Australian Students' Lexical Errors in Korean: Type, Frequency and Cause

Author:
Shin, Seong-Chul

Australian Students' Lexical Errors in Korean:

Type, Frequency and Cause[*]

Seong-Chul Shin

*School of Languages and Linguistics, Faculty of Arts and Social Sciences,
University of New South Wales, Sydney, 2052. s.shin@unsw.edu.au*

신성철. 2002. 호주 한국어 학습자의 어휘 오류 분석 연구. *Journal of Korean Language Education* 13-1: 307-338. 본 연구는 영어를 제 1 언어로 하는 호주내 한국어 학습자들이 빈번하게 범하는 어휘 오류가 어떤 것들인지 알아보고, 이들 어휘 오류를 형태와 빈도에 따라 분류해보며, 그런 오류들이 일어나는 원인을 잠정적으로나마 밝혀 보려는 것이다. 영어 원어민 한국어 학습자 71 명 (대학 2, 3 학년 학생들)이 작성한 정규 시험답안지 141 개를 조사하였고, 이 가운데 Corder (1981:38-9)에 의거하여 자유작문식 텍스트와 재구성형식 텍스트로부터 총 305 개의 어휘 오류를 추출하여 분석하였다. 분석 결과, 어휘 오류를 11 개의 오류 형태와 4 개의 품사별로 분류하고 각각 빈도수에 따라 그 순위를 알아보았다. 분석 결과에 대해서 학습자의 한국어 수준과 작문의 난이도라는 두 가지 측면에서 해석을 시도해 보았다. 오류의 원인으로서는 그 형태를 분석해 볼 때 크게 영어-한국어의 언어간 요인과 한국어의 내적요인, 복합적 요인, 기타 개별 요인 등 네 가지 범주로 묶을 수 있었으며, 전체적으로는 언어간 요인보다는 한국어 자체의 내적요인에 기인한 오류가 많은 것으로 나타났다. 이를 토대로 본 연구는 언어습득이론, 교육과정 및 교재, 교수자와 학습자 등 네 가지 측면에서 한국어 교육의 효율성을 논하였다. (뉴사우스웨일즈대학교)

주제어: 오류분석, 어휘 오류, 한국어 어휘 교육

1. INTRODUCTION

This paper is based on the result of a pilot study, which was carried out as part of a larger project regarding error analysis, whose aim is to investigate the

---

key lexical areas of difficulty for Australian students of Korean as a foreign language (KFL). Specifically, this study intends: 1) to identify the lexical features that present particular difficulties to English native speakers learning Korean; 2) to classify those lexical errors in terms of their type and frequency; and 3) to provide possible explanations for the cause of those lexical problems.

The study of learner errors has been a part of language pedagogy for a long time. Language instructors are constantly concerned about the errors made by their students and with the ways they can improve language teaching. Error Analysis (EA) as a method of the study of errors played a new role in second language (L2) acquisition research in the 1970s. EA became the principal methodology used for investigating learner language and L2 acquisition, supplanting Contrastive Analysis (CA) method. This was mainly due to the weakness of CA and the desire to improve pedagogy through the study of errors (Corder 1975, 1981). CA looked only at the contrastive characteristics of two languages: the first language (L1) of the learner and the target language (TL), while EA is concerned with the learner language and the process of language learning. The significance of errors is well documented in Corder (1967), whose work is regarded as a major contribution to the early development of EA.[1]

Corder (1981: 36) suggests three key steps of EA research: 1) identification of errors, 2) description of errors, and 3) explanation of errors. In his other paper (1974 cited in Ellis 1994), Corder suggests two additional steps that should be completed before and after the key steps: collection of samples and evaluation of errors. These five steps constitute a general framework of EA studies, though the last step, the evaluation of errors, is often handled separately. Among the steps, explanation for cause of lexical errors is regarded as the most important stage as it deals with how and why errors are made, which will then give insights into what and how a learner learns (i.e. the process of L2 acquisition).

There have been a number of studies identifying different causes of lexical errors. Although the terms used in each study may be different, the method of

---

determining the source of error is similar and is largely applicable to different linguistic levels, including the lexical level, with some specific modifications. It has been common to categorize errors into three or four general types: interlingual (transfer), intralingual, unique (e.g. induced) and/or other types of error e.g. developmental and communication-strategy (Coder, 1975, 1981; Ellis 1994; James 1998). Interlingual Errors, which are generally referred to as Transfer Errors, occur when the learner uses L1 features rather than those of L2. In other words, errors in this category are largely caused by the learner using their first language's structure and applying it to the target language (i.e. L1 interference). Intralingual Errors, on the other hand, reflect the complex characteristics of the target language and arise when the learner fails to comprehend fully conditions under which its rules and restrictions apply. Overgeneralization is a good example of such an error type. Induced Errors (Ellis, 1994) refer to those errors made because of inappropriate instruction or instructional materials, while Developmental Errors (Richards, 1971b) occur when the learner falsely hypothesizes rules and concepts on the basis of earlier learning experiences, and thus this reflects on the stage of his/her language development. Communication-Strategy Errors (James 1998) arise when the learner attempts to use an approximate form of the required word or an indirect expression called *circumlocution*.

It is not easy to distinguish between interlingual and intralingual errors, and it is even more difficult to determine whether an error type is interlingual or intralingual, or some other category, as errors can be caused by a number of different reasons. Some studies (for example, George 1972 and Sohn 1986) report that interlingual transfer errors are more universal and frequent than intralingual errors, and others (for example, Dulay and Burt 1974b, Taylor 1975 and Wang 1995) find a higher proportion of intralingual and developmental errors in learners at a particular level. However, this discrepancy in the source of errors seems to be quite understandable because expected results will vary according to various factors. For example, the task used to elicit the samples, the linguistic level or area that was investigated and the profile of subjects used in

the error studies. It is not difficult to claim that even within the same or similar conditions, the findings can be different according to the focus and method of the analysis.   In fact, error classification has been generally an arbitrary matter, relying largely on the researcher's individual point of view, and thus the findings can be considerably different from study to study as to proportion of the error types and their causes.   Nevertheless, there are some good pedagogical values to investigating errors and establishing the sources, though tentative, by using EA as a tool.

Error analysis research in Korean is relatively new and small, reflecting the education history of Korean as a foreign or second language.   It was very recent (late 1980s through 90s) that EA became a recognised part of KFL research areas, a development that came with the establishments, expansion and /or consolidation of KFL programs in several universities in Korea and abroad. Researchers now seem to be getting more engaged in EA research, in an attempt to discover more about L2 (Korean) acquisition and to improve pedagogy based on the findings. Thus far, the majority of EA research outputs on KFL are based on errors produced by English and Japanese speakers learning Korean in Korea. Studies on errors by English speakers include Sohn (1986), 왕혜숙 (1995) and 김미옥 (2001). And studies on errors by Japanese speakers include 김정숙 (1988), 이수경 (1996) and 최우영 (1997), while 전은주 (1994), 김미옥 (1994) and 김유미 (2000) examine errors by both English and Japanese speakers. Other studies include 김영아 (1990) analysing errors by Chinese speakers and 이정희 (2001) dealing with errors by a multiple language background group. Some of these studies contain cross-sectional analysis but most of them focus on one or two linguistic areas. For example, Sohn (1986) presents error patterns at six linguistic levels including orthography, syntax, morphology and lexicon, while 김정숙 (1988) focuses on listening-based phonological errors, 최우영 (1997) on syntactic errors, 김유미 (2000) on case markers and 이정희 (2001) on tense errors.

The main focus of this paper is to analyze and discuss lexical errors. James (1998: 142-54) provides five good reasons[2] for undertaking lexical EA and summarizes (suggestions cited from Richard 1976) the seven characteristics of a lexical item: 1) its morphology including spelling and pronunciation, 2) its syntactic behaviour, 3) its functional or situational restrictions, 4) its semantic value(s), or denotations, 5) its secondary meaning or connotations, 6) other words it is associated with, and 7) its frequency. For classification of errors, he suggests the formal vs semantic dichotomy and subdivides the categories. Under the formal error category, he distinguishes three sub-categories: formal misselection, misformations and distortions, while in semantic errors, he suggests two main types: confusion of sense relations and collocational errors. In each category, a number of specific error types are identified. (A summary of his classification is given in the footnotes.)[3] James sees the source of formal misselection and misformation errors as either interlingual or intralingual, and distortions as intralingual, while confusion of sense relations are intralingual, and collocational errors are either intralingual or interlingual.

As discussed above, there are a number of different ways to describe, classify and explain L2 learners' errors. Some studies employ a general structure of EA research and modify it to meet its own condition, while others adapt a more specific one prepared for a particular linguistics level and task. For a lexical level, the outline proposed by James is a good starting point. Although it is not the intention of this pilot study to modify his framework, I find it more useful and relevant for a detailed study of lexical errors.

---

[2] 'Morphological aspects of words, which used to be treated as part of grammar, can just as well be viewed as part of the word', 2) 'learners themselves believe that vocabulary is very important in language learning, sometimes equating a language with its vocabulary', 3) 'for some learner groups, lexical errors are the most frequent category of error', 4) native speakers consider the lexical errors in learners' IL [interlanguage] to be more disruptive and irritating than other types', and 5) 'vocabulary carries a particularly heavy functional load, especially in early IL.' (pp. 143-4)

[3] Formal misselection: suffix, prefixing, vowel-based and consonant-based, 2) Misformations: borrowing, coinage and calque, 3) Distortions: omission, overinclusion, misselection, misordering and blending, 4) Confusion of sense relations: use of a more general term, use of too specific a term, use of the less apt of two specific terms, and use of the wrong near-synonym, and 5) Collocational errors: semantic-bound selection, statistical preference and arbitrary combination (see James 1998: 144-54 for definitions and examples).

Now we turn to a couple of specific error studies of the Korean language. As in other levels of EA research, much less attention has been given to lexical errors in Korean. Two studies with subjects having English as their L1 are briefly reviewed. Sohn (1986) observed compositions written by second- and third-year American students of Korean and classified the lexical errors into four categories according to word class and usage, namely, errors in nouns, verbs, adverbials and Sino-Korean collocations. By giving examples in each category, he finds a number of different causes of errors, which include wrong choices of words, interference from English, poor knowledge about semantic restrictions and overgeneralization. His classification is simple as it is based on word class of the errors, rather than the type of errors, but he offers intuitive linguistic explanations of the different causes. What seems to prevail in his findings is interlingual transfer errors. He claims that interference from English accounts for problematic features such as missing nouns, unnatural expressions from translation, confusions in the use of transitive and intransitive verbs, partial or non-use of some action and existential verbs, and confusions between the existential verb and the copula and in the use of psychological verbs.

왕혜숙 (1995) analysed 224 lexical errors from 40 intermediate level compositions written by American students of Korean and classified them into eight types of errors.[4] Although her classification was adapted from previous studies based on European languages (e.g. Larsen-Freeman and Long, 1991), it seems to fit reasonably well into Korean as well, with some modifications. Wang finds that the most frequent are lexical shift / code switch, confusion of similar meaning, overgeneralization and collocation / idiomaticity in the order. For a bigger picture and the source of errors, she groups the type of errors into three categories: Intralingual, Interlingual and Combination of both, and finds that 51% of the errors are attributable to Intralingual and 34% to Interlingual. In regards to word class errors, nouns were most frequent (42%), followed by

---

[4] These are errors caused by confusion of similar meaning, errors caused by formal similarity in TL, lexical shift / code switch, collocation / idiomaticity, word coinage, simplification or redundancy, overgeneralization, and literal translation.

verbs, adjectives and idiomatic expressions. Her findings, in which intralingual errors were more common, were contrary to previous studies such as Sohn (1986), where interlingual errors were more frequently mentioned.

In the present study, I will define the terms of error types that are used and necessary for the classification of errors by modifying them from other studies (e.g. 왕혜숙 1995) and refining them to fit the current study. Based on the classification, the frequency of errors has been examined, first as a whole, by comparing the data sources, and then by the word class.   Some high frequency and 'unusual' errors have been chosen for detailed description and explanation. A brief discussion is offered for pedagogic strategies to be used in the teaching and learning of Korean in the English-speaking environment.   The result of this study will make a partial but useful contribution to the field of Korean lexical error analysis, in particular for the construction of a working hypothesis, a systematic error classification and an effective pedagogical method.

2. METHOD

2.1. Subjects

The subjects selected in this study are 71 second-and third-year students from three universities (identified as G, M and N).[5] They are native speakers of English or are believed to have English as their first language. Those who are not considered to be native speakers of English have been excluded from the subject selection process. The subjects have learned Korean as a foreign language for about two to three years. Among the 71 students in this study, 26 are third-year students and 45 are second-year students.

2.2. Data

The data used in this study come from written examination papers administered during the First and Second semesters in 1999 and 2000 at the

---

[5] The majority of them come from 'N' and 'G' Universities. 'M' data are very small and only come from the work of 2nd Year students.

three universities. To maintain the reliability, the data have been obtained only from short written, mid-semester and end-of-semester examination papers and do not include homework-type data such as worksheets and take-home essays. Based on the above-mentioned ground, 141 papers were initially selected to identify or detect errors. To complement the constraints of textual data and thus to increase the validity of the findings, two classes of composition data were selected: *free composition* data and *reformulation* data (Corder 1981:38-9).[6] Accordingly, the textual data selected for analysis come from short answer questions, translation tests and free/ essay composition tests. In all, 305 lexical errors have been identified for analysis, of which 197 come from N, 97 from G and 11 from M data.

2.3. Procedure

To identify lexical errors, only content words and their meanings were taken into consideration, thereby excluding other levels of errors such as orthographic, morphological and syntactic errors, which are separately dealt with in other papers. After identifying the lexical errors, the phrases and sentences containing errors were listed along with correct forms and words. Each error was carefully examined according to its nature, and coded by error type by utilizing classification methods used in previous studies (e.g. 왕혜숙 1995; Sohn, 1986). In this process, every care was taken to make sure that each error was appropriately classified, which required a painstaking decision-making process as they could be assigned to one of several error types. This was partly because the definition of each error type used in other studies was somewhat ambiguous, partly because some errors had more than one nature of cause, and partly because it was difficult to accurately pinpoint the intention of the student or how he/she came to obtain such an erroneous lexical item.

---

[6] According to Corder, it is those which represent the learner's attempt to reformulate, in one way or another, the ideas and intentions of others (e.g. translations, resumes and retelling of stories).

2.4. Error Types

In this study, an attempt was made to refine the definition of each error type, and the errors identified have been classified into eleven types based on the working definitions as follows:[7]

(1) Errors of wrong word choice - where a wrong lexical item is chosen in place of the correct one, and by having that item there, the whole sentence does not make sense at all. This happens particularly when the student selects a wrong or inappropriate item from the dictionary entry, which lists a multiple of L2 (i.e. Korean) equivalents.

Ex) a. Meon.jeo *hyu.il (hyu.ga reul)* ga.go.sip.seum.ni.da. (recreational holidays)

b. Ho.ju.sa.ram.eun a.ju *sa.hoe.jeog... (sa.gyo.jeog...)* (to be social)

(2) Errors of literal translation - where a lexical item is literally translated by sticking to its literal meaning or the way the student's L1 (i.e. English) is expressed.   By having that sort of item there, the message is understood but it sounds awkward at best. This normally arises when the student literally transfers the individual meaning of an item without knowing the set expressions or term-like equivalents.

Ex) a. Sae chin.gu.deul.do *man.deul.go (sa.gwi.go)* sip.eo.yo. (make friends)

b. Yeol.se.sal jjeum e go.deung.hag.gyo *si.jag.hae.yo (deul.eo.ga.yo)* (to enter)

(3) Errors of omission or incompletion - where a lexical item, which should be present, is omitted or some lexical element, which should be complete, is

---

incomplete. With a missing lexical item in the sentence, it only partially makes sense or it sounds unnatural and incomplete. This often happens when the missing or incomplete item is not so important or necessary in the student's L1, English.

Ex) a. Yeo.reum.eul (i) *je.il (je.il joh.a.ha.neun)* gye.jeol i.ra.seo.. (my best/ favourite)

b. Dae.hag.gyo.e.seo *chi.mi ga (chi.mi.saeng.hwal eul)* mahn.i …(hobby activities)

(4) Errors of semantic similarity - where a lexical item with a similar definitional or semantic element is used, and such an item does not fit precisely with another pair of item in the sentence, though communicable and sometimes broadly acceptable. This type of error is often caused when the student is confused by two words of similar meanings as they share some semantic features.

Ex) a. Jeo.neun *ja.sig.eun (a.i.deul.eul)* an joh.a.ha.ni.kka (children)

b. Ho.ju.neun in.gu.ga *jag.a.yo (jeog.eo.yo)* (small in number)

(5) Errors of overgeneralization - where a lexical item or items learnt in the earlier learning sequence are overly applied to other target situations producing unnatural or deviant expressions. This type of error usually arises when the student mistakenly generalizes the use of the item or finds no other item in his/her knowledge and so creates an inappropriate item or a set of items on the basis of other lexical items already learnt in the target language.

Ex) a. *I.nyeon.e (ol.hae)* gal gye.hweg.i.eoss.ji.man (this year)

b. Ga.jog (gwa) gat.ji (i) sal.a.seo *jib.bi (jib.se)*.neun eobs.eo.yo. (rent money)

(6) Errors of idiomatic collocation - where a lexical item used in a sentence is not matched or collocated with another pair of lexical item in the sentence, making the whole sentence unnatural or inappropriate.   This type of error is generally caused by the student's literal conversion of the particular item into L2 (Korean), without knowing the matching item required by the idiomaticity of or the concord relationship between the pair items.

Ex)   a.   Si.gan.eul *ju.syeoss.seo (nae.ju.syeo.seo)* gam.sa… (for making it available)

b.Eon.je.na *sig.sa.reul yo.ri.ha.go (jun.bi.ha.go)* (to prepare /cook a meal)

(7) Errors of code-shifting - where a L1 (English) word or lexical item is used instead of a L2 word or item.   It may be directly switched to a L1 spelling or transcribed in L2 letters.   This happens mostly because the student cannot find the L2 word or lexical item in his/her knowledge.   This may arise because he/she wrongly assumes that it is used as a loan word in the target language community or because he/she frequently switches codes between languages.

Ex)   a.   *Sportscentre (seu.po.cheu.sen.ta)*.neun   un.dong   ha.il.la.i.teu.reul bo.yeo….

b.   Ga.kkeum   ta.ni.kka   *ig.sa.i.ting   hab.ni.da   (jae.mi.iss.seub.ni.da)* (exciting)

(8) Errors of word coinage - where a newly invented lexical item or phrase is not appropriate enough to form a matching or set phrase.   In this type of error, the phrase usually sounds unnatural or redundant.   This type of error arises when the student is aware of the individual words but unaware of concise pair expressions or phrases.

Ex) a. *Nong.sa ma.eul.e.neun (Nong.chon.e.neun)* (in the agricultural community)

b. *Il.ha.go jib.saeng.hwal.i. (jig.jang saeng.hwal.gwa ga.jeong.saeng.hwal.i)* (working life and family life)

(9) Errors of redundancy - where a lexical item or its constituent is unnecessarily repeated or paraphrased. Accordingly, the phrase sounds repetitive, redundant and unnatural. This type of error is often found when students are unaware that the preceding word contains the semantic meaning of the subsequent word, especially when they attempt to make a combined word group or phrase such as a combination of Pure-Korean and Sino-Korean words.

Ex) a. Han.gug. *sang.sa hoe.sa.e.seo (sang.sa.e.seo)* (in a trading company)
b. Han.gug.eun *Ho.ju na.ra (Ho.ju)* bo.da jag.go (Australia)

(10) Errors of Sino-Korean numeral collocation - where a Sino-Korean (SK) and pure Korean (PK) lexical items are wrongly mixed or collocated, particularly in numeral compounds. With this type of error, the phrase may be able to deliver the intention of the writer but it often hinders the fluency and naturalness of the phrase. This sort of error is often caused when the student is unaware of the match or mismatch between SK and PK compound items.

Ex) a. Cho.deung.hag.gyo.e *il.gob.nyeon.e (chil.nyeon.dong.an)* … (for 7 years)
b. Nae.nyeon.e *il.dal.e.seo (il.weol.e)* jeo.neun. gyeol.hon.hal… (January)

(11) Errors of formal similarity, where a wrong lexical item is used due to the formal similarity in the phonetic or orthographic aspect in L2. This type of error is differentiated from orthographic errors in that the error item

carries a correct form and a sensible meaning in itself but semantically it is a completely different lexical item. Generally this sort of error might be caused due to the student's confusion or insufficient knowledge of the target item in both form and meaning.

Ex) a. Gil.eul *ij.eo.beo.ryoss.eul tae(ilh.eo.beo.ryoss.eul tae)* (when someone is lost)
b. Yeo.haeng.eul joh.a.ha.ni.kka *gwang.go.hak.eul (gwan.gwang.hak.eul)* jeongn
(jeon).gong.hae.yo (tourism study)

3. RESULTS

3.1. Frequency of Error Types

In all, 305 lexical error items were selected and analysed. The most frequent four error types were errors of wrong word choice, semantic similarity, overgeneralisation and literal translation in the order of frequency. The least frequent types were code-shift, redundancy, Sino-Korean numeral collocation, word coinage and idiomatic collocation or concord in the order.

Table 1. Frequency of Error Types As a Whole

Wrong word choice	85 (N)	26 (%)
Semantic similarity	45	14
Overgeneralization	39	12
Literal translation	33	10
Formal similarity	23	7
Omission or incompletion	18	6
Word coinage	15	5
Idiomatic collocation	15	5
Sino-Korean numeral collocation	14	4

| Redundancy | 10 | 3 |
| Code-shifting | 8 | 2 |

As Table 1 shows, students were having more difficulties in selecting words appropriate for particular contexts or situations and in differentiating lexical items with similar meaning. They also tend to over-generalize the lexical item that they have learnt and to literally translate the reformulated item or the item stored in their L1. However these errors occurred, it seems that these four error types have something in common, that is, the errors primarily associated with the definitional concepts or with the lack of knowledge about semantic restrictions. They are errors of inappropriateness, which are heavily restricted by the particular semantic contexts. Other error types, such as formal similarity and omission, appear to be attributable to confusion and poor knowledge factors. The students also have some problems, though relatively minimal, with forming set or matching phrases as seen in such types as word coinage, idiomatic collocation and Sino-Korean numeral collocations. Errors of redundancy and code-shifting were the types with the lowest frequency.

For the source of the errors, this study assigned the error types to four broad categories: 1. Interlingual (literal translation, code shifting), 2. Intralingual (semantic similarity, overgeneralization, formal similarity, idiomatic collocation and Sino-Korean numeral collocation), 3. Combination of both (omission / incompletion, word coinage, and redundancy) and 4. Unique (wrong word choice). The reason for the assignment of 'wrong choice' to the category of 'unique' errors is because it can hardly be said that those errors originated from either L1 or L2 itself in and of itself. They are more likely to be related to poor knowledge about the semantic concept of a word, and many of them seem to be attributable to the wrong selection of an appropriate meaning out of multiple meanings given in dictionary. If we assign the source of errors in this way, though arbitrary and tentative, the results are: 42% intralingual; 26% unique; 14% combination; and 12% interlingual. This result both supports and

contradicts previous studies, depending on the conditions set by each study (eg. learner profile, task and linguistic level). The high proportion of intralingual errors seems to reflect the intermediate-advanced level of learning experiences by the subjects in the study.

When we cross-examine the N and G data (see Table 2 below), we find that there is a general agreement in the most and least frequent error types, though there are a couple of interesting features to note. The single most frequent error type found in both data sources was wrong choice, which seemed to have been caused by a lack of knowledge about definitional concepts and restrictions, thereby resulting in the selection of an incorrect lexical item from dictionaries or memories. The most frequent errors found in N data were errors of wrong choice and semantic similarity, followed by literal translation and overgeneralization, while in G data, wrong choice, overgeneralization, semantic similarity and Sino-Korean collocations were most frequent. The least frequent errors produced in N data were Sino-Korean numeral collocations, redundancy, code-shift and word coinage, while code-shift, idiomatic collocation, redundancy and omission were the least frequent error types in G data. It is noted that the errors found in the N data are more concentrated in the first four error types and those in the G data are relatively sporadic. One explanation for this might be that at N, students were allowed to consult a dictionary during their examinations, whereas G students were not. This is interesting in that it gives a hypothetical idea that the use of a dictionary in a foreign language composition may lead to more frequent production of a particular error type or types. When students refer to a dictionary, which lists multiple definitions, there seems to be more of a risk to choose the wrong meaning. This is probably because a dictionary does not give much information on semantic restrictions and conceptual differences of multiple equivalents, and also because students have no or little previous knowledge about the lexical item, so they may randomly choose any given definition. The use of a dictionary may help in reducing certain types of error (probably, such types as code-shifting, formal similarity), but at the same time, it can create

more errors in other areas. Whether the students consulted their dictionaries or not, however, errors of wrong choice occurred with the highest percentage of frequency in both data, and this gives us an idea about the stage of students' lexical developments. Although there were similar error types with much higher concentration in the G data also (i.e. wrong word choice and overgeneralization), it seems that the heaviest concentration of certain types of errors (e.g. wrong word choice, literal translation and semantic similarity) was in the N results, most likely due to the dictionary factor. Also, it is interesting that overgeneralization was one of the most common error types. Students often tend to over-generalize what they learn, and the findings of this study prove this to be the case.

Table 2. Frequency of Error Types By Comparison (N/%)

Error Types	N	Rank	G	Rank
Wrong choice	58 (30)	1	25 (26)	1
Semantic similarity	36 (18)	2	9 (9)	3
Overgeneralisation	21 (11)	4	18 (19)	2
Literal translation	25 (13)	3	7 (7)	4
Formal similarity	9 (5)	6	9 (9)	3
Omission or incompletion	12 (6)	5	5 (5)	5
Word coinage	8 (4)	7	7 (7)	4
Idiomatic collocation	12 (6)	5	3 (3)	7
Sino-Korean numeral collocation	5 (3)	9	9 (9)	3
Redundancy	5 (3)	9	4 (4)	6
Code-shifting	6 (3)	8	1 (1)	8

When the errors were classified by word class (Table 3), the overwhelmingly dominant errors came from nouns and verbs. Nouns (53%) were the most common form of error, followed by verbs (33%). Errors from the other word classes were very minimal. This result supports previous studies such as Wang (1995) in the dominance and frequency of errors, though much more intensively concentrated in nouns and verbs in this study. There can be two or more possible

explanations for this result. One interpretation might be the general understanding about foreign language learning. Language learners tend to learn nouns and verbs before other classes of words and use them more frequently in the early stage of their learning. On the other hand, descriptive words such as adjectives and adverbs tend to be used less frequently or often avoided by the learners when they feel unsure about how to use these words. Obviously, familiarity and avoidance strategies seem to be strongly utilized in compositions by students, thereby creating an overwhelming proportion of errors in nouns and verbs. A second explanation might be in the students' lack of knowledge about the words they employ. As the learning process continues, students are expected to learn a number of ways to express abstract or conceptual ideas, which requires them to use an advanced form of nouns and verbs to describe objects and actions. Students then tend to be over ambitious about how to communicate in L2, and so they try to express themselves in the same way as they think and talk in their L1. Unfortunately, there is an undeniable gap between L1 familiarity and L2 knowledge. To fill this gap, he/she tends to rely on every possible source, from either their memory or dictionary, often without knowing the proper usage or the semantic restrictions of those words. Another possible explanation can be made in regards to translation, which was used as one of the principal written data types in this study. Translation tasks require additional skills and knowledge, and it often requires them to interpret a sentence carrying abstract ideas and concepts. This sort of arbitrary task as a data source may have assisted in the production of noun and verb errors. As a whole, one might say that the subjects in this study simply have not reached a level of competence in vocabulary to perform at an expected level, but this may be an oversimplification of the results, while ignoring the other part of the picture - what they have written correctly.

Table 3. Frequency of Errors by Word Class (N/%)

Word Class	N	G	M	Total

Nouns	110 (55)	47 (50)	5 (46)	162 (53)
Verbs	61 (31)	35 (37)	5 (46)	101 (33)
Adjectives	16 (8)	5 (5)	1 (9)	22 (7)
Adverbs	8 (4)	7 (7)		15 (5)
Idioms	4 (2)	1 (1)		5 (2)
Total				305 (100)

3.2. Examples of Errors and Explanations

This section presents some examples of lexical errors, along with attempts to explain the cause of such errors under each category. The following presents only the erroneous parts, which have been extracted from the full phrase or sentence containing the error. The examples under each category are listed in order of nouns or noun phrases, verbs, adjectives, adverbs and idiomatic expressions, where applicable. Corrections are in parentheses, while the intended meanings are in single quotation marks. Explanations are made in relation to some examples, which need more linguistic and pedagogical attention.

3.2.1. Wrong Word Choice

The following errors are all attributable to wrong word selection in essence. Some errors appear to be related to confusion factors, but most of them are simply due to wrong word choice. Many of them are irrevocably far from the context and non-sensical in the given context, as noticed in such examples as *gyo.je* for '(trading) company', *su.ryeong* for 'reception desk', *han.cheung gug.ga* for 'single-race country', *bang.beob* for 'manner', *yeon.seol* for '(mailing) address', *sag.je.ha.da* for 'to cancel', *gyeol.hab.ha.da* for 'to join (organization)', etc. in (1) and (2). In some cases, the wrong word choice may be due to failure to differentiate between subtle semantic differences, as observed in examples such as *sa.hoe.jeog.i.da* for *sa.gyo.jeog.i.da*, *teul.lin* for *da.reun* in (2) and (3). Also there are phrases consisting of more than one wrong choice of word as in *gong.gong su.song* for 'public transport' in (1). Many of these errors

are believed to be due to the wrong selection of a dictionary definition (as students in one data source were allowed to consult a dictionary. Most errors of this kind can be regarded as mistakes, rather than systematic errors, and for this reason, these "errors" have been assigned to 'Unique' category in the present study. If we look at the Table 2 in the preceding section, however, the wrong use of a dictionary definition is not the only cause for blame. Errors of word selection have taken the highest percentage of frequency in other data also as well, which was produced without dictionary consultation. Students often choose wrong nouns and verbs, and the consultation of a dictionary might have simply aided in producing more errors of this type.

(1)   gong.gong su.song (dae.jung gyo.tong) 'public transport'; gyo.je (hoe.sa) '(trading) company'; gyeol.hon hyeob.dong.ja (gyeol.hon sang.dae.ja) 'marriage partner'; se.gye hyu.il (se.gye.yeo.haeng) 'round-the-world trip'; su.ryeong (an.nae.chang.gu) 'reception desk'; gong.jung (gong.gong jang.so) 'in public'; han.cheung gug.ga (dan.il.min.jog.gug.ga) 'single-race country'; sog.dal.beo.su (go.sog.beo.seu) 'express bus'; bang.beob (mae.neo) ' manner'; pung.seub (mae.neo) 'manner'; yeon.seol (ju.so) 'address'; ji.jeong (yag.sog) 'appointment'; su.eob (gyo.sil) 'classroom'.

(2)   sag.je.ha.da (chwi.so.ha.da) 'to cancel'; kkeo.ji.da (myeol.jong.doe.da) 'to become extinct'; gyeol.hab.ha.da (cham.yeo.ha.da) 'to join (organization)'; sa.hoe.jeog.i.da (sa.gyo.jeog.i.da) ' to be social'; byeon.ha.da (ba.kku.da) 'to change (job)'; jeog.yong.ha.da (ji.won.ha.da) ' to apply ';

(3)   bu.jog.han (ga.nan.han) 'poor'; teul.lin (da.reun) 'different';

(4)   him.deul.ge (yeol.sim.hi) '(study) hard'.

3.2.2. Confusion by Semantic Similarity

While some errors are caused by selecting wrong words or forms, other errors can be caused by occasional or consistent confusion due to semantic similarity.

These errors seem to be related to a lack of knowledge about the conceptual differences between the competing words, rather than the result of random choice or the complete ignorance of the meanings. For example, *hyu.il, hyu.ga* and *gong.hyu.il* in (5) are all related to 'holiday', but there are definitional differences and they are discernible in actual use. The confusion between *si.gan* and *si* in (5), and between *joh.a.ha.da / joh.da* pairs in (6) are frequently seen in learners' compositions. It is probably due to the similarity in phonology or orthography, but it also may be due to the semantic association of 'time of period' with 'o'clock' and 'to be good' with 'to like' for each case, as each pair shares or is believed to share the same semantic origin. As a pedagogical suggestion, it may be better to treat and teach them as completely different semantic components. The source of confusion in other errors also appear to be the insufficient knowledge of subtle conceptual differences and usage of two similar words, as noticed in the 'pair' examples such as *jib / ga.jeong*, *ja.sig / a.i.deul*, *se.sang / se gye*, *i.min / i.ju* in (5), *geo.jeol.ha.da / geo.bu.ha.da, jag.da / jeog.da, kkag.da / be.da* in (6), *jeon.jin.jeog.in / jin.bo.jeog.in* in (7) and *dae.bu.bun / geo.ui* in (8). This confusion might also have been caused by induced factors, where students were not provided with clear-cut instructions, including the usage and practice.

(5) hyu.il (hyu.ga) 'recreational leave / holiday'; gong.hyu.il (bang.hag) 'school holidays'; han.gug jib. (han.gug.in ga.jeong) 'Korean home'; han.gug.sa.hoe (han.in.sa.hoe) 'Korean community'; ja.sig (a.i.deul) 'children'; se.sang yeo.haeng (se.gye.yeo.haeng) 'round-the-world-trip'; i.min (i.ju) 'migrant (animals)'; 3.si.gan (3.si); '3 o'clock'; wi.sa.ram (sa.jang / nop.eun sa.ram) 'boss'.

(6) joh.a.ha.da (joh.da) 'to be good'; joh.da. (joh.a.ha.da) 'to like'; chul.bal.ha.da (si.jag.ha.da). 'to start (a hobby)'; geo.jeol.ha.da (geo.bu.ha.da) 'to refuse'; jag.da (jeog.da) 'to be few / small in number'; keo.ji.da (jeung.ga.ha.da / neul.eo.na.da) 'to increase'; kkag.da (be.da / ja.reu.da) 'to cut (a tree)'; sam.ki.da (ma.si.da / meog.da) 'to drink'.

(7) jeon.jin.jeog.in (jin.bo.jeog.in) 'progressive'.

(8) dae.bu.bun (geo.ui) 'almost', cheos.beon.jjae (meon.jeo) 'first'.

3.2.3. Overgeneralization

A typical category of error is overgeneralization, where students overgeneralize their knowledge on the basis of their earlier learning. Students produce frequent errors in the construction of time phrases such as *da.eum.nyeon, i.nyeon and ji.nan.nyeon* in (9). Also, the counters or suffixes that are used with noun phrases are typically generalized as observed in the examples such as *bi.haeng.gi.se* ('airfare') and *jib.bi* ('rent'). Overgeneralization of such nouns as *il* (for *jig.jang* or *jig.eop*) and *si.nae* (for *do.si*) seems to be prompted by English expressions, and the incomplete nouns such as *myeong* and *geos* due to the lack of knowledge of grammatical restrictions. Some verbs appear to be overgeneralized due to both the interference from English and the intraligual influence. For example, *iss.go.sip.da* (for 'to wish to have') in (10) was produced by overgeneralizing the verb *iss.da* (existence and possession), without knowing that the item changes to *gaj.da* or *ga.ji.da* (possession) when a desire to possess something is expressed. Some other verb errors appear to have been made due to the insufficient knowledge about semantic restrictions. These verbs have two or more similar meanings that are largely integrated into one in English. *Neulg.da, nalg.da* and *o.rae.doe.da* are good examples for such case. Idiomatic or set expressions are also applied to alternative contexts, and it appears to be the result of a false hypothesis about the concepts or the overexpansion of the expression as part of communication strategies.

(9) da.eum nyeon (da.eum.hae / nae.nyeon) 'next year'; i.nyeon / i.beon.nyeon (ol.hae / keum.nyeon) 'this year'; ji.nan.nyeon (ji.nan.hae / jag.nyeon) 'last year'; i.hag.nyeon (i.beon hag.nyeon) 'this school year'; mo.deun ju (mae.ju) 'every week'; ho.ju.bun (ho.hu.sa.ram.deul) 'Australians (general)'; bi.haeng.gi.se (bi.haeng.gi.yo.geum) 'airfare'; bi.haeng.bi

(bi.haeng.gi yo.geum) 'airfare'; jib.bi (jib.se) 'rent'; il (jig.jang /jig.eop) 'workplace / job'; si.nae (do.si) 'city'; myeong (sa.ram) 'person (independent noun); geos (mul.geon / hyu.dae.pum) 'objects (personal belongings)'.

(10)   iss.go.sip.da (gaj.go.sip.da) 'to wish to have'; neulg.da (nalg.da) 'to be old (objects); neulg.da (o.rae.doe.da /) 'to have been long (time) / used for long (objects)'; manh.da (gil.da) 'to be long (length).

(11) go.jang.i.na.da (da.ddeol.eo.ji.da) 'to run out of (objects)'; bu.tag.ha.da (si.kyeo meog.da) 'to order (something to eat)'.

3.2.4. Literal Translation

Another common lexical errors occur by literally converting L1 items, and such errors are due to a strong interference from English.   As seen in (12), (13), (14) and (15), the errors have been made by sticking to the literal meaning of the English version or to the way the words or phrases are expressed in English. Some examples are typical, and others are new.   For example, errors such as *gyeol.gwa* and *hwal.dong* in (12), *nol.da* and *ga.da* in (13) and *man.deul.go.sip.da* in (15) are frequently found in students' compositions at a similar level.   Other errors such as *sseu.da*, *gat.i.sseu.da*, *keu.da* in (13), *ssan* in (14) and *seong.jil.i.jjalb.da* in (15) are also interesting to note.   Some errors are caused by the literal construction of terms and phrases such as *in.gan.ja.won* and *neulg.eun.se.dae* in (12). Items in this category may be relate to other sources (eg. selection or semantic similarity), but the nature of the errors are closely related to literal translations.

(12) ab.ryeog (seu.teu.re.seu) 'pressure'; sang.eob.hoe.sa (sang.sa); 'trading company'; gyeol.gwa (seong.jeog) 'results /record'; oe.gug.eon.eo (oe.gug.eo) 'foreign language'; hwal.dong (haeng.sa) 'activity (event)'; gug.je ju.sik.hoe.sa (gug.je.jeog.in hoe.sa) 'international company'; in.gan ja.won (in.ryeog gwan.ri) 'human resources (organization)'; yeo.geub.sa

(we.i.teu.re.seu) 'waitress'; neulg.eun se.dae (no.in.se.dae/ no.in.deul) 'old generation / people'; o.rae sa.ra.deul (no.in.deul) 'old people'.

(13) nol.da (ha.da) 'to play (sports)'; so.ri.reul man.deul.da (tteo.deul.da) 'to make noise'; ga.da (da.ni.da) 'to attend (regular visit)'; sseu.da (ta.da / il.yong.ha.da) 'to use (public transport)'; gat.i sseu.da (gat.i.ta.da / hab.seung ha.da) 'to share (taxi)'; bal.gyeon.ha.da (al.a.bo.da) 'to find out (facts)'; sig.jag.ha.da / chul.bal.ha.da (deul.eo.ga.da) 'to enter / start (school)'; keu.da (manh.da) 'to be big / many (family)'.

(14) ssan (don.i an deu.neun) 'cheap (hobby)'.

(15) man.deul.go.sip.da (sa.gwi.go.sip.da) 'to make (friends)'; keun pa.ti.neun iss.da (pa.ti.rul keu.ge yeol.da) 'there will be a big party'; yuk.che.jeog.eu.ro bo.yeo.ju.da (geot.eu.ro pyo.hyeon.ha.da) 'to physically show'; seong.jil.i jjalb.da. (seong.jil.i.jo.geub.ha.da) 'to have a short temper'.

3.2.5. Confusion by Formal Similarity

As seen in the following examples, some errors are attributable to confusion by formal similarities. These errors are beyond orthographic or spelling errors in that they normally maintain correct forms and meanings by themselves. Whether the errors are simple mistakes or systematic errors, they are all related to occasional or habitual confusion due to the similarities in the formal and phonological aspects. In particular, words ending with nasal sounds seem to carry phonological confusion factors as observed in *seon.saeng.oe(hoe)*, *gwang.go.hag*. and *gong.jeon* in (16). Other errors seem to be caused by a combination of formal / phonological and semantic factors. Errors such as *ij.eo.meog.d*a (or *ij.eo.beo.ri.da*) in (17), which are observed even in native speakers' speech, might be more consistent as it shares the semantic association with *ilh.eo.beo.ri.da* in that both indicate 'something gone'. A good example of this combination factor is found in the pair, *noh.da* and *neoh.da*, where both

carry the meaning of 'placing something'.   The similarity in form and sound is the primary cause of confusion in this case.

(16) seon.saeng.oe   (saeng.seon.hoe)   'raw   fish';   gwang.go.hag (gwan.gwang.hag) 'tourism study'; gong.jeon   / jeong.gong (jeon.gong) 'major study'; da.yeo.seos (dae.yeo.seos) 'five or six'.

(17) ij.eo.meog.da (ilh.eo.beo.ri.da) 'to lose'; noh.da (neoh.da) 'to put something in'

3.2.6. Omission or Incompletion

Students produce many sentences where appropriate nouns and forms are missing or simplified. Frequently, such missing words or forms are attributable to interlingual transfer factors. Observe the examples in (18) where parts of some phrases can be or are often omitted in casual English-speaking contexts. For example, *taeg.si, chi.mi.* and *pil.su.gwa.mog* in (18) can be enough in the given contexts, but they are incomplete and inappropriate in the parallel Korean context. The superlative *je.il* in (19) needs adjective content words to support it and cannot occur by itself in the noun phrases. In this case, the omission may be due to either poor knowledge about its usage or interference from English, which has special uses of *je.il* or 'the best'. Other errors in this category include missing syllables within a word as in (20), but it is impossible to interpret in such case, whether it is the result of a false hypothesis or a simple mistake.

(18) taeg.si (taeg.si.yo.keum) 'taxi fare'; chi.mi. (chi.mi saeng.hwal) 'hobby (activity)'; pil.su gwa.mog (pil.su gwa.mog jeom.su); jeon.hwa (jeon.hwa.beon.ho)   'telephone   number';   ja.dong   (ja.dong.cha) 'automobile'.

(19) je.il (je.il joh.a.ha.neun) 'favourite'; je.il (je.il keun) 'the biggest'.

(20) kkae.kkeus (kkae.kkeus.han) 'clean (air)'.

3.2.7. Word Coinage

(21) and (22) illustrate errors of word coinage, which requires concise matching of expressions. The first or second part of the pair expressions is a mismatched word, and in some cases, the paraphrased part is an inappropriate one for the pair. These errors are largely due to an insufficient knowledge about pair expressions, but they seem to be caused by other influences as well, such as literal translation and overgeneralization. Still in other cases, communication strategies such as circumlocution or paraphrasing can be the cause of the ill-formed part of set expressions, as noticed in *da.reun sa.ram.gwa taeg.si.reul ta.gi* and *taeg.si na.nwo.seo ta.neun.geo* in (22). From the examples, it is fair to say that both interlingual and intralingual factors are associated with the cause of the errors in this category.

(21) nong.sa ma.eul (nong.chon) 'farming country'; jib saeng.hwal (ga.jeong.saeng.hwal) 'family life'; gong.hwa na.ra (gong.hwa.gug) 'republic'.

(22) il.ha.neun saeng.hwal.(jig.jang saeng.hwal) 'working life'; da.reun sa.ram.gwa taeg.si.reul ta.gi (taeg.si hab.seung) 'sharing taxi'; taeg.si na.nwo.seo ta.neun.geo (tae.si.hab.seung) 'sharing taxi'.

3.2.8. Idiomatic Collocation

In idiomatic or term-like expressions, no part can be simply replaced with some other items if the intended idiomatic implications are to be maintained. Students, however, tend to produce non-idiomatic expressions without knowing the idiomaticity of the items. As observed in (23) and (24), these errors are caused by various reasons such as incorrect paraphrasing, wrong word selection, semantic confusion or circumlocution. Typical examples include *sig.sa.reul yo.ri.ha.da* in (24). The second part, *yo.ri.ha.da*, cannot occur with general *sig.sa* 'meal', and it normally comes with a specific dish or type of cuisine, eg. *bul.go.gi*, Chinese, Italian, etc. Also, the noun *yeong.hyang* 'influence' forms

an idiomatic phrase in accordance with *mi.chi.da, kki.chi.da*. or *ju.da*. Some causes of these errors seem to be related to interlingual interference, but primarily, they tend to be more related to the complexity of the target language (Korean) itself and to the instructional content (eg. induced).

(23) neo.mu iss.neun don (nam.eun.don / jan.don) 'left-over money / change'.

(24) sig.sa.reul yo.ri.ha.da (sig.sa.reul jun.bi.ha.da) 'to prepare a meal'; chi.u.da (seol.geoj.i ha.da) 'to wash dishes'; in.sang.eul saeng.gag.ha.da (in.sang.eul bad.da) 'to get impression'; yeong.hyang.i.da (yeong.hyang.eul mi.chi.da) 'to have an effect /influence on'; gal.su eob.da (go.jang na.da) 'to be broken down (vehicle)'.

3.2.9. Sino-Korean Numeral Collocation

As Korean uses two numeral systems (i.e. Pure-Korean and Sino-Korean), the way the numerals are counted is decided by the counters, as observed in (25). Also, the counters are decided by the numeral system that is being used or the intended utterance, as in (26). For example, *il.dal* could be interpreted as *il.gae.wol* 'one month' and *6.wol* as 'June' but the counters of each item should change because of the intended meanings, e.i. 'January' and 'six months' It is typical to see students frequently confused with this kind of numeral-counter collocation and making consistent errors, which are clearly intralingual errors. This is one of the troublesome areas that need more effective pedagogical strategies to tackle this problem.

(25) il.gob.nyeon (7/chil.nyeon) '7 years'; sib.i sal (yeol.du.sal) '12 years old'; du.nyeon ban (2/i.nyeon.ban) '2 and a half years'; 6.dal (yeo.seos dal / 6/yug gae.wol) '6 months'.

(26) il.dal (1.wol) '1st month /Janunary'; 6.wol (6 gae.wol / yeo.seos.dal) '6 months'; 10.nal (10.il) '10 days'; du.gae (du.myeong) 'two persons'.

3.2.10. Redundancy

Occasionally, students add unnecessary words or lexical forms mainly because they are unaware of the repetition or redundancy. For example, in (27), *ho.ju.na.ra* and *sang.sa hoe.sa* do not need the second part of each phrase as it was already understood or implied in the preceding lexical item. Also, in (28), *don* 'money' was already contained in *byeong.won.bi* 'hospital charge', and the mood of *heung.mi.rob.da* 'appealing' largely in the preceding expression *jae.mi.iss.da* 'interesting', though there is a subtle difference. The redundant *keun* 'big' in (29) was added probably due to the false association about something big or excessive, and the insertion of the incomplete noun, *geos,* 'a thing' is due to immature syntactic knowledge. Whatever the pattern is, these errors all sound repetitive and redundant, and they can be caused by either interference from English or intralingual developmental influences.

(27)   ho.ju na.ra (ho.ju) 'Australia'; sang.sa.hoe.sa (sang.sa) '(international) trading company';

(28)   byeong.won.bi don.eul (byeong.won.bi.reul) nae.da 'to pay the hospital charge'; jae.mi.iss.go heung.mi.han.da (jae.mi.iss.da) 'to be interesting';

(29)   ga.jang   keun   bi.ssan   (ga.jang   bi.ssan)   'the   most   expensive'; joh.a.ha.neun.geos.eun chwi.mi (joh.a.ha.neun.chwi.mi) 'one's favourite hobby'.

3.2.11. Code-Shifting

In a second language speaking situation, it is not difficult to observe code-switching or code-shifting in the speakers. In L2 compositions, students, though infrequent, borrow L1 codes to replace L2 lexical components, and they are transcribed in L2 letters as noticed in (30) or written in L1 spelling as in (31). Frequently, these errors are produced when students do not know the L2 word, and this produces negative transfer errors.

(30) peu.ri.jeu (naeng.jang.go) 'fridge / refrigerator'; aeg (ig).sa.i.ting.ha.da (a.ju jae.mi.iss.da) 'to be exciting'; sa.keo (chug.gu) 'soccer'; e.ti.kes (ye.jeol) 'manners'.

(31) panic   ha.da   (dang.hwang.ha.da)   'to   be   panic';   sportscenter (seu.po.cheu.sen.ta) 'sportscenter'.

Thus far, we have examined the 11 categories of lexical errors in composition. There are still some errors that are ambiguous, and are difficult to classify or explain the possible causes. We can assign those errors to another category and identify the cause from another source, as they may share some common features of two or more categories. In fact, we have observed that there are quite a good number of errors whose possible cause can be multiple, and thus can be explained in different ways. We can also select a particular category and explain the nature of those errors in more detail, but that is not the intention of this study.

4. CONCLUSION

There are always risks of certain inaccuracy in identifying, analyzing and classifying errors. This study is not an exception. The limitation of the present study includes the lack of systematic elicitation of data, the ambiguity of some classification criteria, and the incomprehensive nature of explanations. This study is not intended to make any vigorous claims as the classification of errors is largely a subjective matter and the categorization of error source is still ambiguous as pointed out by Ellis (1994: 62-63). Nevertheless, the present analysis of the textual data from the written tests does provide some insights into the areas of weakness in students' lexical knowledge, the patterns of lexical errors in composition, and possible methodological guidelines for future studies in this field.

The findings in the present analysis can be summarised in the following points. First, among the 11 error categories identified in this study, wrong word choice caused an overwhelmingly high percentage of errors, and this was the

case regardless of whether or not a dictionary was consulted. Other error categories with high frequency were: confusion by semantic similarity, overgeneralization and literal translation. Second, more than half of the total errors involved nouns, followed by verbs and adjectives. It is possible to interpret these findings in two ways: that students' lexical developments are still in the early stages, where nouns and verbs are more frequently used to form a basic sentence, or that students attempted to or were asked to attempt to make sentences requiring abstract nouns for concepts that are beyond their L2 proficiency. Third, a large number of the errors produced are intralingual. Interlingual errors are relatively small among the four error sources assigned to the 11 error categories.   The large proportion of intralingual errors present an encouraging sign in students' lexical developments, in that as learning proceeds, intralingual errors are generally more produced than interlingual transfer.

The result of this study provides us with some important implications. Theoretically, EA continues to attract SLA researchers who want a methodology to deal with learner data and who want to establish new hypotheses on a theory of acquisition and other linguistic aspects. This is particularly the case for lexical EA as lexis is now playing an important role in language study and language learning (James, 1998). In this regard, the present error data and the results of this analysis would be useful, though limited, to those who are considering exploring research in the acquisition of Korean as a foreign language. Pedagogically, we have been able to identify, tentatively, the patterned lexical weaknesses of Australian students and the possible causes of these problems. The findings would be useful for the design of remedial programs and the development of teaching materials including a learner dictionary. For instructors, there is a need to devise pedagogically effective learning and teaching strategies that prevent fossilization of certain errors in students' lexical developments. For students, it is desirable to understand and be able to use vocabulary with multiple meanings, for example, by reading L2 language materials as much as they can, rather than relying solely on a dictionary. In terms of research, this study recognizes the importance of a highly

systematic data collection method and a comprehensive error categorization system that other investigators could readily adapt to their purposes and conditions. Such an elicitation method and a categorization system should be able to provide grounds to better account for students' production of lexical errors. Also, there is a need for a parallel longitudinal study that deals with the same lexical error patterns and compares error types and frequency of different learner groups, so we can control a number of variables in individual learners and thus gain a better understanding of lexical developments. Finally, questions raised from the present study must await answers from further research, which will provide significant grounds to make any generalization on students' lexical development.

References

강화진. 2000. KFL 학습자들의 작문 오류분석을 통한 효율적인 작문지도. 고려대학교 석사학위논문.

김미옥. 1994. 한국어 학습에 나타난 오류 분석. 한국말 교육 5. 국제한국어교육학회.

김미옥. 2001. 맞춤법 오류에 관한 연구. 외국어로서의 한국어 교육 Vol 25-26. 연세대학교 언어교육연구원.

김영아. 1990. 재중 교포들의 한국어 학습시 나타나는 오류 분석. 이중언어학회지 7. 이중언어학회.

김유미. 2000. 학습자 말뭉치를 이용한 한국어 학습자 오류 분석 연구. 연세대 교육대학원 석사학위논문.

김정숙. 1988. 일본인의 한국어 학습시 나타나는 음운론적 오류분석. 한국어학 신연구. 한신문화사.

왕혜숙. 1995. 영어화자의 한국어 작문에 나타난 어휘상 오류 분석. 이중언어학회지 12. 이중언어학회.

이수경. 1996. 일본어를 모어로 하는 한국어 학습자의 오류에 경향에 관하여. 한국말 교육 7. 국제한국어교육학회.

이정희. 2001. 한국어 학습자의 시제 오류 연구. 이중언어학 18. 이중언어학회.

전은주. 1994. 한국어 학습시 나타나는 학습자의 모국어 영향. 한국말 교육 5. 국제한국어교육학회.

최우영. 1997. 외국어로서의 한국어 학습자의 오류에 대한 연구. 이화여대 석사학위논문.

Corder, S.P. 1967. The significance of learner's errors. *International Reviews of Applied Linguistics* 5, and also in Corder (1981).

Corder, S.P. 1975. Error analysis, interlanguage and second language acquisition. *Language Teaching and Linguistics Abstracts* Vol 8.

Corder, S.P. 1981. *Error Analysis and Interlanguage*, Oxford: Oxford University Press.

Dulay, H. and M. Burt. 1974b. Errors and strategies in child second language acquisition. *TESOL Quartterly* 8: 129-36.

Ellis, R. 1994. *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

James, C. (1998). *Errors in Language Learning and Use*. London: Longman.

Geroge, H. 1972. *Common Errors in Language Learning: Insights from English.* Rowley, Mass.: Newbury House.

Larsen-Freeman, D. and M. Long. 1991. *An Introduction to Second Language Acquisition Research*. London: Longman.

Norris, J. 1983. *Language Learners and their Errors*. London: Macmillan Publishers.

Richards, J. C. 1971b. A non-contrastive approach to error analysis, in Richards R. C.(ed.). (1974). *Error Analysis: Perspectives on Second Language Acquisition*. London: Longman.

Shin, S-C. 2001. An error analysis: orthographic errors produced by Australian KFL learners. *Teaching Korean as a foreign language* Vol. 25-26. Yonsei University Institute of Language Research and Education.

Sohn, H.M. 1986. Patterns of American students' composition errors in Korean, *Linguistics Expeditions*. Seoul: Hanshin Publishing Co.

Taylor, B. 1975. The use of overgeneralization and transfer learning strategies by elementary and intermediate students of ESL. *Language Learning* 25: 73-107.