

Constructing the English Loanwords Contents (ELC) Database

Seung-mi Cheon

(Hankuk University of Foreign Studies)

Cheon, Seung-mi. Constructing the English Loanwords Contents (ELC) Database. *The Linguistic Association of Korea Journal*, 13(2), 97-119. This paper explores the complexities of constructing the English Loanwords Contents (ELC) database, which is a listing of English loanwords culled from four dictionaries and hundreds of articles in Korean newspapers. The database is examined for its use as an educational tool for English as a Second Language learners in Korea, particularly with respect to vocabulary.

Key words: English loanwords, educational tool, vocabulary, ELC, database, Corpora, LADL, UNITEX, INTEX, FSGraph, CALL

1. Introduction

The English language has made an enormous impact on Korea. Indeed, it has become the first foreign language to exercise a great influence on the language life of Korean people. A survey by Yang (2002) shows that television has the greatest influence on newspapers, radio, and magazines in regards to the spread and use of English in Korea. Terms borrowed from English can also be found frequently in Korean documents. The growth of technology in specific domains produces a new terminology¹⁾, which corresponds to new concepts. The problem we must consider is that most of this terminology is not registered in dictionaries and this can result in serious errors in natural language processing systems, such as information retrieval²⁾.

1) The linguistic expressions of concepts in the specific fields.

2) Search engines such as Internet Information Retrieval or Retrieval Systems

A more profound problem in dealing with English loanwords is the transliteration of their loanwords from foreign words, particularly with respect to proper nouns and technical terms. This problem is significant, especially in countries where English is not an official language, such as Korea, where the majority of these new technical words and phrases are created in English. In Korea, many texts often contain English words in their original form (namely, in Roman character set) or in Hangeul following transliteration.

To render the problem more complex, diverse transliterations of one loanword can be used in a variety of ways within one document. For example, the word *carpet* [ka:rpit] may appear in English or it may appear as (pronounced [kapet]), a Koreanized loanword originating from the English word. The loanword and its English original, *carpet*, refer to the same concept however, there can be any number of variants of any given loanword. For example, English loanwords, such as (pronounced [kapet]), and (pronounced [kapit]), both originate from the English word *carpet*. This makes any type of search of loanwords problematic because they are not always found in a Korean or Korean-English dictionary.

While there are standard transliteration principles for foreign words more generally and loanwords that are contained in Korean dictionaries, writers often do not always respect these principles (Chung, 2002). Moreover, requiring writers to use standard transliteration principles for using any given foreign word is also unrealistic. Therefore, it is essential that a method must be utilized in managing these problems and their application. Without an effective methodology, the error rate in assigning a loanword or its variants to its original form in English, will be greater, and this in turn will have implications for ESL (English as a Second Language) learners.

To make matters more difficult, loanwords can have multinational origins. While some loanwords have their origins in English, others are derived from German, Russian, Japanese and so on. For this reason,

on any tool.

language users may encounter considerable frustration in finding the spelling and original language of any given loanword.

Such orthographic, phonological, and other forms of interference are a very common phenomenon in the situation of borrowing. Although, in some cases, new segments or sequences of foreign words are adopted without any modification, in most cases they tend to be modified in accordance with the phonology of the borrowing language (Chung, 1984 1986). Besides the phonological implications of English loanwords, deletions and abbreviations in their morphological and syntactical forms are often used. With respect to semantics, the meanings of original words are often expanded, contracted, or abased, and in this sense, such meanings can be shaped by factors such as psychological, linguistic, social, or historical.

With the aforementioned in mind, it is necessary to construct a positive attitude towards English loanwords and this can be facilitated through the use of a loanwords tool such as CALL³⁾ courseware or other such educational loanwords tools. A more important point to emphasize is that the use of these loanwords in various materials can be an effective device for vocabulary learning. When students learn loanwords in textbooks, the pronunciation of these words can be confused with that of English words. This occurs because the phonetic forms of English loanwords, pronounced by Korean speakers, can be imported from Korean phonological rules which might be entirely different to the phonetic rules for English.

This paper will trace the construction of the English Loanwords Contents (ELC) Database, a project which entailed trawling four dictionaries as well as hundreds of articles in Korean newspapers, for English loanwords. These loanwords were then entered into the ELC Database. The paper will also extrapolate on the use of the ELC Database for ESL learners in Korea, particularly with respect to the teaching of vocabulary. The culling of loanwords occurred in two distinct phases. The first, the Dictionaries Database, entailed the four

3) Computer-Assisted Language Learning

dictionaries and the second, the Newspaper Corpora Database. The two databases were then merged into what became known as the ELC Database.

In order to construct the ELC Database, over 35,000 English loanwords were culled from a variety of lexical sources and classified into 19 categories. Selection was based on four criteria: 1) words with variants in transliterating into Korean, 2) words presented in at least two sources, 3) words that were so changed in Korean that it was difficult to match them to their original English forms, 4) words used frequently in daily life.

A project two years in the making, the ELC Database now has approximately 35,000 word entries, the largest English loanwords database available on/off-line so far. It includes simple nouns, compound nouns, technical terms, and proper nouns, such as the names of people, geographical designations, titles of books, names of specific objects and so on. It is the first bilingual word-aligned Korean English loanwords database, comprised of Korean English loanwords and their original English words.

To expand the ELC Database, various types of Korean transliterations of English loanwords were constructed utilizing the UNITEX graph editor FSGraph, a software based on the Local-grammar framework. The UNITEX graph succeeded the first version which was INTEX developed by LADL (Laboratoire d'Automatique Documentaire et Linguistique) at University Paris 7.

The ELC Database is designed for an educational context, specifically to meet the needs of ESL learners. The objective is to provide a good foundation for students which would enable them to make further progress in their study of English. Although learning English can be rendered more interesting and exciting when loanwords are a focus of study, particularly for technology-savvy learners, it is imperative that more careful attention be paid to the narrow limits of these loanwords to avoid the gap that exists between an original English word and its corresponding English loanword deployed in Korean.

2. Construction of English Loanwords Database

The first step in constructing the ELC Database was to compile a list of English loanwords in Korean. This culling of English loanwords was most important for the construction of the Electronic Dictionary, which can recognize all possible variable loanwords in Korean. We used two sources mainly, the corpus and existing dictionaries.

2.1. Dictionaries Database

2.1.1. Compiling English Loanwords Database

First, from the existing dictionaries, we obtained about 35,000 loanword entries, the largest so far, which included simple nouns, compound nouns, technical terms, proper nouns (names of people, geographical designations, titles of books, names of specific objects and so on) and others⁴⁾.

The simple English loanwords nouns in Korean were extracted from both *The Dictionary of Simple Nouns in Korean* (Nam, 1994) and *The Standard Korean Dictionary on Hangeul* (2001). Over 7,300 English loanwords which are simple nouns were identified and classified into two categories, a 'NS' code and a 'EXT' code, which denote *The Dictionary of Simple Nouns in Korean* (1994) and *The Standard Korean Dictionary on Hangeul* (2001)⁵⁾. The code 'NE' denotes about 50 overlapping words in 'NS' and 'EXT'.

A total of 43,307 English loanwords proper nouns⁶⁾ in Korean were

4) 3,000 simple nouns, 7,400 compound words, 9,000 technical terms, 26,000 proper nouns. Overlapped words are not counted in each part.

5) 2,338 simple nouns of English loanwords are extracted from *The Dictionary of Simple Nouns in Korean* (1994) and 4,378 simple nouns are extracted from *The Standard Korean Dictionary on Hangeul* (2001). For further details of databases of English loanwords, see Cheon (2004).

6) 9,853 proper nouns from NJS, the symbol for *The Dictionary of Proper Nouns in Korean* (2003), and 26,738 proper nouns from PN, *Database of Nouns on Hangeul*.

extracted from *The Dictionary of Proper Nouns in Korean* (Nam, 2003) and *Database of Nouns on Hangeul*⁷⁾. The databases of the code 'PN-NJS' denote 10,336 overlapping words in 'NJS' (*The Dictionary of Proper Nouns in Korean* (2003)) and 'PN' (*Database of Nouns on Hangeul*). The following table indicates source, code, category, and the number of DB of English loanwords in detail.

Table 1: Source, Code, Category, and Number of English loanwords database

Code	Number	Category	Source
NS	30	common noun	The Dictionary of Simple Nouns in Korean
EXT	2,472	common noun	The Standard Korean Dictionary on Hangeul
NE	47	common noun	Overlapping words in NS and EXT
NJS	4,914	proper noun	The Dictionary of Proper Nouns in Korean
PN	18,323	proper noun	The Databases of Nouns on Hangeul
NJS-NS	1	proper noun	Overlapping words in NJS and NS
NJS-EXT	3	proper noun	Overlapping words in NJS and EXT
PN-NS	260	proper noun	Overlapping words in PN and NS
PN-EXT	1,399	proper noun	Overlapping words in PN and EXT
PN-NE	632	proper noun	Overlapping words in PN and NE
PN-NJS	4,888	proper noun	Overlapping words in PN and NJS
	Total 32,969		

2.1.2. The Headwords of Various Korean Transliterations

While the actual definitions of the present mark-way of borrowed English loanwords⁸⁾ were investigated, it was found out that they were marked differently in random samples of textbooks, dictionaries, as well as newspapers and magazines.

7) *Database of Nouns on Hangeul* does not published yet.

8) The mark-way of borrowed words means the principle of mark-way, which is unswerving and is in accord with the system of the natural language.

**Table 29): Frequency of the Use of Variants
in Transliterating of English Loanwords**

The Original Word	Variants	Frequency of the Use of Each Variants	The Original Word	Variants	Frequency of the Use of Each Variants
net		9	racket		12
		23			0
		1			9
napkin		4	leisure		4
		1			1
roket		5	gauze		3
		2		가	1

The second step in developing Dictionaries Database of English Loanwords in Korean entailed sorting the headwords of the various Korean transliterations of an English loanword. This was done using the criteria outlined in *The System of Romanization for the Korean Language*¹⁰⁾ -- rules of transcribing loanwords -- from the databases of four dictionaries used in this study. The following Table describes each DB code.

Table 3: Description of English Loanwords Code

(1) VAR-PN-EXT	The symbol for an English loanword, which is transliterated differently in PN and EXT.
(2) VAR-PN-NJS	The symbol for an English loanword, which is transliterated differently in PN and NJS.
(3) VAR-NJS-EXT	The symbol for an English loanword, which is transliterated differently in NJS and EXT.
(4) VAR-PN	The symbol for an English loanword, which is transliterated several ways in PN.
(5) VAR-NJS	The symbol for an English loanword, which is transliterated several ways in NJS.
(6) VAR-EXT	The symbol for an English loanword, which is transliterated several ways in EXT.

9) For further details of English loanwords database in this paper, see "Cheon (2005)".

10) published by National Academy of Korean Language (2001).

2.1.3. Loanwords and Their Original Words in English

The third step in constructing the Dictionaries Database of English Loanwords involved searching for the original English words of the English Loanwords. This entailed a sweep of English dictionaries as well as the internet and textbooks which might include technical terms that employed loanwords. It proved to be an arduous and tedious task which was nevertheless an indispensable step in constructing the ELC Database.

Table 4, below, shows the correct English form of loanwords that could not be found in several Korean or Korean-English dictionaries. What is essential to understand is that with globalization, the English language has made its way into the Korean vernacular, such that Koreans now identify the use of English as an integral part of their speech. Indeed, this phenomenon has its own Korean term, namely *Konglish*.

Table 4¹¹⁾: The Correct English Form of Loanwords

Loanword	The Correct English Form
	substitute
	a player's number
	rearview mirror
	window shade

The database shown in Table 4 above may well compromise the intelligibility of any given communicative encounter with foreigners.

2.1.4. The Classification of English Loanwords

The fourth step in constructing the Dictionaries Database necessitated identification and classification of 32,969 English loanwords into 11 categories. These categories were further divided into seven categories according to structure and additional information found in the database.

11) The source of the some databases: Ahn & Park (1994); Oh (1996); and Hong (1983).

Table 5: Description of Code

CODE		DESCRIPTION	CRITERION	EXAMPLE	NUMBER
SIMP	Simple Noun	The original word of loanwords containing one single English word	English	(alibi)	24,735 /32,138
COMP	Compound Noun	(1) The original words of loanwords containing two or more English words or including the symbols such as hyphen and more than one space (2) A compound noun made up of a. [English + English], b. [English + Korean], c. [Korean + English], d. [English + English + English], e. [English + English + Korean], and so on.	English	(1) (crayon pastel) (2) a. b. c. d. e.	7,403 /32,138
S-VAR	Source Language-Variableness	The loanwords that can be equivalent to a variety of English words	English	(cord, code, chord)	290 /32,138
S-INV	Source Language-Invariableness	The loanwords that have one original English word	English	(alcohol)	31,848 /32,138
REG	Regularity	The loanwords in Korean that can be found out the equivalent English words easily	Korean	(banana)	26,518 /32,138
EXC	Exception	The loanwords in Korean that go against students' expectation	Korean	(wafer)	5,620 /32,138
S-CUR	Source Language-Current	The loanwords in Korean that are in common use	Korean	(aspirin)	8,749 /32,138
S-RAR	Source Language-Rare	The loanwords that are not in common use	Korean	(spool)	23,389 /32,138
TERM	Term	The technical terms	Korean	(channel)	8,620 /32,138
CLP	Clipping	The loanwords that are shortened	Korean	(air conditioner)	100 /32,138
ABB	Abbreviation	A short form of a word and the process of abbreviating a word	English	(OS, Operating System)	1323 /32,138
FOR	Foreign Word	The loanwords that are from another language except English	The original language	(hof, German)	80 /32,138

2.2. Newspaper Corpora Database

To complete our actual English loanwords database we also observed the large Korean corpora. To do this, we began by constructing a corpus of reliable size, and then we had to decide between the use of ready-made corpus resources and the development of our own corpora. Before constructing the corpora, we took a random sample of Korean newspapers and we counted the number of English loanwords in each newspaper.

We found a total of 359 words, both Korean and English loanwords, in the international section of the *Chosun Ilbo*, for example, on June 23rd, 2003. Of that total, we found 40 English loanwords, or 11 percent. When we looked up these English loanwords in our Dictionaries Database, we found that 88 percent of them were listed; however, five loanwords, or 11 percent, were not listed. In terms of frequency of English loanwords by section in the *Chosun Ilbo*, Table 6 shows that the Sports section used English loanwords most often, while the Politics section used these words least often.

Table 6: Percentage of English Loanwords Covered by Dictionaries Database

SECTION	The Figure of English Loanword	Total words	Percentages of English Loanwords	English Loanwords Do Not occur in DIC-DB	Percentage	Percentage of English Loanwords Covered by DIC-DB
International	40	359	11.14%	5	13%	88%
Culture	33	726	4.55%	5	15%	85%
Culture	29	310	9.35%	4	14%	86%
National	15	330	4.55%	0	0%	100%
National	9	361	2.49%	0	0%	100%
Sports	27	317	8.52%	3	11%	89%
Sports	45	263	17.11%	9	20%	80%
Sports	31	179	17.32%	4	13%	87%
Sports	38	181	20.99%	13	34%	66%
Sports	64	150	42.67%	12	19%	81%
Politics	2	174	1.15%	0	0%	100%
Politics	3	230	1.30%	0	0%	100%
Local	15	288	5.21%	0	0%	100%
Local	7	174	4.02%	0	0%	100%

(Source: June 23rd, 2003. *Chosun Ilbo*.)

English loanwords can be found in many fields and disciplines as well as in popular culture. However, these loanwords are likely to be found in some areas more than others, such as economics, science, IT, and so on. Such fields are in a constant state of flux and with the new developments that are happening at a very fast rate, new English loanwords are also emerging which are used to describe technical terms. To establish and keep our Newspaper Corpora Database up to date, it is therefore absolutely essential that we trawl the fields mentioned above as often as possible.

To ascertain the frequency of English technical loanwords, we examined the economy section of two Korean daily newspapers, the *Chosun Ilbo* and the *Chungang Ilbo*. We chose this section specifically because we anticipated that we would find English loanwords to describe or explain technical terms. We found 44 articles in the corpora of these two newspapers between June 26th, 2003 and August 4th, 2003. The corpora was comprised of a total of 10,740 Korean words and English loanwords. Of that total, 844 English loanwords were found, or 8 percent.

To expand our Newspaper Corpora Database utilizing newspapers, we examined the *Yonhap News* corpora¹²⁾ and we culled a total of 9,975 English loanwords from articles dealing with IT and the economy. Excluding words we already found in the other newspapers, a total of 2,396 new loanwords emerged.

When we examined the loanwords culled from *Yonhap News*, the *Chungang Ilbo* and the *Chosun Ilbo*, we found an interesting phenomenon. There were many instances where various transliterations for the same English loanword were used. Table 7 represents various transliterations in Korean of one English loanword found in *Yonhap News* and *Chungang Ilbo* and *Chosun Ilbo*.

12) The size of this corpora is over 50 megabyte.

**Table 7: Different Transliterations
in *Yonhap News* & *Chungang Ilbo* and *Chosun Ilbo***

When we examined the corpora of all the newspapers mentioned above, we found a total of 3,435 English loanwords, 2396 from *Yonhap News* and 1039 from the *Chosun Ilbo* and the *Chungang Ilbo*. After we subtracted 408 English loanwords we found common to all three newspapers from the total, the number of new loanwords we found should come to 3,027.

However, the figure we arrived at was in fact 2,966 new loanwords. This is because we found it necessary to exclude 61 loanwords which had different transliteration for the same English original word; for example, the original English word *data* can be written in Korean in any number of ways, such as , or .

Table 8: The figure of Headwords from newspapers

TYPE	Yonhapnews	Chungang(Chosun) Ilbo	Overlapped words	TOTAL
Figure	2,396	1,039	408	2,966

The 2,173 English loanwords to which we have referred comprised 74 percent of the loanwords culled from the Newspapers Corpora. These loanwords were not listed in the Dictionaries Database which was constructed from the four dictionaries mentioned earlier. This demonstrates the importance of making good use of as many corpora as possible, especially texts using technical terms and proper nouns which tend to utilize English loanwords more frequently.

At this point it was necessary to ascertain the effectiveness of the data we had gathered from the Newspaper Corpora Database in terms of its inclusion of English loanwords in Korea. We decided to examine

all the English loanwords in the Economy section of the *Chosun Ilbo* for one day, September 3, 2003 and to compare our results with the Newspaper Corpora Database. The aim was to establish how many loanwords found on that day were already in our Newspaper Corpora Database.

We found a total of 778 English loanwords in all the articles in the Economy section of the *Chosun Ilbo* on that day. We then entered these words into our Newspaper Corpora Database and found 702 loanwords already accounted for. This demonstrated that our Newspaper Corpora Database was over 90 percent efficient. The 76 words that were not found in the Newspaper Corpora Database were then entered into the Dictionaries Database to test the viability of Dictionaries Database. All the words were found, with the exception of 11. Table 9 shows some of 11 words not listed in either in the Newspaper Corpora Database or Dictionaries Database.

Table 9: English Loanwords not registered on the Newspaper Corpora Database & Dictionaries Database

NIE ()
(가)
(power shift)
PBEC ()
(pension)

In conclusion, the Dictionaries Database comprised of 32,969 English loanwords, culled from the four dictionaries, was 70 percent effective in terms of inclusion of English loanwords in Korea. When the 2,173 English loanwords entered separately into the Newspaper Corpora Database were added to the total of 32,969 in the ELC Database, to reach a total of 35,142¹³⁾, total lexical coverage of loanwords was over 90 percent.¹⁴⁾

13) For further details of English loanwords database, see "Transcriptional English Loanwords Database (Seung-mi Cheon, 2004)".

3. Findings the Variants of Loanwords in the ELC Database

The core of the ELC Database, outlined in Part 3, consists of English loanwords and their variants. To accomplish the difficult task of finding the variants of existing Korean transliterations of the English loanwords already in the database, we applied UNITEX graph editor FSGraph, a software based on Local-grammar methodology, which was first built by LADL (Laboratoire d'Automatique Documentaire et Linguistique) researchers under the name of INTEX.

3.1. ELC Graphs

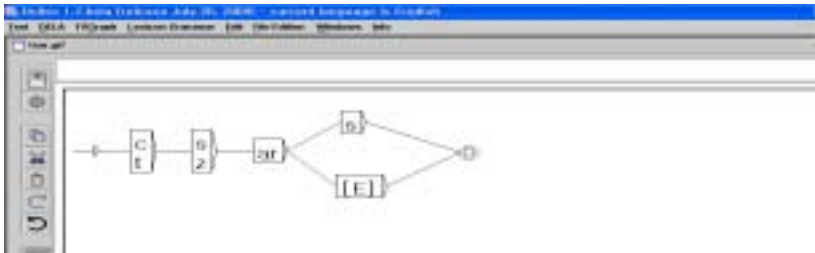
Finite state graphs (FSTs) were originally developed by INTEX software¹⁵). The more recent version is the UNITEX¹⁶) graph editor FSGraph. The 'input' part of an FST is used to identify specific sequences in texts; the 'output' part of an FST is used to perform substitutions in the text, to associate each identified occurrence with some information, or to enrich the text by inserting linguistic markers into it. In many cases, FSTs represent linguistic units more effectively than paper dictionaries. For example, FSTs may be used to bring together graphical variants of a word. Such an equivalence class can be used to check the spelling coherency of a word, as in the following graph:

14) This comes close to Laufer's (1989) estimate of a 92% lexical coverage in written texts, a minimum requirement for reasonable L2 reading comprehension. This pedagogic suggestion is also endorsed by Sutarsyah, Nation & Kennedy's (1994) corpus-based case study of the two corpora, one from a university textbook and the other from Section J (Learned and Scientific texts) from the LOB corpus and the parallel Wellington corpus.

15) Useful information on the system INTEX is given by Silberztein, 1999.

16) UNITEX program allows users to rapidly construct, test and maintain descriptions of specific patterns that occur in texts written in natural language. Each description is represented by a local grammar, usually entered via the UNITEX graph editor.

Figure 1. The Composition of FST (*csar*)



The above graph represents a finite-state automaton. It has one initial state (left-most arrow) and one final state (right-most square). This representation has been devised in order to facilitate the effective construction of grammars by linguists¹⁷⁾.

It departs from classical representation in that states are not overtly represented. The nodes are not the states and there is no symbol for them. [E] stands for the null word. After applying this graph, the eight variants (=2*2*1*2) of forms, *csar*, *czar*, *tsar*, *tzar* and *csars*, *czars*, *tsars*, *tzars*, will be indexed with the same key *tsar* (meaning Russian Emperor).

A more elaborate example of an FST graph is . One of the variants can be tagged as , .NFS(carpet)CON;NV¹⁸⁾. It can be inferred from the above example that these 18 forms (=1*2*3*3*1), , , , and others like them, can be indexed with the same key *carpet* on applying this graph to any e-dictionary.

17) M. Silberztein (1993) has written a graph editor for this purpose: FSGRAPH in INTEX.

18)

Code	Description
	variants
	headword
NFS (carpet)	+Noun, +Foreign, +Simple Noun (the original English word)
CON	Concrete Noun
NV	Noun ended with Vowel

Figure 2. Graph of *carpet*

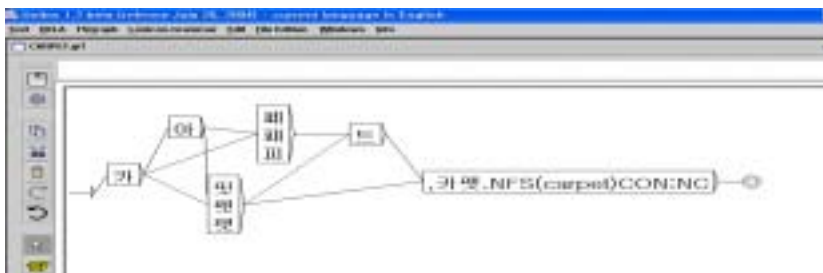


Table 10: Variants of Transliteration of *carpet*

TAG	
,	.NFS(carpet)CON;NC
,	.NFS(carpet)CON;NC
,	.NFS(carpet)CON;NC
,	.NFS(carpet)CON;NC
,	.NFS(carpet)CON;NC
.....	

The FST graph proved to be the most efficient method of obtaining English loanwords variants between a graph (Figure 2) and lists (Table 10). Obtaining the variants by means of the FST graph was unquestionably less painstaking than doing this work manually. From our actual database, about 2,500 English loanwords were recognized as having more than two variants of transliteration.

We therefore focused on these variants in order to complete the whole ELC Database. All the possible variants for each loanword are represented under one graph, equivalent to a finite-state transducer (FST), using the UNITEK graph editor FSGraph. These graphs can be merged automatically into one huge graph in the UNITEK system and in this way constitute the main part of our ELC Database.

4. The Role of ELC in the Classroom

There is no doubt that ELC can make a valuable contribution to the

ESL classroom. ELC falls into the category of CALL software, so popular in the modern classroom. ELC could be a very useful tool for teaching vocabulary. The critical point here is that vocabulary is more meaningful when learners are invited to explore words that have authentic meaning in their everyday lives. This is particularly the case with young people, for example, who might be overwhelmed attempting to learn English but can become suddenly motivated when they find their favorite PC Bang words, such as Star Craft or Lineage in the classroom. In this way the learning of English is more accessible.

4.1. English Loanwords in Teaching Vocabulary

English loanwords found in various materials, can be an effective device for English language teaching. This methodology which can employ the loanwords list has the following advantages:

1. Learners can be more motivated to learn English because they are already familiar with English loanwords used in the course of speaking Korean.
2. Students have a better understanding and new affection for the Korean language through English learning (Yoon, 2002).
3. Teachers may find the vocabulary list of loanwords an invaluable asset on the textbook list for beginner classes, in that newcomers to English will feel less daunted by the language when they find loanwords they recognize.

The pronunciation of English loanwords is a key issue in teaching vocabulary. When students learn English loanwords from their textbooks, the pronunciation of loanwords is often confused with that of English words. This is because the phonetic form of English loanwords, when pronounced by Korean speakers, is governed by Korean phonological rules which may be different to those that apply to English (Chung, 1988).

One study on the misaccentuation of the English loanwords spoken by speakers of Korean shows that most speakers of Korean tend to pronounce English loanwords in the same kind of pitch contour as that

found in Korean words, rather than in the English contour (Kim & Moon, 1998). To find types of pitch contour of English loanwords, ten native speakers of Korean were asked to pronounce 120 words, which were then analyzed by CSL (Model 4300B¹⁹). This study shows that the types of pitch contour are generally determined by the CV structure of the syllables. Koreans have a tendency to pronounce English in a Korean way in that they usually insert a certain vowel into a cluster of English consonants. Koreans tend to pronounce the syllables in each English word individually, while native speakers speak rhythmically, almost ignoring unstressed syllables.

5. Limitations and Possibilities for the ELC Database

Several design considerations and technical restrictions were evident in the development of the ELC Database. They are as follows:

1. Because ELC was designed as an educational tool, it is necessary to test it in an educational environment. Feedback from students would be invaluable in the form of questionnaires, observation, and discussion.
2. Because of the rapid changes which occur in English loanwords lexicon, particularly in regards to science and technology, it is imperative that the database be updated on a regular basis.
3. The ELC Database with its capacity to store an enormous amount of data on English loanwords can be invaluable tool in the classroom. It can be used to supplement reading texts and grammar references in the classroom and it can also be useful for ESL learners working alone. This database system can save teachers time and effort and it can be worked into the curriculum in such a way that students are motivated to learn English because English loanwords have meaning in their everyday lives.
4. A variety of computer-adaptive vocabulary comprehension tests need to be developed to measure foreign or second-language

19) See Kim & Moon (1998) for a full account of "CSL."

- proficiency with appropriate complementary evaluation instruments.
5. The low number of loanwords and their variants can also be a limitation. The ELC Database has a total of 35,142 English loanwords. Of this total, only 2,500 have been treated to decipher their variants. This is, of course, a very limited number. It is very simple, however, to explore the variants of additional loanwords by means of the *Automatic Foreign Word Transliteration Model*, developed by Lee (1999).

6. Conclusion

English loanwords are encroaching on the Korean language at a very rapid rate, such that they have almost the same status as the Korean language. The fact of the matter is that Korean language users, bombarded with English language and culture, choose to borrow English words when Korean vocabulary cannot adequately accommodate itself to new concepts and terms. When foreign words are received in Korean, however, their pronunciation, form, and meanings are often manipulated to conform to Korean phonological rules. Such linguistic strategies sometimes result in the English words being changed beyond recognition. It would appear that this transcription has been a law unto itself, and indeed this may be the case considering the limited success of the Korean government's efforts to standardize the transcription of English words into Korean (Chung, 2003).

In spite of the importance of teaching English loanwords to Korean EFL students, it is rather disappointing and discouraging that there is a paucity of research in this area over the last decade. With regard to L2 vocabulary acquisition, for example, very few studies have been reported in Korean EFL research literature (Jung, 2000; Lee, 1996). This is in sharp contrast with a resurgence of interest in this field of research in other countries in the last few decades (Carter & McCarthy, 1988; Coady & Huckin, 1997; Hatch & Brown, 1995; Huckin, Haynes, & Coady, 1993; Nation, 2001; Read, 2000; Schmitt, 2000; Schmitt &

McCarthy, 1997; Singleton, 1999). It is clear that more research needs to be conducted into English loanwords to understand their significance in facilitating the education process of Korean EFL students.

Over the past few years, a considerable number of studies have explored the transliteration of foreign words into the Korean alphabet system and there has been considerable discussion about this issue. However, very few attempts have been made to address the practical use of English loanwords in teaching English. It is for this reason that this ELC Database was constructed.

This paper has presented an account of the construction of the ELC Database and the complexities of this task in terms of English loanwords and their variants. The paper has also examined the value of the ELC Database as an educational tool for English language learners, particularly its capacity to motivate students who are already familiar with many English loanwords in Korean.

Although the focus of this paper has been on the use of the ELC Database in education, it is important to point out that the ELC Database has many possibilities beyond this sphere. Further investigation needs to be done into the use of this database in many other fields; for example, machine translation systems, natural language processing systems, and information retrieval systems, to name a few. The ELC Database might be particularly useful in improving the quality of information retrieval and machine translation systems by assigning the various transliteration of English loanwords to their original English forms. These are possibilities for further research.

References

- Ahn, B. G., & Park, H. M. (1994). Studies of modern Japanese loan word. *The Japanese Culture Association of Korea*. Vol.1994 No.1
- Carter, R., & McCarthy, M. (1988). *Vocabulary and language teaching*. London: Longman.
- Cheon, S.-M. (2004). English Loanwords Database. TR-04-05. *DICORA*.

- Cheon, S.-M. (2005). *Construction of English loanwords contents for the development of educational tools: a step towards the prosperity of CALL courseware*. Hankuk University of Foreign Studies. *Chosun Ilbo*, June-September. (2003).
- Chosun Ilbo*, September, 23rd, (2004)
- Chung, K. (1984). Five Problems in Korean speakers' Pronunciation of English. *Korean journal of Linguistics* 10.
- Chung, K. (1986). The Recognition and Reproduction of Foreign Sounds; A Generative Approach. *Korean journal of Linguistics* 12.
- Chung, K. (1988). Phonological Theories and the Teaching of Foreign Language Pronunciation. *The Journal of English Language and Literature* 34. 2.
- Chung, K. (2002). Problems in transcription of foreign words/loanwords. *Studies in Foreign Language Education* 34. Hankuk University of Foreign Studies
- Chung, K. (2003). Transcription of loanwords and pronunciation of foreign languages. *Studies in Foreign Language Education*, 17. Hankuk University of Foreign Studies.
- Coady, J., & Huckin, T. Eds. (1997). *Second language vocabulary acquisition: A rationale for pedagogy*. Cambridge: Cambridge University Press.
- Hatch, E., & Brown, C. (1995). *Vocabulary, semantics, and language education*. Cambridge: Cambridge University Press.
- Hong, Y.-S. (1983). Problems in Japanese loanwords. *Mil-mool*. Vol.1 No.1.
- Huckin, T. N., Haynes, M., & Coady, J. (1993). *Second language reading and vocabulary learning* Norwood, N.J.: Ablex Pub. Corp.
- Jung, Y.-S. (2000). The impact of diglot reader method on vocabulary acquisition. *English Teaching*, 512, 177-199.
- Kim, S.-M. & Moon, S.-M. (1998). A Note on Prosodic Differences between Korean and English in loan words from English. *Malsori* 35, 25-36.
- Lee, J.-S. (1999). *An English-Korean transliteration and retransliteration model for cross-lingual information retrieval*. Korea Advanced

- Institute of Science and Technology.
- Lee, H.-J. (1996). An effective English teaching method for secondary school students. *English Teaching*, 512, 77-105.
- Nam J.-S. (1994). *The Dictionary of Simple Nouns in Korean*. Dictionnaire Des Noms Simples Du Coreen. LADL, Universite Paris 7 - CNRS.
- Nam J. S. (2003). The Dictionary of Proper Nouns in Korean. Technical Report 03-03. *DICORA*.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge Univeristy Press.
- National Academy of Korean Language. (2001). *The System of Romanization for the Korean Language*.
- National Academy of Korean Language, (2001). *The Standard Korean Dictionary on Hangeul*.
- National Academy of Korean Language. *Database of Nouns on Hangeul*. (does not published yet)
- Oh, J.-R. (1996). *Acronyms in Japanese loanwords*. Seoul: Hyung Seol Publishing Company.
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge Univeristy Press.
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- Schmitt, N., & McCarthy, M. Eds.. (1997). *Vocabulary: Description, acquisition and pedagogy*. Cambridge: Cambridge University Press.
- Silberztein, M. (1999). Text Indexation with INTEX. *Computer and the Humanities*, 33. Kluwer Academic Publishers: Amsteram
- Silberztein. Max D. (1993). Dictionnaires electroniques et analyse automatique de textes. *Le systeme INTEX*, Paris, Masson, 234 p.
- Singleton, D. (1999). *Exploring the second language mental lexicon*. Cambridge: Cambridge Univeristy Press.
- Yang, W.-M. (2002). *The characteristics of borrowed words in magazines and advertisement*. Seoul: Han Yang university Press.
- Yonhap News*. (1998-2003).
- Yoon, H.-S. (2002). The use of loanwords for enhancing word recognition

in elementary school English.

Seung-mi Cheon

College of English, Hankuk University of Foreign Studies

270 Imun-dong, Dongdaemun-gu, Seoul 130-791, South Korea

Phone: 82-2-6227-8109

Email: smcheon@hufs.ac.kr

Received: 30 Mar, 2005

Revised: 17 Jun, 2005

Accepted: 19 Jun, 2005