	通過	통과하다 통과되다		(-하다) to pass, to skip, to be approved; (-되다) to be passed, to be approved
277	통제 02	統制	통제적 통제하다 통제되다		control, regulation; (-적) controlled; (-하다) to control, to regulate; (-되다) to be controlled
278	통합	統合	통합적 통합하다 통합되다	통합 총비용, 선박 통합	combination, integration; (-적) integrative; (-하다) to integrate, to consolidate; (-되다) to be integrated
279	투자 02	投資	투자하다 투자되다	확장 투자, 초기 투자, 투자 비용	investment; (-하다) to invest; (-되다) to be invested
280	특성 01	特性		절연 특성, 특성 곡선, 거동 특성, 유동 특성, 제어 특성	characteristic, nature, quality, attribute
281	특수 02	特殊	특수하다		(-하다) unusual, unique
282	특이	特異	특이하다		(-하다) unusual, unique
283	특정	特定	특정하다 특정되다		(-하다) particular, specific, certain; (-되다) to be specified

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
284	파괴	破壞	파괴적 파괴하다 파괴되다		destruction, demolition; (-적) destructive; (-하다) to destroy, to demolish; (-되다) to be destroyed
285	파악 01	把握	파악하다 파악되다		grasp, understanding; (-하다) to grasp, to understand; (-되다) to understand, to figure out
286	판 06	板			plank, board, record, disk, disc
287	편의 02	便宜	편의적		convenience; (-적) convenient
288	평가	評價	평가하다 평가되다	성능 평가, 평가 항목, 평가 단위, 평가 기준, 평가 대상, 평가 방법, 평가 결과, 평가 시스템	assessment, evaluation; (-하다) to assess, to evaluate; (-되다) to be assessed
289	평균	平均	평균적 평균하다	시간 평균	average, mean; (-적) average; (-하다) to take an average
290	표면	表面	표면적	표면이 거치다, 외부 표면	surface; (-적) superficial, external
291	표준 01	標準	표준적	국제 표준, 기술 표준	standard, average, norm; (-적) standard

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
292	품질 03	品質		품질 경영, 품질 관리, 서비스 품질, 품질 향상, 품질 모형, 품질 비용, 개발 품질	quality
293	풍속 02	風速			wind speed
294	피해	被害	피해하다 피해되다	인명 피해	harm, damage; (-하다) to harm, to damage; (-되다) to be damaged
295	필수 02	必須	필수적		necessity; (-적) essential, necessary
296	하천 02	河川		하천 제방	river, stream, brook
297	한계	限界			limit, limitation(s)
298	한정	限定	한정적 한정하다 한정되다		limit; (-적) limited; (-하다) to limit, to restrict; (-되다) to be limited
299	항공	航空			aviation, flight
300	해당 05	該當	해당하다 해당되다		(-하다) be relevant, be applicable (to), apply (to), correspond (to); (-되다) correspond (to)
301	핵심	核心	핵심적		core; (-적) core, essential, key
302	행위	行爲			act, action
303	향상 01	向上	향상하다 향상되다	품질 향상, 성능 향상	improvement, advancement, progress; (-하다) to improve, to advance;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
					(-되다) improve, enhance
304	현상 06	現像	현상하다 현상되다		(-하다) to develop (-되다) to be developed
305	현실 02	現實	현실적		reality, actualities; (-적) realistic
306	현장 03	現場		산업 현장	site, scene (of an accident)
307	협력	協力	협력하다		cooperation, collaboration; (-하다) to cooperate
308	형성 01	形成	형성하다 형성되다		formation, development; (-하다) to form, to build up; (-되다) to be formed, to be built
309	형식 01	形式	형식적		form, formality; (-적) formal
310	형태	形態	형태적	가지 형태	shape, form; (-적) physical
311	화면 05	畫面			picture, screen, monitor
312	화학 01	化學	화학적	물리 화학, 생물 화학, 화학 공학, 화학 반응, 화학 물질, 전기 화학, 화학 분석	chemistry; (-적) chemical
313	확보 01	確保	확보하다 확보되다		(-하다) to secure, to retain; (-되다) to be secured, to be retained
314	확장	擴張	확장하다 확장되다	확장 투자	extension; (-하다) to extend; (-되다) to be extended

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
315	활발히	活潑히		활발히 진행하다	vigorously, energetically, actively
316	횟수	回數			number (of times)
317	훈련	訓練	훈련하다 훈련되다	훈련 기관, 교육 훈련	training, drill, discipline, exercise; (-하다) to train, to drill; (-되다) to be trained
318	흐름	(native)		현금 흐름	flow stream
319	가공 01	加工	가공하다 가공되다		process, manufacture; (-하다) to process; (-되다) to be processed
320	가동 07	稼動	가동하다 가동되다		operation; (-하다) to operate; (-되다) to be operated
321	가설 04	假說			hypothesis
322	가속 01	加速	가속하다		acceleration; (-하다) to accelerate;
323	가시	可視	가시적		visibility; (-적) visible
324	가열 02	加熱	가열하다 가열되다		Heat; (-하다) to heat; (-되다) to be heated
325	가중치	加重値			weight
326	각도 01	角度			angle
327	감성 02	感性	감성적		sensitivity, sensibility
328	감쇠	減衰	감쇠하다	감쇠 계수	decrement, decrease, attenuation; (-하다) to damp, to be attenuated
329	감시 02	監視	감시하다 감시되다		surveillance, observation, watch; (-하다) to watch, to monitor;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
					(-되다) to be watched
330	감지 04	感知	감지하다 감지되다		perception, sensing; (-하다) to perceive, sense, detect, be aware of; (-되다) to be sensed, to be detected
331	강성 02	剛性			hardness, stiffness, rigidity
332	강재 02	鋼材			structural steel
333	강하 02	降下	강하하다	압력 강하	falling, dropping, fall; (-하다) to fall, to drop
334	개념	槪念	개념적	기본 개념	concept, idea, notion, conception; (-적) conceptual
335	개소 01	個所			place, site
336	개수 03	個數			number (of)
337	개시 06	開始	개시하다 개시되다	개시 재령, 건조 개시	beginning; (-하다) to begin, to start, to open; (-되다) to begin, to be started
338	개요	槪要			outline, summary
339	개정 02	改正	개정하다 개정되다		revision, amendment; (-하다) to revise, to amend; (-되다) be revised, be amended
340	개체 02	個體	개체적		entity, object; (-적) entity-based
341	객체 02	客體			object
342	거동 03	擧動	거동하다	거동 특성	behavior, conduct; (-하다) to move, to behave

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
343	거래	去來	거래하다 거래되다		deal, trade, business; (-하다) to trade, to deal; (-되다) be traded, be dealt
344	검증	檢證	검증하다 검증되다		verification; (-하다) to verify, to prove; (-되다) be verified, be proven
345	결함 01	缺陷			flaw, fault, defect
346	결합	結合	결합적 결합하다 결합되다	공유 결합, 유도 결합	combination, union, fusion; (-적) cohesive; (-하다) to combine; (-되다) to be united, to be combined
347	경계층	境界層			boundary layer
348	경과	經過	경과하다 경과되다		progress; (-하다) to pass; (-되다) to pass
349	경로 03	經路		차량 경로, 경로 계획	course, route (to), channel
350	경사 06	傾斜			slope, slant, incline, inclination
351	계수 03	係數		탄성 계수, 감쇠 계수, 상관 계수, 열전달 계수, 마찰 계수, 항력 계수, 확산 계수	coefficient, coefficient, modulus, factor
352	계열	系列	계열적		affiliation; (-적) systematic

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
353	고압	高壓	고압적		high-handedness, coercion, pressure; (-적) high-handed, coercive
354	고온 02	高溫			high temperature
355	고유 03	固有	고유하다	고유 진동수	characteristic; (-하다) to be unique
356	고정 05	固定	고정적 고정하다 고정되다		(-적) fixed; (-하다) to be fixed, to be fastened; (-되다) to be fixed
357	고찰 02	考察	고찰하다 고찰되다		consideration, contemplation, (-하다) to consider, to contemplate; (-되다) to be considered, to be contemplated
358	곡선 02	曲線		화재 곡선, 특성 곡선	curve
359	공극 03	空隙			opening, gap, break
360	공급자	供給者			supplier, provider
361	공법 01	工法		해체 공법, 공법 선정	method of construction
362	공유 02	共有	공유하다	공유 결합	sharing; (-하다) to share
363	공정 01	工程		포토 공정, 생물 공정, 제조 공정, 분리 공정, 공정 조건, 공정과 설비, 생산	process

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
				공정, 처리 공정, 공정 제어	
364	공진 02	共振		공진 주파수	resonance
365	공학 01	工學		공학 회지, 토목 공학, 생물 공학, 인간 공학, 유전 공학, 화학 공학, 산업 공학, 전자 공학, 기계 공학, 공학 분야, 자동차 공학, 환경 공학	engineering
366	과도 03	過度	과도하다		excess; (-하다) excessive;
367	관측	觀測	관측하다 관측되다		observation; (-하다) to observe; (-되다) to be observed
368	광 03	光			shine, light, light emission
369	광학 01	光學			optics, optical science
370	교차 01	交叉	교차하다 교차되다		cross, intersect; (-하다) to intersect; (-되다) to be mingled
371	구간 04	區間			section
372	구동	驅動	구동하다		drive; (-하다) to drive
373	구리 02	(native)			copper
374	구배 02	勾配			grade

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
375	구조물	構造物		콘크리트 구조물	structure, construction
376	구체 02	具體	구체적		concreteness; (-적) detailed, definite, specific
377	구축 03	構築	구축하다 구축되다	산업단지 구축, 구축 사업	construction; (-하다) to build, to construct; (-되다) to be constructed, to be built
378	구현 04	具現	구현하다 구현되다		realization, materialization; (-하다) to realize, to materialize; (-되다) to be materialized
379	구획	區劃	구획하다 구획되다		division; (-하다) to divide; (-되다) to be divided
380	국부 01	局部	국부적		local; (-적) local
381	규격	規格			standard
382	규명	糾明	규명하다 규명되다		investigation; (-하다) to investigate; (-되다) to be investigated
383	균열 02	龜裂	균열하다 균열되다	균열 선단, 파쇄 균열, 자연 균열	crack (in), crevice, fissure; (-하다) to crack; (-되다) to crack, to fall out
384	균일	均一	균일하다		equal (to), even; (-하다) to be equal
385	그래프	graph			graph
386	극대 01	極大	극대하다		the greatest, the largest;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
					(-하다) to be greatest
387	근접 01	近接	근접하다		proximity, approach; (-하다) to come close to, to approach
388	긍정	肯定	긍정적 긍정하다 긍정되다		affirmation; (-적) affirmative, positive; (-하다) to affirm; (-되다) to be affirmed
389	기동 04	機動	기동하다		maneuver; (-하다) to move
390	기반 01	基盤	기반하다	규칙 기반	base, footing, foothold, groundwork; (-하다) to base on
391	기상 02	氣象		기상 조건	weather, weather conditions
392	기인 03	起因	기인하다 기인되다		(-하다) to result from, to arise from; (-되다) to be caused
393	기존	旣存	기존하다	기존 건축물, 기존 연구	existence; (-하다) to exist
394	기체 03	氣體			gas
395	기판 04	基板		전극 기판	standard, vexil, vexillum
396	난류 04	暖流		난류 유동, 난류 모델	warm current
397	내구 07	耐久	내구하다		endurance, sustenance, persistence; (-하다) to endure, to sustain
398	내장 05	內藏	내장하다 내장되다		being built-in; (-하다) to be built-in; (-되다) to be built in, to be equipped with
399	냉각 01	冷却	냉각하다		cooling;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
			냉각되다		(-하다) to cool, to freeze; (-되다) to be cooled, to be frozen
400	노드	node		중계 노드, 모바일 노드	node
401	노이즈	noise			noise
402	노출	露出	노출하다 노출되다	노출 시간	exposure, disclosure; (-하다) to expose, to reveal; (-되다) to be exposed, to be revealed
403	누적	累積	누적하다 누적되다		accumulation; (-하다) to accumulate; (-되다) to be accumulated
404	다중 01	多重			multiplex, multiple
405	단말기	端末機			terminal (unit)
406	단면 02	斷面	단면적	날개 단면	cross section, aspect; (-적) fragmentary
407	단일 01	單一	단일하다		singleness, unity; (-하다) to be single
408	단축 02	短縮	단축하다 단축되다		shortage (-하다) to shorten, cut (down); (-되다) to be shortened
409	단층 02	斷層		단층 촬영	layer
410	달성	達成	달성하다 달성되다		achievement, accomplishment; (-하다) to achieve, to accomplish; (-되다) to be achieved
411	대류	對流		대류 열전달	convection current

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
412	대상 11	對象	대상적	자극 대상, 평가 대상, 분석 대상, 연구 대상	object, target, subject; (-적) targeted
413	대수 14	臺數			number
414	대안 03	代案			alternative
415	대역폭	帶域幅			bandwidth
416	대칭 02	對稱	대칭적 대칭하다 대칭되다		symmetry; (-적) being symmetrical (-하다) to be symmetric; (-되다) to be symmetric
417	데이터	data		데이터 마이닝, 데이터 버전, 데이터 링크, 계측 데이터, 데이터 전송, 데이터 통신, 입력 데이터, 데이터 처리, 데이터 모델, 실험 데이터, 데이터 분석	data(datum)
418	도식 06	圖式	도식적		diagram; (-적) graphic, formulaic
419	도출 02	導出	도출하다 도출되다		deduction; (-하다) to deduct; (-되다) to be deduced, to be drawn
420	독성 01	毒性		흡입 독성, 독성 가스, 독성 시험	toxicity
421	동영상	動映像			video

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
422	동적 01	動的			dynamic, kinetic
423	동향 04	動向		연구 동향	trend
424	등급	等級			class, grade, rating
425	디젤	diesel		디젤 엔진	diesel
426	레벨	level			level
427	루프	loop			loop
428	링크	link		데이터 링크	link
429	망 07	網			net, network
430	매립	埋立	매립하다		(land) reclamation; (-하다) to reclaim;
431	멀티미디어	multimedia			multimedia
432	메모리	memory			memory
433	메커니즘	mechanism			mechanism
434	메탄	methane			methane
435	명시 02	明示	명시적 명시하다 명시되다		clear statement; (-적) explicit, clear (-하다) to specify, to state; (-되다) to be stated explicitly
436	모니터링	monitoring		전력 모니터링, 모니터링 시스템	monitoring
437	모델링	modeling		프로세스 모델링	modeling
438	모듈	module		통신 모듈	module
439	모드	mode			mode
440	모멘트	moment		모멘트 텐서, 굽힘 모멘트	moment

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
441	모바일	mobile		모바일 로봇, 모바일 노드	mobile
442	모터	motor			motor
443	무선 02	無線		무선 통신	radio
444	무인 04	無人		무인 잠수정, 무인 항공기	manless
445	문헌	文獻		참고 문헌	literature (on)
446	물류	物流			distribution
447	물리	物理	물리적	물리 화학, 물리 계층	physics; (-적) physical
448	물성	物性			a property of matter
449	미세	微細	미세하다		(-하다) fine, tiny
450	민감	敏感	민감하다		(-하다) sensitive
451	밀도 01	密度		연기 밀도, 교통 밀도, 상대 밀도, 공기 밀도	density
452	바이오	bio			bio
453	박리	剝離	박리하다		desquamation; (-하다) to desquamate, to peel off
454	반경 02	半徑		반경 방향	radius
455	반도체	半導體		반도체 물질	semiconductor
456	반사 01	反射	반사적 반사하다 반사되다		reflection; (-적) reflexive; (-하다) to reflect; (-되다) to be reflected
457	반응기	反應器			effector
458	발사 01	發射	발사하다 발사되다		launch, discharge; (-하다) to launch, to set off;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
					(-되다) to be launched, to be set off
459	발현	發現		강도 발현	manifest
460	방법론	方法論			methodology
461	방사 07	放射	방사하다 방사되다	방사 패턴	radiation; (-하다) to radiate; (-되다) to radiate
462	방정식	方程式		지배 방정식, 운동 방정식	equation
463	방출 01	放出	방출하다 방출되다		release; (-하다) to release, to distribute; (-되다) to be released
464	배수 07	排水	배수하다 배수되다	광산 배수	drainage; (-하다) to drain; (-되다) to be drained
465	배열 02	排列	배열하다 배열되다	수신기 배열	arrangement, sequence; (-하다) to arrange; (-되다) to be arranged
466	배출 02	排出	배출하다 배출되다	탄소 배출, 오염 배출, 배출 가스, 물질 배출	emission, discharge; (-하다) to emit; (-되다) to be emitted
467	베이스	base			base
468	벡터	vector		속도 벡터	vector
469	벽면	壁面			the surface of a wall
470	변수 06	變數		종속 변수, 독립 변수, 설비 변수, 입력 변수, 결정 변수, 상태 변수, 설계 변수	variable

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
471	변위	變位	변위하다		displacement; (-하다) to displace
472	변형	變形	변형하다 변형되다	파단 변형, 수축 변형, 전단 변형	modification; (-하다) to modify, to transform; (-되다) to be modified
473	변환 03	變換	변환하다	변환 효율, 에너지 변환	change, conversion, diversion; (-하다) to convert
474	병렬	竝列	병렬적 병렬하다 병렬되다	병렬 전단	line, row; (-적) parallel; (-하다) to be in a row, to stand in a line; (-되다) to be in a row
475	보급 02	補給	보급하다 보급되다		supply; (-하다) to supply; (-되다) to be supplied
476	보유 01	保有			possession; (-하다) to possess; (-되다) to be possessed
477	보정 02	補正	보정하다		revision, correction, compensation; (-하다) to revise, to correct, to compensate
478	복원 02	復元/復原	복원하다 복원되다		restoration; (-하다) to restore, reconstruct; (-되다) to be restored, to be reconstructed
479	복귀	復歸	복귀하다 복귀되다		return; (-하다) to return, to come back; (-되다) to be returned, to be restored,
480	복합 02	複合	복합적	복합 재료	compositeness, complex, compound;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
			복합하다 복합되다		(-적) complex; (-하다) to compound, to mix; (-되다) to compound, to be mixed
481	볼트	bolt			bolt
482	부가 02	附加	부가적 부가하다 부가되다		addition, annexation, supplement; (-적) additional, supplementary; (-하다) to add; (-되다) to be added
483	부식 07	腐蝕	부식하다 부식되다	외면 부식	corrosion; (-하다) to decay; (-되다) to be corroded
484	부여 04	附與	부여하다 부여되다		(-하다) to give, to grant; (-되다) to be given, to be granted
485	부착	附着	부착하다 부착되다		sticking, adherence, adhesion, cohesion; (-하다) to stick, to attach; (-되다) to be stuck, to be attached
486	부하 02	負荷	부하하다	부하 전류	load; (-하다) to strain
487	분류 03	分類	분류하다 분류되다	유형 분류	classification, categorization; (-하다) to classify, to group, to categorize; (-되다) to be sorted, to be classified
488	분말 01	粉末			powder
489	분배	分配	분배하다		distribution, division;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
			분배되다		(-하다) to distribute, to divide; (-되다) to be allocated
490	분산 02	分散	분산적 분산하다 분산되다	분산 전원	breakup, dispersion; (-적) dispersed; (-하다) to disperse; (-되다) to be dispersed
491	분할	分割	분할하다 분할되다		division, partition, dismemberment; (-하다) to divide; (-되다) to be divided, to be partitioned
492	분해	分解	분해하다 분해되다		disassemble, dismantle; (-하다) to disassemble; (-되다) to be disassembled
493	불량 01	不良	불량하다		(-하다) poor, faulty
494	붕괴	崩壞	붕괴하다 붕괴되다		collapse; (-하다) to collapse, to break down; (-되다) to be collapsed
495	블록	block		캐쉬 블록, 블록 공중	block
496	비대칭	非對稱			asymmetry
497	비례 01	比例	비례적 비례하다		proportion; (-적) proportional; (-하다) to be in proportion;
498	빈도 02	頻度			frequency
499	사료 03	思料	사료하다 사료되다		consideration, thinking; (-하다) to think, to consider; (-되다) to be considered

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
500	사슬 01	(native)		가치 사슬	chain
501	사용량	使用量		용수 사용량	amount (of a thing) used
502	사이클	cycle		사이클 슬립	cycle
503	사이트	site			site
504	사항 02	事項		요구 사항, 고려 사항	matter, subject
505	산정 09	算定	산정하다 산정되다		calculation, estimate, assessment; (-하다) to compute, to calculate; (-되다) to be calculated, to be computed
506	산출 02	算出	산출하다 산출되다	산출 요소	calculation, computation; (-하다) to compute, to calculate; (-되다) to be calculated, to be computed
507	산화 06	酸化	산화하다 산화되다		oxidation, oxidize; (-하다) to oxidize; (-되다) to be oxidized
508	상기 02	上記	상기하다		above statements; (-하다) to write above, to be aforementioned
509	상부 03	上部			superior, upper
510	상세 08	詳細	상세하다		details, particulars; (-하다) detailed;
511	상수 09	常數		속도 상수	constant, invariable
512	상승	上昇	산승하다 산승되다	온도 상승	rise, increase; (-하다) to rise, to increase; (-되다) be risen, to be included
513	상용 03	常用	상용하다		common use, everyday use;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
			상용되다		(-하다) to use every day; (-되다) to be widely used
514	상위 01	上位			high rank
515	상호 04	相互		상호 작용, 상호 운용, 상호 인증	mutual
516	샘플	sample			sample
517	샘플링	sampling			sampling
518	생물학	生物學			biology
519	생산량	生産量			output
520	생성	生成	생성하다 생성되다		creation, formation; (-하다) to create, to generate; (-되다) to be created
521	서버	server			server
522	석탄	石炭			coal
523	선도 05	先導	선도적 선도하다		(-적) leading (-하다) to lead, to guide
524	선로 04	線路			tracks, railroad
525	선행 01	先行	선행하다 선행되다	선행 연구	precedence; (-하다) to precede; (-되다) to be preceded
526	선형 05	線形			linear
527	설계 02	設計	설계하다 설계되다	내진 설계, 참여 설계, 설계 기준, 설계 단계, 건축 설계, 설계 변수, 기본	plan, map, design; (-하다) to plan, to design; (-되다) to be planned, to be designed

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
				설계, 구조 설계, 설계 과정, 시스템 설계	
528	설비 02	設備	설비하다 설비되다	피뢰 설비, 소화 설비, 설비 변수, 전기 설비, 공정 설비	facilities, equipment, system; (-하다) to equip; (-되다) to be equipped
529	설정 02	設定	설정하다 설정되다		set-up; (-하다) to set up; (-되다) to be set up
530	섬유 02	纖維		섬유 염색	fiber, textile
531	성능	性能		내진 성능, 내화 성능, 성능 보증, 조종 성능, 성능 향상, 성능 평가, 성능 개선, 제어 성능	performance
532	성분 01	成分		구성 성분	ingredient, component, constituent
533	세부	細部	세부적		details, particulars, fine parts; (-적) detailed
534	센서	sensor			sensor
535	셀	cell			cell
536	소모 02	消耗	소모적 소모하다 소모되다	전력 소모	consumption; (-적) consuming; (-하다) to consume; (-되다) to be consumed
537	소성 07	塑性			plasticity

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
538	소요 02	所要	소요하다 소요되다	소요 시간	requirement; (-하다) to take, to cost; (-되다) to be required
539	소자 07	素子		전자 소자	element, device
540	속성 05	屬性			attribute, properties, quality
541	송수신	送受信			transmitting and receiving
542	송신 01	送信	송신하다 송신되다		transmission; (-하다) to transmit; (-되다) to be transmitted
543	수렴 03	收斂	수렴하다 수렴되다		convergence; (-하다) to collect, to converge; (-되다) to be collected, to be converged
544	수립 02	樹立	수립하다 수립되다	계획 수립	establishment, founding; (-하다) to establish; (-되다) to be established
545	수분 01	水分			water, moisture
546	수소 03	水素		불화 수소	hydrogen
547	수송	輸送	수송하다		transportation; (-하다) to transport
548	수식 04	修飾	수식하다 수식되다		modification; (-하다) to modify; (-되다) to be modified
549	수신 07	受信	수신하다 수신되다		reception; (-하다) receive; (-되다) to be received
550	수압 02	水壓		수압 파쇄	hydraulic pressure
551	수용 06	受容	수용하다		(-하다) accept;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
			수용되다		(-되다) to be accepted
552	수익	收益	수익하다		profit, earnings; (-하다) to make money
553	수정 09/08	修訂/修正	수정하다 수정되다		modification; correction, rectification; (-하다) to correct, to modify; (-되다) to be corrected, to be modified
554	수지 04	收支		교환 수지	income and expenditure, earnings and expenses
555	수직 06	垂直	수직적	수직 방향, 수직 속도	verticality; (-적) vertical
556	수축 01	收縮	수축하다 수축되다	건조 수축, 자기 수축, 수축 변형	contraction, contract, shrink; (-하다) to be contracted; (-되다) to be contracted
557	수치 05	數値		수치 해석, 수치 계산	numerical value, figure;
558	수평 02	水平		수평 안정판, 수평 방향	horizontality
559	순 12	順			order
560	순차 02	順次	순차적		order, turn; (-적) being sequential
561	순환 01	循環	순환하다 순환되다		circulation, rotation; (-하다) to circulate, to rotate; (-되다) to be circulated, to be rotated
562	스마트	smart		스마트 무인기	smart
563	스마트폰	smartphone			smartphone
564	스펙트럼	spectrum			spectrum

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
565	시뮬레이션	simulation		시뮬레이션 결과	simulation
566	시설물	施設物		교육 시설물	facility, establishment
567	시스	sheath			sheath
568	시편 03	試片			test specimen
569	시행 01	施行	시행하다 시행되다		enforcement, implementation; (-하다) to enforce, to conduct; (-되다) to enforced, to be carried out
570	식별	識別	식별하다 식별되다	식별 규칙	(-하다) to distinguish, to discern; (-되다) to be discerned
571	신경망	神經網		인공 신경망, 신경망 모델	a nerve network
572	신뢰 02	信賴	신뢰하다		confidence; (-하다) to trust (in), to have confidence (in)
573	신뢰도	信賴度			credibility, reliability
574	실린더	cylinder			cylinder
575	실시간	實時間			real-time
576	실용	實用	실용적 실용하다 실용되다		practical use, utility; (-적) practical; (-하다) to put to practical use; (-되다) to be practically used
577	실질	實質	실질적		material, matter, essence, substance; (-적) actual, real, being essential, being substantial
578	실행 02	實行	실행하다		practice, action;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
			실행되다		(-하다) to carry out, to implement; (-되다) to be carried out, to be executed
579	알고리즘	algorithm		휴리스틱 알고리즘, 유전 알고리즘	algorithm
580	알칼리	alkali			alkali
581	압	壓			pressure
582	압축	壓縮	압축하다 압축되다		compression, summarization; (-하다) to compress, to condense, to summarize; (-되다) to be condensed, to be summarized
583	압축강도	壓縮强度			compression strength
584	압축기	壓縮機			compressor
585	액체	液體			liquid, fluid
586	야기 02	惹起	야기하다 야기되다		(-하다) to cause, to arouse; (-되다) to be caused
587	양상 07	樣相			aspect, appearance, look
588	억제	抑制	억제하다 억제되다		control; (-하다) to control, to suppress, to restrain; (-되다) to be controlled, to be suppressed
589	언급	言及	언급하다 언급되다	앞서 언급	reference, mention; (-하다) to refer, to mention; (-되다) to be mentioned, to be stated

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
590	여부	與否			whether
591	역량 02	力量			capability; competence (in/of), ability (to do), capacity
592	역학 01	力學		유체 역학, 공기 역학	dynamics
593	연계 03	連繫	연계하다 연계되다		ties, connection, link, liaison; (-하다) to connect, to link; (-되다) to be connected, to be linked
594	연료	燃料		화석 연료	fuel
595	연비 03	燃比			(gas) mileage, fuel efficiency
596	연성 01	延性			ductility, malleability
597	열전달	熱傳達		대류 열전달, 열전달 계수	heat transfer (transmission)
598	열화 03	熱火			blazing fire, furious flames
599	오차	誤差			error, mistake
600	와류 02	渦流	와류하다		whirlpool, maelstrom; (-하다) to swirl, to flow in whirls
601	완료	完了	완료하다 완료되다	완료 시간	completion; (-하다) to complete; (-되다) to be completed, to be finished
602	완전 01	完全	완전하다		(-하다) complete, full
603	외기 02	外氣		외기 조건	(open) air
604	외피	外皮		이중 외피, 외피 구조	integument, skin, outer cover, rind, shuck, crust, shell
605	요소 04	要素		유한 요소, 투입 요소, 구성	element, constituent, factor

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
				요소, 산출 요소, 위험 요소, 요소 기술	
606	요인 02	要因		결정 요인, 요인 분석	key factor, important factor
607	용량 02	容量		정전 용량	capacity, volume
608	용액	溶液			solution
609	용적	容積			capacity
610	용해 02	溶解	용해하다 용해되다		dissolution, melting, liquefaction; (-하다) to dissolve, to lyse; (-되다) to melt, to dissolve
611	운용	運用	운용하다 운용되다	상호 운용	management; (-하다) to apply, to employ; (-되다) to be applied, to be operated
612	원가 05	原價		대체 원가	production cost, prime cost, cost price
613	원격 02	遠隔	원격하다		remote; (-하다) to be far apart, to be widely separated; remote
614	원료 02	原料			raw material, base material
615	원리 02	原理	원리적		principle, fundamentals; (-적) fundamental
616	원소 02	元素			(chemical) element
617	원자 02	原子			atom
618	원형 03	原形			original form
619	원활	圓滑	원활하다		smoothness, harmony; (-하다) smooth

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
620	위상 01	位相			status
621	유기물	有機物			organic matter
622	유도 07	誘導	유도하다 유도되다	유도 동력, 유도 결합,	inducement, induce, lead; (-하다) to guide, to induce; (-되다) to be guided, to be induced
623	유동 02	流動	유동적 유동하다	난류 유동, 유동 해석, 유동 제어, 유동 특성	flow, flowing, floating, flowage; (-적) being fluid; (-하다) to flow, to circulate
624	유량 03	流量		질량 유량, 공기 유량	flux
625	유무 02	有無			existence and nonexistence
626	유속 02	流速			velocity[speed] of a moving[running] fluid
627	유용 01	有用	유용하다		(-하다) useful, helpful;
628	유입 01	流入	유입하다 유입되다		inflow, influx; (-하다) to flow in; (-되다) to flow into
629	유전 12	遺傳	유전적 유전하다 유전되다	유전 알고리즘, 유전 공학	heredity, inheritance, hereditary transmission; (-적) being hereditary; (-하다) to be inherited, to be hereditary; (-되다) to be inherited, to be hereditary
630	유체 05	流體		유체 역학	fluid
631	유출 01	流出	유출하다 유출되다		spill; (-하다) to be spilled, to leak;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
					(-되다) to be spilled, to be leaked
632	유해 02	有害	유해하다		(-하다) harmful, noxious, hazardous
633	유효 01	有效		조직 유효	(-하다) valid, available, good, effective
634	응력	應力		잔류 응력, 인장 응력, 전단 응력	stress, internal stress
635	응용	應用	응용하다 응용되다	응용 프로그램, 응용 분야	application of smth; (-하다) to apply; (-되다) to be applied
636	이론	理論	이론적		theory; (-적) theoretical
637	이산화	二酸化		이산화 탄소	dioxide
638	이온	ion		리튬 이온, 이온 교환, 이온 전지	ion
639	인가 05	認可	인가하다 인가되다		permission; (-하다) to permit, to approve; (-되다) to be permitted, to be approved
640	인력 01	人力			manpower, labor force
641	인자 06	因子			factor
642	인장 02	引張		인장 강도, 인장 응력	tension
643	인접 02	鄰接	인접하다 인접되다		(-하다) to adjoin, to border on, (-되다) to adjoin, to be contiguous to
644	인증 03	認證	인증하다 인증되다	인증 기법, 상호 인증	certification, accreditation, confirmation; (-하다) to certify;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
					(-되다) to be certified
645	인터페이스	interface			interface
646	인프라	infrastructure			infrastructure
647	임계 02	臨界		임계 속도	critical
648	임의	任意	임의적		option, discretion, voluntariness; (-적) optional
649	입자 02	粒子		나노 입자	particle
650	입출력	入出力			input/output
651	작동 01	作動	작동하다 작동되다		operation; (-하다) to operate, to run, to work; (-되다) to be operated
652	잔류 01	殘留	잔류하다	잔류 응력	(-하다) remain, stay (behind)
653	장기 09	長期			long period of time; (-적) long-term
654	장력 02	張力		장력 철근	tension
655	장애 02	障礙		독서 장애	disability, defect, impediment, impairment, disorder
656	장애물	障礙物			obstacle
657	장착	裝着	장착하다 장착되다		installation; (-하다) to equip, to install; (-되다) to be equipped
658	장치 05	裝置	장치하다 장치되다	단말 장치, 기억 장치, 장치 정비, 보호 장치, 저장 장치, 입력 장치, 처리 장치, 제어	equipment, installation, device; (-하다) to set up, to install; (-되다) to be set up, to be embedded

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
				장치, 실험 장치, 측정 장치	
659	재질 02	材質			quality of the material
660	저감	低減	저감하다		reduction, diminution, decrease; (-하다) to reduce
661	저장 03	貯藏	저장하다 저장되다	저장 탱크, 저장 장치, 에너지 저장	storage, preservation; (-하다) to store; (-되다) to be stored
662	저하 01	低下	저하하다 저하되다		decline; (-하다) to fall, to decline; (-되다) to drop, to decline
663	저항	抵抗	저항하다	접촉 저항, 회전 저항	resistance; (-하다) resist
664	적재 04	積載	적재하다 적재되다		(-하다) to load (up); (-되다) to be loaded with
665	적정 06	適正	적정하다		(-하다) optimal, appropriate
666	전극 02	電極		전극 기판	electrode, (magnetic) pole
667	전단 06	剪斷	전단하다	병렬 전단, 전단 보강, 전단 강도, 전단 응력, 전단 변형	shear; (-하다) to shear
668	전략 03	戰略	전략적		strategy, tactic; (-적) tactical
669	전산 03	電算			(electronic) data processing
670	전이 03	轉移	전이하다 전이되다	상태 전이	metastasis, spread; (-하다) to move, to spread;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
					(-되다) to be spread, to be passed on
671	전진 04	前進	전진적 전진하다		forward movement, advance; (-적) progressive; (-하다) to move forward, to improve
672	전파 04	電波			radio wave[signal], electromagnetic wave[signal]
673	전형 04	典型	전형적		model; (-적) typical
674	절감 02	節減	절감하다 절감되다	에너지 절감, 비용 절감, 절감 효과	reduction, retrenchment; (-하다) to cut (down), to reduce, to trim; (-되다) to be reduced, to be cut
675	절단 01	切斷	절단하다 절단되다		(-하다) to cut; (-되다) to be cut
676	점성 03	粘性			viscosity
677	점토	粘土			clay
678	접지 01	接地	접지하다		ground[earth] connection, grounding, earthing; (-하다) to ground, to earth
679	접합	接合	접합하다 접합되다		connection, inosculation, anastomosis; (-하다) to join, to replant; (-되다) to be joined, to be replanted
680	정규 01	正規	정규적	정규 분포	regularity, formality, legality; (-적) regular, formal

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
681	정량 03	定量	정량적 정량하다		fixed quantity, standard capacity, dose; (-적) quantitatively; (-하다) to quantify;
682	정밀 02	精密	정밀하다		(-하다) detailed, close, precise, accurate, exact
683	정수 21	整數			integer, whole number
684	정의 05	定義	정의하다 정의되다	정의 영향	definition; (-하다) define (smth as smth); (-되다) to be defined
685	정적 08	靜的			still, static
686	정책 02	政策	정책적		policy; (-적) policy-level
687	정체 03	停滯	정체적 정체하다 정체되다		tie-up, stagnation; (-적) stagnant; (-하다) to be stalled; (-되다) to be stalled
688	정확도	正確度			accuracy
689	제도 01	制度	제도적		system; (-적) institutional
690	제어 01	制御	제어하다 제어되다	제어 이득, 열차 제어, 유동 제어, 전자 제어, 제어 시스템, 제어 장치, 산업 제어, 제어 기능, 공정 제어, 제어 성능, 제어	control; (-하다) to control, to keep back; (-되다) to be controlled

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
				특성, 제어 가능, 제어 기술	
691	제어기	制御器		예측 제어기	controller
692	제정 01	制定	제정하다 제정되다		enactment; (-하다) to establish, to enact; (-되다) to be enacted, to be designated
693	제조 07	製造	제조하다 제조되다	제조 업자, 제조 업체, 제조 공정	manufacture, production; (-하다) to manufacture, to produce; (-되다) to be manufactured, to be produced
694	조류 02	潮流			tidal current
695	조립 02	組立	조립하다 조립되다	자기 조립	assembling; (-하다) to assemble; (-되다) to be assembled
696	조성 04	造成	조성하다 조성되다		(-하다) to build, to develop, to make; (-되다) to be built, to be developed
697	조작 05	操作	조작하다 조작되다		(-하다) to operate, to control, to handle, to work, to manipulate; (-되다) to be manipulated
698	조합 04	調合	조합하다 조합되다	모델 조합	combination; (-하다) to mix, to combine, to compound; (-되다) to be combined
699	조화 07	調和	조화하다 조화되다		harmony, balance; (-하다) to get along with, to go with;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
					(-되다) to get along with, to go with
700	종료	終了	종료하다 종료되다		end, close, conclusion; (-하다) to end, to close; (-되다) to be finished, to expire
701	주기 13	週期	주기적		period, cycle; (-적) periodical, cyclical
702	주도 02	主導	주도적 주도하다	자기 주도	lead; (-적) leading; (-하다) to lead
703	주목 03	注目	주목하다 주목되다		attention; (-하다) to pay[give] attention (to); (-되다) to be watched
704	주파수	周波數		공진 주파수, 주파수 대역, 주파수 응답, 주파수 영역	frequency
705	주행 02	走行	주행하다	주행 거리	(-하다) to drive, run
706	중력 03	重力		중력 중심	gravity, gravitation, force of gravity, gravitational pull
707	증대	增大	증대하다 증대되다		increase; (-하다) to increase, to grow; (-되다) to increase, to grow
708	증폭	增幅	증폭하다 증폭되다	증폭 회로	amplification; (-하다) to enlarge, to intensify; (-되다) to be enlarged, to be intensified
709	지름 02	(native)			diameter

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
710	지상 01	地上			ground
711	지속 01	持續	지속적 지속하다 지속되다	지속 시간, 지속 가능	(적) continuous; (-하다) to persist, to carry on; (-되다) to last, to endure
712	지수 06	指數			index, quotient
713	지연 03	遲延	지연하다 지연되다	지연 시간	postponement, delay; (-하다) to delay, to postpone; (-되다) be delayed; to be postponed
714	지정 13	指定	지정하다 지정되다		designation, appointment; (-하다) to designate; (-되다) to be designated
715	지침	指針			guidelines
716	지향 02	志向	지향하다		aim; (-하다) to aim
717	지형 01	地形			topography, geographical features
718	직경 01	直徑			diameter
719	직관 03	直觀	직관적 직관하다	전문 직관	Intuition; (-적) intuitive; (-하다) to have an intuition
720	직류 01	直流			direct current
721	진동수	振動數		고유 진동수	the number of vibrations
722	진입	進入	진입하다		entry, penetration; (-하다) to enter, to penetrate
723	진폭 01	振幅			amplitude (of vibration)
724	질량	質量	질량적	질량 유량	mass;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
					(-적) massive
725	질소 01	窒素		질소 산화물	nitrogen
726	집합 01	集合	집합하다 집합되다		set; (-하다) to assemble, to gather; (-되다) to be assembled
727	차원 02	次元		차원 모델	level, dimension
728	차지 08	charge			charge
729	창출 01	創出	창출하다 창출되다		(-하다) to create; (-되다) to be created
730	채취	採取	채취하다 채취되다		collection; (-하다) to collect, to gather; (-되다) to be collected, to be gathered
731	채택	採擇	채택하다 채택되다		adoption; (-하다) to adopt, to choose; (-되다) to be selected, to be chosen
732	척도	尺度			criterion
733	천연 01	天然	천연적	천연 물질	nature; (-적) naturally
734	천연가스	天然 gas			natural gas
735	철골 03	鐵骨		철골 커플링	steel frame, steel skeleton
736	체계 03	體系	체계적	지원 체계	system; (-적) systematic
737	체적 02	體積		체적 변화	volume, bulk
738	초과 02	超過	초과하다 초과되다		excess; (-하다) to exceed; (-되다) to be exceeded
739	초래 01	招來	초래하다		cause, lead;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
			초래되다		(-하다) to cause, to lead to, to result in; (-되다) to be incurred
740	촉진 01	促進	촉진하다 촉진되다		expedite promotion, acceleration; (-하다) to accelerate, to quicken; (-되다) to be expedited, to be quickened
741	최적	最適			optimum
742	최적화	最適化			optimization
743	추구 01	追求	추구하다		pursuit; (-하다) to pursuit, to chase
744	추세 03	趨勢			trend, tendency, drift
745	추정 02	推定	추정하다 추정되다	일사량 추정	estimation, assumption, presumption; (-하다) to estimate, to assume; (-되다) to be estimated, to be assumed
746	추출	抽出	추출하다 추출되다		extraction; (-하다) to extract; (-되다) to be extracted
747	축 11	軸			axis (axes)
748	축적	蓄積	축적하다 축적되다		accumulation; (-하다) to accumulate; (-되다) to be accumulated
749	출력	出力	출력하다	출력 전압	output; (-하다) to print, to output, to display
750	충족	充足	충족하다 충족되다		fulfillment; (-하다) to satisfy, to fill, to meet;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
					(-되다) to be satisfied, to be filled, to be met
751	취급	取扱	취급하다 취급되다		handling; (-하다) to treat, to handle, to serve; (-되다) to be treated, to be served
752	취득	取得	취득하다		acquisition; (-하다) to acquire, to gain, to get, to obtain
753	취약	脆弱	취약하다		(-하다) to weak, to vulnerable
754	측	側			side
755	측면	側面			side, flank, aspect
756	측정 01	測定	측정하다 측정되다	압력 측정, 측정 장치, 측정 방법, 측정 결과	measurement; (-하다) to measure, to survey; (-되다) to be measured, to be surveyed
757	치수 01	치數			size, measurement
758	침투	浸透	침투하다 침투되다		invasion; (-하다) infiltrate, invade, penetrate; (-되다) to get through, to infiltrate, to spread
759	칩	chip			chip
760	컴퓨팅	computing			computing
761	케이블	cable			cable
762	코어	core			core
763	클러스터	cluster			cluster
764	탄성 01	彈性		탄성 계수	elasticity
765	탄소 01	炭素		이산화 탄소, 탄소 함유량,	carbon

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
				탄소 나노, 탄소 배출	
766	탐색 02	探索	탐색하다		exploration; (-하다) to explore, to search
767	탑재	搭載	탑재하다		load; (-하다) to load, to embark
768	터빈	turbine			turbine
769	통상 02	通常	통상적		(-적) usually, normally, generally
770	통신망	通信網			communications network
771	투입	投入	투입하다 투입되다	투입 요소	investment, injection; (-하다) to invest, to spend; (-되다) to be invested, to be spent
772	특허	特許	특허하다	특허 활동	patent (on/for); (-하다) to patent
773	파라미터	parameter		형상 파라미터	parameter
774	파장 02	波長			wavelength
775	파형	波形			ripple mark, wavy pattern
776	패턴	pattern		방사 패턴	pattern
777	펌프	pump			pump
778	편차 02	偏差			deviation
779	평균값	平均값			average value
780	평면	平面	평면적		flat surface; (-적) flat, plane
781	평행	平行	평행적 평행하다 평행되다		parallel; (-적) parallel; (-하다) to go side by side, to be parallel; (-되다) to be parallel

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
782	평형 02	平衡			balance
783	폐기물	廢棄物			waste
784	포트	port			port
785	포화 06	飽和	포화하다		saturation; (-하다) to be saturated
786	폼	form			form
787	표본	標本			specimen
788	표적 03	標的			target, mark
789	프레임	frame			frame
790	프로젝트	project		건설 프로젝트	project
791	플랜트	plant			plant
792	플랫	flat			flat
793	피드백	feedback			feedback
794	피크	peak			peak
795	필드	build			build
796	필터	filter			filter
797	하류 01	下流			downstream
798	하부 03	下部			lower part
799	하위 02	下位			low rank
800	하중 02	荷重		익면 하중, 하중 조건	weight, load
801	학문 02	學問	학문적 학문하다		study, learning; (-적) academic, scholarly; (-하다) to study, to learn
802	할당	割當	할당하다 할당되다		assignment, delegation, allotment, allocation; (-하다) to assign, to allot;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
					(-되다) to be assigned, to be allotted
803	함량	含量			content
804	함수 04	函數		목적 함수, 전달 함수	function
805	함유 01	含有	함유하다 함유되다		(-하다) to contain, to include; (-되다) to be included, to be contained
806	합금 01	合金			alloy
807	합리 01	合理	합리적 합리하다		rationality; (-적) rational, logical; (-하다) to be reasonable
808	합성 01	合成	합성하다 합성되다	합성 얼음	synthesis; (-하다) to compound, to synthesize; (-되다) to be compounded, to be synthesized
809	항목	項目		수질 항목, 평가 항목	item
810	해상 02	海上		해상 풍력	on the sea, sea, maritime, marine
811	해양	海洋	해양적		ocean, sea; (-적) oceanic
812	행렬	行列		대기 행렬, 행렬 모형	line, matrix
813	행하다	行 --			to do, to act, to conduct
814	향후	向後		향후 연구	henceforth, henceforward
815	현황 02	現況			present condition[situation], current state of affairs
816	혼합	混合	혼합하다	혼합 경량	mixture, blend; (-하다) to mix, to blend;

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
			혼합되다		(-되다) to be mixed, to be blended
817	홉	hop			hop
818	화물	貨物			freight, cargo, goods
819	화합물	化合物			compound
820	확률	確率		발생 확률	probability, chance, likelihood
821	확산	擴散	확산하다 확산되다	확산 계수	spread, diffusion; (-하다) to spread, to diffuse, to disperse; (-되다) to be dispersed, to be diffused
822	환기 02	換氣	환기하다 환기되다		ventilation; (-하다) to ventilate; (-되다) to be ventilated
823	회로 01	回路		정류 회로, 증폭 회로, 논리 회로	(electrical) circuit, circuitry
824	회수	回收	회수하다 회수되다		collection; (-하다) to retrieve, to recover, to collect; (-되다) to be retrieved, to be recovered, to be collected
825	회전수	回轉數		분당 회전수	rev count
826	획득	獲得	획득하다 획득되다		acquisition; (-하다) to acquire, to obtain, to gain; (-되다) to be acquired, to be obtained
827	횡	橫	횡적		width, crossways; (-적) horizontal
828	효율	效率	효율적	변환 효율, 에너지 효율	efficiency, effectiveness; (-적) horizontal

№	Word	Origin word	Modified Lemmas	Collocations	Translation to English
829	흡수 02	吸收	흡수하다 흡수되다		absorption; (-하다) to absorb (-되다) to be absorbed
830	흡입	吸入	흡입하다	흡입 독성	inhalation; (-하다) to inhale, to breathe in