Identification of Lexical Competence Problems in L2 Writing Process of Korean EFL University students

Yuah Vicky Chon
(University of Essex)


While the nature of productive lexical knowledge may be distinguished as competence-based or performance-based, classification of lexical problems in the translating phase is lacking in the writing strategies literature. With separate interest in competence-based word knowledge, this study investigates the lexical aspect of the writing process by exploring the types and tokens of lexical problems that L2 writers identify in the drafting (translating) stage of writing. To obtain data, 10 Korean university students were asked to think-aloud while writing twice in English on word processors with access to Internet dictionaries. A repeated measures design was used where students wrote for two audiences and two writing topics. Three main types of lexical competence problems were found where use of the Internet dictionary had an effect on the identification of problems. Signs of vocabulary problems indicated that writers had unknown word, partially known word, and more than one word problems where eight subcategories of partially known word problems were found. The findings may provide pedagogical implications for helping L2 writers solve different types of lexical problems such as in contexts where word processors and Internet dictionaries are used as the medium.

I. INTRODUCTION

When it was realized in the context of Korean university EFL writing classes that students are often left to complete their writing assignments outside the classroom (Yuah Chon & Hae-Dong Kim, 2005), it became apparent that the cognitive aspect of the L2 writing process, especially concerning how conceived ideas are transferred to written text (i.e., drafting/translating), does not attract the attention it deserves. In the L2 translating stage, lexical problems are bound to occur since students are asked to write in a language
with a smaller mental lexicon (ML)\(^1\) where an L2 equivalent known in L1 may not exist or where knowledge of an L2 word is only partially known. To solve these lexical problems in writing, students may use a range of communication strategies (Dörnyei & Scott, 1997; Færch & Kasper, 1983), facilitative strategies that allow learners to make use of existing linguistic resources to convey the intended meaning as accurately as possible (Chimpaganda, 2000). However, classification of the types of lexical problems that writers identify is a prerequisite for investigating writing communication strategies since different strategies are used for different problems. In consideration of the translating phase and the vocabulary problems that are identified by L2 writers, this study aims i) to devise a coding scheme for classifying lexical problems that occur due to lack of lexical knowledge (i.e., lexical competence problems), ii) to classify the types of lexical competence problems that are conscious to L2 writers, and iii) to investigate the frequency of these vocabulary problems.

II. LITERATURE REVIEW

1. Definition of a Word

Since this study is on writing vocabulary problems, there is need to review the definition of a word and what knowing a word for production entails. It has been pointed out that trying to define a ‘word’ is not simple due to several problematic aspects. Schmitt (2000) explains that it is too general a term to encapsulate the various forms vocabulary takes which may include single words (e.g., *die, expire*), phrasal verbs\(^2\) (e.g., *pass away*), and idioms\(^3\) (e.g., *bite the dust, kick the bucket, give up the ghost*). As such, Schmitt points out that defining the term ‘word’ is difficult due to the possible lack of correspondence between individual words and individual meanings. Words may not have a one-to-one correspondence between a meaning and a single word so that meanings in English are often represented by multiple words. To handle these multiword units, the interchangeable terms *lexeme, lexical unit, or lexical items* are used which are defined as “an item that functions as a single meaning, regardless of the number of words it contains” (p. 2). Thus all words ranging from single units to idioms are *lexemes*.

Defining the term ‘word’ also causes difficulties due to the various grammatical and

---

\(^1\) The ‘mental lexicon’ is a person’s mental store of words, their meanings and associations

\(^2\) A ‘phrasal verb’ is a verb plus one or more other prepositions or adverbial particles, which has an idiosyncratic meaning compared to the component words.

\(^3\) An ‘idiom’ is a string of words which taken together has a different meaning than the individual component words.
morphological variations of vocabulary (p. 2). For instance, it is not clear whether walk, walked, walking, and walks should be counted as a single word or four. In these examples, there is a base, root, or stem word that is the simplest form of that word. When inflections are added to the stems for grammatical purposes, the resulting word is inflected (e.g., walked, walking, and walks are inflections of the root word walk). However, if the affixes change the word class of a stem (where exceptions are that some derivations may result to be the same word class; e.g., happy, unhappy), the result is a derivative. Thus stimulative (adjective) and stimulation (noun) are derivatives of stimulate (verb). Sets of such words that are related in meaning with slight difference in orthographic form are word families that include the base word, all of its inflections, and its common derivatives. In comparison, the term lemma is more restricted and includes only the base word and its inflections. Ultimately in this study, the term ‘word’ refers to various variants of lexeme, and to lexical items such as multi-word verbs, phrasal verbs and idioms.

2. Knowing a Word for Production

Nation (1990, 2001) describes what ‘knowing a word’ requires through the receptive (listening and reading) and productive (speaking and writing) distinction of word knowledge. Productive knowledge of a word includes receptive knowledge and extends it. It involves knowing how to pronounce the word, how to write and spell it, how to use it in correct grammatical patterns along with the words it usually collocates with. Productive knowledge also involves not using the word too often if it is typically a low-frequency word, and using it in suitable situations. It involves using the word to stand for the meaning it represents and being able to think of suitable substitutes for the word if there are any. For a more systematic classification of the features entailed in knowing a word, Nation (2001) divides ‘what is involved in knowing a word’ into aspects of ‘form,’ ‘meaning’ and ‘use.’ With our specific interest in production, the aspects of productive knowledge and use are presented with examples in Table 1.

Another way of classifying word knowledge is the dichotomy of competence and performance which has been discussed by language scholars to deal with the nature of language ever since Chomsky (1965) first made the distinction concerned with the theory of syntax. He made the fundamental distinction between competence as “the speaker-hearer’s knowledge of the language” and performance as “the actual use of language in concrete situations” (p. 4).

Also within the context of the study, problems and errors need to be distinguished for discussing vocabulary problems. Errors are competence problems which are “wrong forms that the pupil could not correct even if their wrongness were to be pointed out” (Edge, 1989, p. 10) so that when errors are committed writers have no awareness as to their lack
of knowledge. While errors are left unattended by the learners since they are not identified as being problematic in the course of L2 communication, problems can be noticed or identified by L2 learners in the process of communication (e.g., I don’t know how to say this word in English). Problems are self-initiated by the learner particularly in writing (cf. they may be other-initiated by an interlocutor in speaking). In the case of a problem, the L2 learner will try to bridge the gap between the problem and intended goal by trying to reach a solution (e.g., a strategy). With errors there is no realization of this kind and they can be corrected only by another such as a teacher/native speaker. In this study, we are only interested in the more or less conscious problems, not errors, since only those problems that are identified by the writer can motivate writers to find a solution (via strategies).

<table>
<thead>
<tr>
<th>Form</th>
<th>Meaning</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>spoken</td>
<td>form &amp; meaning</td>
<td>grammatical functions</td>
</tr>
<tr>
<td>written</td>
<td>concept &amp; referents</td>
<td>collocations</td>
</tr>
<tr>
<td>word parts</td>
<td>associations</td>
<td>constraints on use (register, frequency....)</td>
</tr>
</tbody>
</table>

TABLE 1
Knowing the Productive Aspect of the Word (adapted from Nation, 2001)

<table>
<thead>
<tr>
<th>Productive aspect of the word</th>
<th>Form</th>
<th>spoken</th>
<th>How is the word pronounced?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(e.g., being able to say underdeveloped with correct pronunciation including stress)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>written</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>How is the word written and spelled? (e.g., being able to write it with correct spelling)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>word parts</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>What word parts are needed to express the meaning? (e.g., being able to construct it using the right word parts in their appropriate forms)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Meaning</td>
</tr>
<tr>
<td></td>
<td></td>
<td>form &amp; meaning</td>
<td>What word form can be used to express this meaning? (e.g., being able to produce the word to express the meaning underdeveloped)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>concept &amp; referents</td>
<td>What items can the concept refer to? (e.g., being able to produce the word in different contexts to express the range of meanings of underdeveloped)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>associations</td>
<td>What other words could we use instead of this one? (e.g., being able to produce synonyms and opposites for underdeveloped)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Use</td>
<td>In what patterns must we use this word? (e.g., being able to use the word correctly in an original sentence)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>collocations</td>
<td>What words or types of words must we use with this word? (e.g., being able to produce the words that commonly occur with it)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>constraints on use (register, frequency....)</td>
<td>Where, when, and how often can we use this word? (e.g., being able to decide to use the word to suit the degree of formality of the situation, such that developing is more acceptable than underdeveloped which carries a slightly negative meaning)</td>
</tr>
</tbody>
</table>

3. Writing Strategies and Vocabulary Problems

Foremost for the present study, we are interested in the cognitive aspects of the L2
writing process when writers may be under lexical constraints to construct text for their intended message. As such, reviewing L2 writing process literature that gives reference to any lexical aspects of the writing process with identification of the types of problems is of interest.

There have been some references to vocabulary problems in the L2 writing process literature, but they have been treated marginally and there has not been any systematic classification of the types of writing vocabulary problems (e.g., Cumming 1989, 1990). For instance, models of how language is processed in written production have been popularly represented by Flower and Hayes’ model of writing (1981) which characterized much of the work of the 1980’s within the cognitive process approach to writing. However, the model, in spite of its claimed contribution for explaining the cognitive process of L2 writing, fails to give any account of how lexical production is achieved with regard to vocabulary problems which is most likely to occur in the ‘translating’ phase of the writing process (unless it has been realized in the planning stage). Others have also noticed the flaw of the translating (formulating) phase positing that it is in fact the non-optional stage of writing (McCutchen, Covil, Hoyne, & Mildes, 1994; Roca de Larios, Manchon, & Murphy, 2006; Scardamalia & Bereiter, 1986; Witte & Cherry, 1986) where vocabulary issues are bound to arise.

There have been references to writing vocabulary problems in cognitively-oriented studies which examine L2 learners’ ongoing thinking episodes or decision-making processes where writers search for appropriate words or phrases (Cumming, 1989, 1990; Whalen & Ménard, 1995; Roca de Larios, Murphy, & Manchon, 1999; Roca de Larios, Marin, & Murphy, 2001; Roca de Larios et al., 2006; Yeon Hee Choi & Jieun Lee, 2006). Yeon Hee Choi and Jieun Lee (2006) conducted a study on Korean university students to investigate the impact of L2 proficiency and writing task difficulty on their L1 use. Although writing vocabulary problems were not of interest, think-aloud protocols obtained from the students indicated that lower proficiency writers depended more on the use of L1 for translating L2 into L1 in order to generate ideas, where vocabulary problems are likely to occur, and their reliance on L1 increased for lexical searching and language use in the argumentative task.

Cumming (1990) conducted a study with the aim for describing episodes of concurrent metalinguistic and ideational thinking of the 23 adult Francophone ESL learners composing on two tasks, writing an informal letter and an expository argument. The findings of the study indicated that the ESL students devoted higher percentage of their decisions to thinking concurrently about gist and language use (where vocabulary is one of the components) when working on the more cognitively demanding complex argument task. In fact, in these writing processes two goals appeared to have guided the vast majority of these behaviors: (a) searching for and assessing appropriate words and phrases and (b)
comparing cross-linguistic equivalents which all imply existence of vocabulary problems, however, without any classification of them. The most conspicuous cognitive activity in these thinking episodes involved searching for the ‘right’ word to express an idea where a lot of mental effort was put into “seeking out a word or phrase, generating and assessing possible alternatives, and then reaching a choice that satisfied the sense appropriate to their discourse” (p. 491). Learners appeared “compelled to generate mentally a small thesaurus of alternatives, calling up numerous categorically related items from memory” (p. 500) which are expected to cause vocabulary problems.

Whalen and Ménard (1995) in their study on L2 writing strategies attempted to see how L2 writers employ their strategic knowledge while writing and to characterize the interactions between strategic knowledge and linguistic knowledge during writing. They examined three factors of ‘writing strategies’ (i.e., planning, evaluation and revision) at different ‘discourse processing levels’ (pragmatic, textual and linguistic) when 12 Anglophone second-year French undergraduate students were asked to write argumentative texts in both English (L1) and French (L2) while thinking-aloud. The results of their study showed that limited knowledge of L2, by which vocabulary problems are expected to occur, constrains the L2 writers’ planning, evaluation and revision strategies from being employed at more global levels of discourse processing. The researchers report that there was use of circumlocution (p. 384) and the dictionary (p. 409) which indicates that there was identification of lexical gap problems by the writers. Also in L2 writing, 77% of all processing which occurred at the linguistic level seemed more problematic when writers had to concentrate on ‘searching, accessing, and retrieving linguistic forms’ (p. 406). One of the explanations for this is that the writers had to constantly search for appropriate lexical items in dictionaries, and verify morphological rules at which point writers are likely to meet vocabulary problems.

Second language writing strategy studies of Spanish EFL writers which pay separate attention to the ‘formulation’ (i.e., translating) stage (Roca de Larios et al., 1999; Roca de Larios et al., 2001; Roca de Larios et al., 2006) make reference to writing vocabulary problems. The studies generally show that writers engage in use of strategies to ‘upgrade expression of meaning’ to find a better match between intention and expression (i.e., fluent formulation expectedly due to performance problems) and to ‘compensate for their lack of linguistic resources’ (i.e., problem-solving formulation due to competence problems). In connection with our interest in lexical competence problems, they adopt Newell’s (1980) notion of problem space as the basic frame of reference to demonstrate that ‘problem-solving formulation processes’ is conceived as ‘the sequence of operations that proceed from the initial state (the way the writer represents the problem to himself) to a goal state (the solution to the problem)’ (Roca de Larios et al., 2001, p. 514). As examples of problem-solving behaviors, the researchers saw that writers retrieved initial syntactic
plans plus a sequence of operations (rereading, pausing, repairs, pre-text, evaluations, etc.) leading to the production of one or more alternative plans. These EFL writers showed explicit signs of their problems by a variety of means: candidate formulations (‘what is the right word, help or helping?’), explicit problem indications (‘I don’t know how to write this’), or repairs that prompt a search episode. These are all signs of vocabulary problems that the researcher in the present study also expected to observe in the think-aloud protocols of Korean university students.

The two types of formulation behavior mentioned previously are also related to how word problems are classified by Hemmati (2001). Using Chomsky’s (1965) dichotomy of competence and performance that has been used to deal with the nature of language, she distinguishes word knowledge as ‘competence-based problems’—writer’s stored knowledge of the language, and ‘performance-based problems’—use of that vocabulary knowledge. It is these notions of competence and performance that is adopted for the coding scheme of vocabulary problems later in the study. Hemmati’s study on vocabulary problems and taxonomy of lexical strategies was conducted on EFL Iranian university students where one of the aims was to explore the types of conscious vocabulary problems that Iranian EFL learners meet in the writing process. As a result of analyzing the think-aloud protocols, she found that the writers met a range of different types of vocabulary problems. With our separate interest in competence lexical problems, the types of word problems that were found in the study were: NO WORD/’unknown word’ problems (i.e., the target word cannot be retrieved at all since the word has either never been learned or once learned, but has been forgotten and is not available anymore), ONE WORD/’partially known one word’ problems (i.e., a writer succeeds in retrieving one lexical item but then realizes that some aspect of it is either unknown or uncertain), and MORE THAN ONE WORD/’word confusion’ problems (i.e., writers are confused between more than one retrieved words and the dilemma is to choose one word).

As a whole, evidence found in previous studies supports our expectation that L2 students will consciously recognize a linguistic problem as a result of producing, or trying to produce the target language (e.g., Cumming, 1989, 1990; Hemmati, 2001; Roca de Larios et al., 1999, 2001). However, vocabulary aspects in these studies are dealt with only in fragments or hinted without any systematic classification of the problems. Hemmati’s classification is also problematic since she lacks consistency in the categorization of ONE WORD problems where she fails to show that the problems can occur as parallel subcategories between competence and performance problems. Having recognized that there is lack of research on any systematic classification of the types of vocabulary

---

4 Report on the writers’ identification of performance-based lexical problems is beyond the context of the present study.
problems that writers identify during the process of L2 writing, the present study attempts to answer the following research questions.

1) What are the types of vocabulary problems that writers consciously identify in the translating (drafting) stage due to lack of vocabulary knowledge (i.e., lexical competence problems)?
2) What is the frequency of these lexical competence problems identified by writers in the translating stage?

III. METHOD

1. Participants

There were 10 students who took part in the study, and they were from an English elective course English through the Internet (ETI) at a university in Seoul, Korea. The course provided an ideal context for data collection where students would be writing in English with access to the Internet. The researcher participated in the study and in the instruction of the course as a researcher and teacher (i.e., a participant observer) to familiarize the students with the writing tasks of the course and to conduct a study on the writing process at the end of the semester. The dual role as a researcher and teacher in the course facilitated the researcher to take control of any intervening variables (e.g., amount, type of instruction and tasks taught) that might have influenced the context of the study tasks. Regarding background profile, the students had learnt English for a mean of ten years in an English as a Foreign Language (EFL) context. The students consisted of majors in English (n=5, 50%), Business (n= 4, 40%) and French culture (n=1, 10%). There were 8 male and 2 female students. The mean age of the subjects was about 24 and they ranged from sophomores to juniors. When the subjects’ vocabulary proficiency was separately measured with Nation’s Productive Vocabulary Levels Test (2001), which is associated with the ability to use vocabulary in writing, the mean vocabulary size (mean= 3650.33, SD= 1059.08; min= 2007.78, max= 5538.89) indicated that the subjects in the study had sufficient vocabulary of English needed to communicate (i.e., write and speak) to make themselves understood in academic contexts where native English is the norm, and be able to read authentic general English texts or academic texts successfully.

The background questionnaire provided further detailed information about the subjects’ previous experience with writing. The majority (n=8, 80%) of the students had received some form of instruction for writing English compositions. With regard to their experience with English writing instruction, the students overall had received instruction after coming
to the university as a compulsory part of their English courses and in some of their English elective courses. With regard to the aspects of writing that they perceived to be important, grammar, vocabulary and word choice were the most important elements besides the non-linguistic elements such as organization and content. With their experience of using dictionaries, the students’ preferred type and medium of dictionaries were English monolingual print dictionaries and Internet bilingual dictionaries which validate findings in another study conducted with a similar group of students (Chun, 2004).

2. Tasks and Procedures

1) Writing Tasks

In order to identify lexical competence problems that were conscious to the writers in the process of L2 writing, data were collected with think-aloud protocols through which the students’ writing process could be viewed. Each subject was asked to write on two writing tasks with an interval of a week by writing with counterbalanced combinations of two audiences and two writing topics (see Table 2). The audiences emerged as a part of the writing tasks of the course where they were asked to write for two types of Internet audiences, local and global peers\(^5\) vs. general readers.\(^6\)

<table>
<thead>
<tr>
<th>No of week</th>
<th>Topic, Audience, and Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st week</td>
<td>Topic 1 Audience 1</td>
</tr>
<tr>
<td></td>
<td>1, 2, 3, 4, 5</td>
</tr>
<tr>
<td></td>
<td>Topic 1 Audience 1</td>
</tr>
<tr>
<td></td>
<td>6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>2nd week</td>
<td>Topic 2 Audience 2</td>
</tr>
<tr>
<td></td>
<td>1, 2, 3, 4, 5</td>
</tr>
</tbody>
</table>

Two writing topics for the tasks were selected based on a preliminary survey on ‘student-chosen writing topics’ that the researcher conducted prior to the think-aloud writing tasks.\(^7\) From the results, the two most popular writing topics differing in the extent

---

\(^5\) The students met the peer audience, who were mostly non-native speakers, through the Internet-based writing project International Writing Exchange (http://www.ruthvilmi.net). The project required the participants to write essays and exchange feedback on the essays with the aim of practicing to write in English.

\(^6\) The general audience, who were expected to consist of mostly native-speakers, was readers of the students’ websites uploaded at Yahoo geocities (http://geocities.yahoo.com/gcp/).

\(^7\) A group of 121 students in the same university and in similar year groups as those that would participate in the main study was asked to imagine writing compositions in English, and rate 79 writing topics on a scale from 1 (very boring) to 5 (very interesting) depending on how interesting it would be to write on them.
of cognitive demand were selected. One writing prompt asked them to ‘compare and contrast’ differences of Korean and Western culture that contribute to a successful relationship which was expected to be the more demanding task. The other writing prompt asked them to ‘describe’ strategies or methods they used to improve their English speaking skills. The topic ‘Relationships’ required the students to manipulate abstract concepts when writing to compare cultural characteristics. In comparison, the topic on ‘Strategies’ allowed more leeway for the students to use concrete words and draw on more familiar experiences as those related to learning to speak English.8

2) Think-Aloud as Method of Data Collection

Data for the study were obtained by analyzing the subjects’ TA (think-aloud) protocols while performing two writing tasks which were conducted at the end of the semester after there had been equal instruction for writing compositions and exposure to each audience at IWE (International Writing Exchange) and Geocities respectively (i.e., 6 weeks). In each week, students in their assigned groups were collectively asked to think-aloud in the language most comfortable for verbalization while recording on digital voice recorders (Forus Electronics, DVR-004) for the 70-minute TA session. They were given instructions on think-aloud and the relevant audience and writing topic. Before the first TA writing session, one session was spent on training the subjects in TA. As a demonstration, the researcher played them a think-aloud recording of a person thinking aloud in L1 while making an itinerary of a business trip after which the subjects were asked to try warm-up exercises in multiplication math problems and thinking of a list of animals (Ericsson & Simon, 1993). For further practice, the students were asked to think-aloud in pairs for a different task, a reading task, so as not to later have an effect on their writing processes. Since the study was trying to simulate the students’ usual writing processes, other facilities that were available to them were word processors with access to Internet dictionaries.9 The students were reminded to continue verbalizing also when dictionaries were used since ‘appealing’ is a part of the L2 writing process (Hemmati, 2001; Scholfield & Katamine, 2000). The researcher remained present during the TA sessions, but did not intervene because of concern that this might interfere with the writing process. However, as a procedure of data collection, subjects were asked to save their documents every 10 minutes

---

8 Discussion on the lexical problems with regard to task variables (i.e., audience and writing topic) is beyond the context of this study.

9 It was realized during the pilot and preliminary studies (Yuah Chon & Hae-Dong Kim, 2005) that when these Korean university students used dictionaries for writing, the preferred type and medium was the Internet bilingual dictionary which was available when the students were using computers to word process their documents.
as a different document (i.e., Gildong 1, Gildong 2....) on a floppy disk given at the
beginning of the session whenever reminded by the researcher since products of the
compositions at different temporal stages were needed for analysis to trace the writing
process together with the protocols.

The think-aloud method has been used in research on writing processes, but it is not
without its criticisms. Researchers have expressed doubts as to whether it may distort
writers’ normal composing processes or whether the process of composing aloud may in
fact be distinct from that of composing silently (Faigley & Witte, 1981; Dobrin, 1986).
Bracewell and Breuleux (1994) also criticize the think-aloud method for causing problems
of reliability. However, Ericsson and Simon (1993) emphasize that verbal reports when
elicted with care and interpreted with full understanding of the circumstances under which
they are obtained are a valuable and thoroughly reliable source of information about
cognitive processes. In a similar vein, Smargorinsky (1989, 1994) maintains that a
think-aloud protocol offers a unique glimpse into the workings of the human mind and
makes an important contribution to our understanding of the composing processes. As with
any other conceivable methods of data collection, think-aloud protocols are neither perfect
nor complete, but when collected with care they provide rich and direct data about the
cognitive processes of composing not afforded by other methods of investigation.

3. Data Analysis

The writing tasks yielded data in the form of think-aloud protocols recorded twice on
digital voice recorders for each of the 10 subjects yielding 20 protocols. Essay drafts and
final products written with Microsoft Word were also analyzed. In order to come up with a
coding scheme for categorizing vocabulary problems, the first step of the analysis was to
transcribe and segment the protocols with indications of what they showed (i.e., what is
written or said in English or Korean; see Appendix for a sample of the protocol and
illustration of the conventions). Through reading and rereading the protocols,
vocabulary-related sequences were extracted from the TA protocols for the categorization
of different types of lexical problems. However, knowing that vocabulary problems (e.g.,
search for words) are most likely to occur at the planning stage (Cumming, 1990, 2001;
Roca de Larios et al., 1999) or in the translating phase (Flower & Hayes, 1981) where
ideas are transferred to written text, there was need to distinguish between when subjects
were having vocabulary problems and the subjects’ process of planning content. For
identifying vocabulary-related sequences, explicit and implicit signs and evidence
identified by Hemmati (2001) were used as a framework, however, with recognition of any
additional signs of vocabulary problems (see Table 3). Of these signs, uses of the dictionary,
whether explicit or implicit, were conspicuous signs of writers’ competence problems. In
identifying competence problems, it may be argued that the writer needed to use the dictionary to check on his/her vocabulary knowledge as an ‘aide-mémoire’ (Scholfield, 1982, p. 92) so that, as mentioned before, signs of vocabulary problems were identified in conjunction with other signs.

<table>
<thead>
<tr>
<th>Explicit signs (said in L1 or L2)</th>
<th>Implicit signs</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘I don’t know’; ‘I am stuck for a word’</td>
<td>-Silence; Hesitations (i.e., er, mm, and uh)</td>
</tr>
<tr>
<td>‘It is hard to translate this word’</td>
<td>-Consulting a dictionary</td>
</tr>
<tr>
<td>‘It sounds like Konglish (=Korean and English).’</td>
<td>-Reading words, a sentence, or a few sentences preceding the target word</td>
</tr>
<tr>
<td>‘Is there no other word for ESPECIALLY, ESPECIALLY?’</td>
<td>-Language shift and saying the word or its synonyms in L1</td>
</tr>
<tr>
<td>‘Is this word an adjective?’</td>
<td>-Self-questioning or repeating a word with question intonation</td>
</tr>
<tr>
<td>‘Let me look it up in the dictionary’</td>
<td>-Mismatch between what the writer says in L1 and what s/he writes in English</td>
</tr>
<tr>
<td>‘These are absolutely not the words I want’</td>
<td>-Saying a series of incomplete and complete words (wrong or right)</td>
</tr>
<tr>
<td>&lt;said after looking at a dictionary entry&gt;</td>
<td>-Retrieving words related in sound or meaning to the intended word</td>
</tr>
<tr>
<td>‘I have used this word too many times.’</td>
<td></td>
</tr>
<tr>
<td>‘I am confused about these words’</td>
<td></td>
</tr>
</tbody>
</table>

With the use of the computer and Internet dictionaries in the writing process, the TA gave additional signs as to when subjects were having vocabulary problems. Repeated verbalization of the problematic word in L1 or L2 followed by the sound of brisk keyboard typing which may be accompanied by immediate report on dictionary entry or on the outcome of the look-up gave direct signs about the presence of word problems. Consider the following example:

If one has a deluxe car or wears an expensive cloth,…then he is, he is a [be treated, be treated, treat ] <brisk keyboard typing heard indicating consulting a dictionary> treat… [okay]

Here the writer repeats unknown words in Korean in the face of an unknown word problem, immediately refers to an online dictionary, and then confirms to himself that the word ‘treat’ is appropriate to use. Hence, as seen in the example, for a protocol sequence to be considered as containing a vocabulary problem, solitary signs standing alone were often not enough. Signs and evidences of vocabulary problems needed to be looked at in combination with other signs to make sure they were clearly problems caused by lack of lexical knowledge. Another point that needs attention in the present study is that to classify any vocabulary problems, the writer had to be ‘conscious’ of the problems so that only ‘self-initiated’ problems were included in the classification. In our analysis, an instance
where a word is used in some wrong way without any sign of the writer being aware of the problem was not included as a vocabulary problem.\footnote{Those word problems that the writer is not aware of and left unnoticed would be ‘errors’.

As one procedure for coding the types of vocabulary problems, samples of the protocols were shown to a second judge, a lecturer in applied linguistics who had conducted various writing studies, for validation of the coding scheme so that any points that the researcher herself had missed could be modified and re-coded. The intra-judge reliability was taken care of in the process. When approximately seven percent (50 instances for the principle of convenience) of all the protocols of the whole items (654 instances) were randomly chosen and checked by the researcher, there was 90\% intra-judge reliability. Regarding inter-judge reliability, the same seven percent of the instances were checked by a third judge, a professor in applied linguistics, and as a result indicated 88\% reliability. The instances of vocabulary problems that had been coded differently were discussed between the judges and modified when deemed necessary.

IV. RESULTS AND DISCUSSION

1. Classification of Vocabulary Problems

Having identified the signs of vocabulary problems, the TA protocols were categorized according to the types of problem. According to Hemmati (2001), if a vocabulary problem is due to lack of writer’s insufficient knowledge of the word, it is referred to as \textit{competence-based}, and when a writer has sufficient knowledge but experiences some problem arriving at the target word it is \textit{performance-based}. The interest of the study here is only to investigate the competence-based problems. Utilizing Hemmati’s (2001) coding scheme of lexical competence problems, the analysis of the TA protocols produced three types of lexical competence problems. The three vocabulary problem situations are classified as NO WORD, ONE WORD, and MORE THAN ONE WORD problems (see Table 4 for details).

\begin{table}[h]
\centering
\caption{Competence Vocabulary Problem Situations in the Writing Process of EFL Students}
\begin{tabular}{l}
\hline
Types of word problem \\
\hline
I. NO WORD problems \\
- The subject does not know a L2 word that s/he knows in L1 \\
II. ONE WORD problems \\
- The subject retrieves a L2 word but some aspect of the word is unknown \\
III. MORE THAN ONE WORD problems \\
- The subject retrieves more than one L2 words and is confused in what word to choose \\
\hline
\end{tabular}
\end{table}
In the following sections, the categories and subcategories of vocabulary problems will be explained in more detail with signs of each problem and examples from the protocols.

1) NO WORD Problems

NO WORD problems are when the target word that the subject has in mind in L1 does not exist in the ML which we call the ‘unknown word’ problem so that it is a ‘competence’ problem. Indications of ‘unknown word’ problems were:

- Explicit questioning such as ‘How can I say XX?’
- Repeating L2 words preceding target word
- Repeating translation of target word in L1
- Mismatch between what the writer says in L1 and what is written in L2
- Shift to L1 to say or question the target word
- Dictionary-consulting
- Pausing

Consider the following example of an unknown word problem:

[bribery cannot be exterminated] ……. this is the same way that the bribe is not disappear, bribe is not disappear. This is the same way that the bribe is not disappear in Korea.

In the example, there is mismatch between what the writer says in L1 and what is written in L2 which indicates a NO WORD vocabulary problem. Here the writer’s target word is ‘exterminated’ but he does not know it. In place of the unknown word, he retrieves the word ‘disappear’ considered closest to his intended word after a pause.

2) ONE WORD Problems

The category of ONE WORD problems can occur when a subject retrieves a word but finds his/her knowledge incomplete in some aspect of the word due to lack of competence so that these are also called ‘partially known one word problems’. The knowledge may have never been learnt or forgotten by the subject. These ONE WORD problems could further be classified as cases when a word retrieved is partially known due to lack of knowledge of the meaning, spelling, lexical grammar, collocation, existence, and when an alternative word is not known for a different style or to avoid repetition or to elevate the sophistication level of a word. The ONE WORD competence problems more or less coincide with what Hemmati found in her study of Iranian university students but three
Identification of Lexical Competence Problems

different categories have been added: style, repetition and unsatisfied with sophistication level. The details of ONE WORD competence problems are explained in the following.

(1) Meaning of Retrieved Word not Completely Known Problems

This is the situation when the writer retrieves a word but its meaning is either not entirely known or the writer is uncertain about it. The signs that indicate this type of problem are explicit statements such as ‘Is this right?’, ‘Can I use the word for this meaning?’, and implicit signs were ‘Is this strange?’, or ‘Shall I look it up/check it in the dictionary?’, some of which can appear as signs of vocabulary problems in other subcategories as well. Consider the following example of the meaning partially known word problem. Here the subject retrieves the word contents but is not sure if that is correct to use for the target meaning ‘content’ of a lecture, and consults the dictionary to check on his unsure knowledge.

Recording3 [the content of a lecture] recording… some contents [Is it correct to say CONTENTS?] recording some contents [I think it would be right to refer to content of a lecture; let me look it up in the Eng-Kor dictionary] contents…

(2) Spelling Partially Known Problems

The spelling partially known situation is when the writer perceives a gap in his knowledge of spelling of the intended word that needs to be produced. The subject may know the meaning and pronunciation of a word retrieved, but is unsure of how to write it. Consider the following example where the writer lacks spelling knowledge so that the writer uses the dictionary to find the spelling of ‘grammar’ when he is not sure whether its ending is ‘-ar’ or ‘-er’.

If you have a good skill, have a good skill to, skills for grammar, [g-r-a-m-m-e-r; Is it ‘-a-r’? g-r-a-m-m, spelling, g, gram-mar, gram..........<long pause and sound of keyboarding indicates subject’s dictionary search instance> ..... [Oh, grammar; ...... spelling, They are all ‘a’.]}

It needs to be noted here that with use of the computer in the writing process, spelling problems recognized by the spellchecker (in Microsoft Word) were not included in the analysis since they were not self-initiated or conscious to the writer. There was interest solely in problems recognized by the subject since they had to be aware of them to use solutions (i.e., looking for spelling information in the dictionary).
(3) Lexical Grammar Partially Known Problems

This refers to the situation when a subject retrieves a word, but perceives lack of knowledge about the grammatical aspect of the word. Hence, this category is labeled lexical grammar (rather than grammar) since this is a problem related to the grammatical traits of a word. Also since grammar is a broad term, they were subcategorized in Hemmati’s study, so that she identified problems of ‘part of speech’, ‘verb form’, and ‘complementation’ (which she labels ‘word’s syntactic pattern in sentence’). In our study, lexical problems also appeared in the same categories, but ‘non-count/count noun’ problems also arose in the subjects’ writing process. This may be due to the influence of the L1, Korean, where there are no words such as articles. Each lexical grammar problem will be explained and exemplified in the following with protocol excerpts.

① ‘Verb Form’ not Known

This refers to the situation when a subject questions the inflected form of a verb that has been retrieved. As Hemmati categorized, this subcategory has to do with tense and transitivity. In the present study, there was only one instance of this problem which was solved through a student appealing to the teacher during the writing task. While the think-aloud writing session was in progress, one student asked the researcher/teacher as to what ‘give lectures’ was in passive form, and the researcher helped him to come up with ‘where lectures are given in English’ after looking at the student’s sentence. The rarity of this type of problem may show that these Korean EFL university students are more confident in lexical grammar compared to other problems of writing like lexical meaning. This may be due to the fact that they have been trained previously with more emphasis on form (e.g., grammar exercises) than on practicing communicating their intended messages.

② ‘Part of Speech’ not Known

This is the situation when a subject retrieves a word but finds that the part of speech is not known for his/her intended meaning.

- Explicit signs are:
  - ‘What is the verb of this word?’
  - ‘Is word A (not) an adjective/verb?’

- Implicit signs are:
  - Shift to L1 to remind subject of the target part of speech via problematic word in L1
  - Solution chosen (dictionary-consulting)

Consider the following example where the subject retrieves ‘response’ but is not sure of the part of speech. He then checks his knowledge in the ML to see if this is a noun, but is still not completely sure so he checks with the dictionary, and succeeds in his attempt to check the information.
stimu-, stimulation, okay, no stimulation, no response [Is RESPONSE not a verb? <types RESPONSE at online dictionary> RESPONSE, RESPONSE, okay, this is a noun, isn’t it?] No stimulation, no response....

3 ‘Non-count/Count noun’ not Known
This is a subcategory where subjects question the countability of a noun retrieved.

- Explicit sign of this problem is:
  - ‘Is the word a countable noun?’
- Implicit signs are:
  - Questioning the word
  - Questioning the two alternative forms (in singular and plural forms of the word)
  - Solution chosen (dictionary-appealing)

There were only two instances of this in the ONE WORD competence category. Consider the example where the subject is not sure whether the noun word ‘industry’ has to be used as a singular or a plural. The subject questions countability of ‘industry’. From his own knowledge he thinks it is a countable noun, but is still unsure so he decides to look up the word in the dictionary.

[Does it have to be INDUSTRY or INDUSTRIES? Shall I look it up in the dictionary? As I know it INDUSTRY is COUNTABLE but since it can be different from what I know, let me look it up in the dictionary]

4 ‘Complementation Pattern’ not Known
Here ‘complementation’ refers to the part of a sentence that follows the verb or a noun and whose choice is dependent on it to complete a sentence. Hence, competence complementation pattern problems occur due to lack of lexical grammar knowledge. For instance, a writer may not know what preposition to use to complement a verb that is retrieved for use. In the following, there are two complementation problems occurring in sequence where the problem word is ‘attended’ after which the writer does not know which preposition to use between ‘in’ or ‘with’. Use of the dictionary validates that it is a competence problem where he finds that ‘attend’ does not need to be complemented by a preposition.

mmm for example 2, I ummm I am am I am attending attending classes [Is it ATTENDING IN THE CLASSES?] attended attended [attend2] attended with attended with attended [if I look up the dictionary, it is just ATTEND] I am attending the classes which ummm which are ummm...
(4) Collocation Partially Known Problems

The collocation partially known problem is a situation occurring when a subject retrieves a word but is not sure what collocates before or after the word. For instance, the subject wants to say ‘learn something by heart’, but is initially not sure what to say before ‘heart’. The writer manages to say ‘by heart’ but is not sure so has to look it up in the dictionary. She, however, wrongly uses ‘by heart’ as a verb.

words…[remember is..] doing heart [what was it?] by heart [I know there was ‘something’ HEART; ….to remember, to remember completely is…was it BY HEART? If I look up the idiom <BY HEART was in the entry>..... ] mmmm,
by heart, [memorize by heart] by heart. When you meet wholly new words or idioms, you can use them later if you byheart.11

(5) Doubt about Existence Problems

This type of problem only occurred as a competence problem in the protocol since the subjects were able to solve these only through alternative sources such as the dictionary rather than via retrieval of the writer’s existing knowledge from the ML. These problems arose out of word coinage, literal translation and loanwords. In other studies (Harvey & Yuill, 1997; Hemmati, 2001) these problem originated from coined words, literal translation and cognates, but in our study they arose from loanwords instead of cognates in the strict sense12 since Korean is not an Indo-European language where there are likely to be words with the same origin as the other. Consider the following example where the problem has arisen from lack of knowledge of loanwords:

For example [for example, for example, drama Friends, for example, for example] ……the movie, movie and, movie that, the movie you watched [nuhhhhh, nuuuhh] for example uh, TV program, for example ahh TV program [Is there such a word as drama, drama? Drama, drama, <looks up drama at Korean-English dictionary> [Ohh, TELEVISION, TELEVISION SERIAL DRAMA] television serial drama, ohhh [that’s so good!!] television serial drama, drama is …a television serial drama, like Friends, f-r-i-e-n-d, Friends or uhhh Sex and City, City, television serial, serial drama is

Here this low-proficiency writer retrieves the Korean loanword ‘drama’, but is hesitant

---

11 Missing the leaving of a space between ‘by’ and ‘heart’ was an error committed by the writer.
12 ‘Cognates’ and ‘loanwords’ are distinguished in historical linguistics, whereas ‘cognates’ in applied linguistics may include ‘loanwords’ and other similar words (Personal communication; Scholfield, 2006).
Identification of Lexical Competence Problems

(6) Word of Alternative Style not Known Problems

This is the situation where the subject finds the style of a retrieved word to be unsuitable for the intended meaning but lacks lexical knowledge to think of an alternative word to suit the intended word style. The only way to detect the problem is through the report from the subject that s/he does not like the style of the retrieved word. Dictionary-consulting is a solution as well as an indication that the problem is in the competence category. There was only one case of this and this was when the writer retrieved ‘on the earth’ (which should have been ‘why on earth’) as an alternative to ‘really’. However, the writer senses that this is an expression to be used when one is surprised about something, implying that the word is not formal enough, and not reflecting the emphasis he wants and therefore looks it up in the dictionary for alternative expressions. The dictionary entry results are not satisfactory for solving the word problem and the following shows that he eventually changed his message to ‘directly or indirectly’, which is not mentioned in the TA report.

let’s take a look at umm on which field on which field they are really cooperating, really cooperating2, then let’s take a look at on which field they are really cooperating umm [why an earth 4? On earth, in which field] On the earth, [ON THE EARTH, not these words, this is to be used when you are surprised about something, let me look up the dictionary <Kor-Eng> <looks up the dictionary and reads the entry ➔ > in the world; on earth; under the sun; in the name of God; the dickens; the hell ..Really you are to blame for it, Does he know any English at all? <hums to himself> What the deuce is the matter? Whatever is the matter? … <continues reading dictionary entry> [Are these all I have? ]

(As a final product writes: Then, let’s take a look at in which field they’re directly or indirectly cooperating one another.)

(7) Alternative not Known for Repetitive Word Problems

This is the situation when the writer retrieves a word but realizes that s/he has used that word repetitively. Subjects are unable to solve problems with their own linguistic knowledge and have to appeal to another source (e.g., thesaurus). In the following, the
writer may retrieve the word ‘difference’ but realizes that the word is being repeated, so tries to solve this by looking it up in the thesaurus. However, when the thesaurus does not give her an immediate solution, she reverts to using the problem word ‘difference’.

[Is there not another word to talk about the difference other than DIFFERENCE? Not DIFFERENCE…. I have used DIFFERENCE too commonly so let me look up the thesaurus. …the difference points 5, Do I have to look up with difference? It is a difference <I have to write about>, but I have to keep using DIFFERENCE. DISTINGUISHABLE2 umm DISTINGUISHABLE2 DIFFERENCE? No that is strange; haaa, Is there no other word other than DIFFERENCE? CONTRAST is contrast, OPPOSITE does not make sense, umm DIFFERENCE, inevitably I have to use DIFFERENCE.] Difference, the difference in…..

(8) Alternative not Known for Word Deemed to be Unsatisfactory with Sophistication Level Problems

This is an instance when the subject retrieves a word but is dissatisfied with the sophistication level of the word. In addition to repetition problems, it was regarded as a problem with a particular instance of style, but they have been categorized separately since it is too general to be included under the term. The signs to indicate the problem may seem similar to the ONE WORD repetition problems, but a word was coded as such when the word retrieved was not used previously in the protocol, and also when the subject specifically states dissatisfaction with the sophistication level of the word. Consider the following example of a competence sophistication problem where the writer wants a different word from the very common word ‘want’ but has to look it up in the Microsoft thesaurus due to lack of lexical knowledge. In this case, he immediately notices an alternative word ‘desire’ and proceeds to use it in his composition.

[Isn’t there another word other than WANT? Let me look at the thesaurus; there is DESIRE, DESIRE! DESIRE is good] desire to? These days many people have the desire to improve their Eng speaking skill for their career or …..

With regard to coding, protocols as such may be subject to argument. Although it has been categorized as a competence problem, it can also be identified as a performance problem by others since he seems to know ‘desire’ when he sees it. He cannot retrieve a better word for ‘want’, so he looks in the thesaurus and this reminds him of a better word he already knows. This process resembles the usual L1 user’s use of a thesaurus as an ‘aide-mémoire’ for retrieval. However, when L2 writers appeal, they are likely to use the dictionary or thesaurus more due to lack of knowledge, and as seen in the example above,
although the writer seemed to know the word ‘desire’, the use of the thesaurus did not convince the researcher to consider it as a performance problem. In fact, while analyzing the protocol, the researcher sometimes felt that students were feigning knowing some of the new words they had found for fear of being graded low for the course they were enrolled in.

3) MORE THAN ONE WORD Problems

This is the situation when a writer simultaneously retrieves more than one candidate word to be used for his intended meaning but is confused about the words so this problem may also be called ‘word confusion’. This is due to lack of knowledge about the words or a competence-based problem. Hemmati (2001) recognizes confused words that are similar in both meaning and form (e.g., society/sociology, compete/competent; dependent/independent, make/made) or similar only in form (e.g., export/expert, well/will). They are similar to what has been coined by Laufer (1991) as ‘synforms’, similar lexical forms, or more generally known in linguistics as ‘near-homonyms’. This term was coined after Laufer studied errors of language learners and native speakers after which she tried to define, illustrate, and classify words with similar forms that cause confusion in language acquisition. She puts two different types of linguistic similarity under the term synformy: “words which have similar form since they are related to each other etymologically, whether the relation is transparent (industrial/industrious) or not (assumption/consumption); and words whose similarity in form is purely accidental (lunch/launch)” (p. 53). Though the two types are quite different linguistically, she treats them as one phenomenon since they result in the same feature of interlanguage, confusion of one word with another which has a similar form.

In the present study, word confusion problems arose more broadly due to writer’s lack of knowledge of near synonyms (similar meaning, different form; e.g., trouble/difficulty) as well as synforms/homonyms (similar form; e.g., personal/personnel). In our study, we identified word confusion problems by seeing that they had been solved via dictionary/thesaurus appealing or belated arbitrary choice accompanied by hesitation or pauses which are indications of competence problems. In the following, the writer retrieves words ‘adopt’ and adapt’ simultaneously, but finds the words confusing. He looks it up in the dictionary via E ‘adapt’, and succeeds in finding that this is the right word of the two but fails to write in active voice with the correct preposition, that is, ‘cannot adapt to’.

Therefore, this situation can be problem to some people who cannot adapt; Is it ADAPT? Or is it ADOPT? Let me find it; I think it was ADAPT] adapt, be adapted for the new relationships and rule in keeping relationships.
2. Frequency of Competence-based Vocabulary Problems

Regarding competence-based vocabulary problems, a total of 368 instances occurred in the 20 protocols written collectively by 10 subjects during the two sessions. This was 56.27% of the total vocabulary problems that the writers had in the study.\(^1\) In comparison, the writers in Hemmati’s study had a higher percentage of competence-based problems (66.84%) which possibly explains that the students in Hemmati’s study may have planned content that was beyond their vocabulary competence so that they had more competence problems compared to the writers in the present study. Another factor may be that since vocabulary problems have to be identified by the writers, this could be due to the different level of ‘awareness’ that the writers exhibited in each study when recognizing word problems in the process of writing. However, when it was realized that the length of essays varied (min=161 words; max=572 words) in spite of the same writing conditions (e.g., time and availability of resources such as the word processor, and Internet dictionary which they were not forced to use), and so of opportunities of problems, it was fairer to work in problems per 100 words rather than raw numbers of problems for the analysis. The calculations for the three types of competence problems are presented in Table 5.

<table>
<thead>
<tr>
<th>Raw Mean Frequency &amp; Mean Frequency per 100 Words of Vocabulary Problems</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total no. of problems=368</strong></td>
</tr>
<tr>
<td>NO WORD Unknown word (167/45.38%)*</td>
</tr>
<tr>
<td>ONE WORD Partially known word (187/50.82%)</td>
</tr>
<tr>
<td>MTOW Word confusion (14/3.80%)</td>
</tr>
</tbody>
</table>

*Note*: *(raw frequency of word problem/ percentage); MTOW = More than one word problems

Results in Table 5 show that while both ‘raw mean frequencies’ and the ‘mean frequencies per 100 words’ ranked in the same order across subjects (for the 2 essays each writer had written), each subcategory of vocabulary problems was highest for ONE WORD partially known words (mean=18.7; mean/100 words=6.56) and lowest for MORE THAN ONE WORD word confusion problems (mean=1.4; mean/100 words=.40). However, from the large SDs we can see that the frequency of the vocabulary problems varied a lot between the subjects. Among these, ONE WORD partially known words could be further divided

---

\(^1\) Here the present study is conducted with the exclusion of other types of lexical problems (i.e., performance-based and dictionary-based problems) that the researcher found in a larger study.
into subcategories. ONE WORD partially known problems consists of not being sure about meaning, spelling, lexical grammar, collocation, existence of word, style, and not being able to solve repetition and sophistication problems with one’s own knowledge where writers are not able to access good synonyms for the intended meaning. One problem could not be identified for its source of problem (see percentages and frequencies noted in Figure 1). With regard to competence lexical grammar problems, subjects were not sure about part of speech (8 instances), verb form (1), countability (2) and complementation (9).

FIGURE 1
ONE WORD Partially Known Problems

<table>
<thead>
<tr>
<th>Problem Type</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meaning not sure</td>
<td>35.1%</td>
</tr>
<tr>
<td>Sophistication</td>
<td>14.1%</td>
</tr>
<tr>
<td>Repetition</td>
<td>17.8%</td>
</tr>
<tr>
<td>Doubt existence</td>
<td>3.8%</td>
</tr>
<tr>
<td>Collocation not sure</td>
<td>7.6%</td>
</tr>
<tr>
<td>LexGrm not sure</td>
<td>10.8%</td>
</tr>
<tr>
<td>Spell not sure</td>
<td>10.3%</td>
</tr>
<tr>
<td>Style</td>
<td>.5%</td>
</tr>
</tbody>
</table>

Note: category unidentified .5% (1) not included in graph

In a similar vein, Harvey and Yuill’s (1997) study reports on what was looked up by EFL writers at intermediate level when they were asked to use a monolingual pedagogical dictionary while engaged in writing. The writers’ introspective report revealed that the look-ups were motivated by gaps in the order of ‘spelling (24%) – meaning (18.3%) – existence (12.8%) – synonymy to solve a repetition problem (10.6%) – grammar (10.5%) – register (9.3%) – collocation (8.2%) and – inflection (5.9%)’ when writers had one, two, or more than two reasons for looking up the dictionary. As such, spelling and meaning were looked up the most in Harvey and Yuill’s study while in our study the reasons for the look-ups were in the order of ‘meaning- repetition- sophistication- lexical grammar- spelling not sure- collocation not sure- doubt about existence- style’ of the ONE WORD partial knowledge problems. The reason for ‘spelling’ and ‘existence’ lagging behind in the order of look-up is due to the different writing medium, the word processor where spellcheckers could solve their spelling problems even before students noticed them. Interestingly however, when we disregard ‘spelling’ and ‘existence’, the order of the first three items looked up across both studies are in the same order as in: ‘meaning –
synonymy (e.g., to solve repetition/sophistication word problems) – grammar’.

V. CONCLUSION

1. Main Research Findings

It was found in the study that Korean EFL writers identified different types of competence lexical problems when translating (Flower & Hayes, 1981) conceived ideas to visible text. Competence problems comprise situations when a word is completely ‘unknown’ (NO WORD), when a word retrieved is ‘partially known’ (ONE WORD) and when ‘more than one words retrieved are confusing’ (MORE THAN ONE WORD) to the writer. We also found that ONE WORD partially known word problems can occur due to lack of knowledge in aspects of meaning, spelling, lexical grammar, collocation, existence, or when alternative words are not known for a different style, to avoid repetition or change the sophistication level of a word. There was reference to lexical competence problems in the ‘problem-solving formulation’ process which acts to compensate for lack of linguistic resources when L2 writers’ interlanguage knowledge is unstable for conveying an intended meaning. (Roca de Larios et al., 1999, 2001, 2006). However, the scope of these previous studies is limited since the researchers neglect to classify lexical problems and confuse them with what are the solutions to the problems. In relation to Hemmati’s (2001) study, the classification derived in the present study is more plausible since parallel competence and performance categories have been derived for ONE WORD problems.

2. Implications of the Study

While there is plethora of L2 writing literature on the composing processes of planning (e.g., Cumming, 1989; Roca de Larios et al., 1999; Sasaki, 2000) and revising (e.g., Hall, 1990; Manchon, Roca de Larios, & Murphy, 2000; Zamel, 1983), current literature lacks analysis of what happens in the translating phase of writing which is where vocabulary problems are most likely to occur. Even the best-known models of written production (Bereiter & Scardamalia, 1987; Flower & Hayes, 1981; Hayes, 1996) which have been influential in L2 writing research have given scant attention to the process (Hemmati, 2001; Roca de Larios et al., 2001 are exceptions). These models make no reference to the complexities involved in trying to convert thoughts into language, perhaps because it is taken for granted that the process of matching intentions and expression is automatic. On the other hand, this study contributes to L2 writing research by trying to provide a psychologically-plausible model of what occurs in the translating phase with regard to
what and how writers identify one main category of vocabulary problems, competence lexical problems.

3. Recommendations for Future Research

This study provides recommendations for several branches of research. In this study, only the competence lexical problems were explained, but other types of problems that the researcher found, such as performance lexical problems (e.g., word retrieval problems that occur due to temporary inaccessibility of lexical knowledge), or dictionary-based problems (e.g., one of the dictionary words is not known or inappropriate for the intended meaning when the dictionary is consulted) are worthy of investigation. Classification of the different types of lexical problems would also provide a foundation for investigating what L2 writing strategies are used to solve the different types of problems. Findings would contribute to model-building of the L2 writing process with respect to translating.

REFERENCES


**APPENDIX**

A sample of think-aloud protocol and conventions

- **Underlined text** indicates the verbalization of what is or has been written by subject in English.
- **Normal text** indicates verbalization said or resaid in English.
- **Series of dots…..** indicate a pause in the subject’s verbalization.
- **Text in [square brackets]** indicates verbalization in L1 that has been translated.
Many students find difficulty [right], find trouble? trouble...[Shall I use TROUBLE or DIFFICULTY?] difficulties in learning the English language, but ahhh many students find trouble or difficulties in learning the English language [I did not use capital letters] learning the English language but...[let me finish the sentence at this point] Many students find trouble trouble, trouble, difficulties [difficulties2, let me use the dictionary...difficulty, DIFFICULTY...DIFFICULTY is better] Many students <remembers to write the name of the dictionary> Many students find it difficult...... to learn the English language, to learn the English language [How do I say that speaking is the most difficult?] many students find it difficult...... [no, no, FIND DIFFICULTIES is better than FIND IT DIFFICULT] learning [right] learning the English language. [no, I will write first] However, speaking is the most difficult part [speaking is the most difficult part] difficult part.

Applicable levels: tertiary education, adult education, general education
Key words: EFL writing, process writing, vocabulary problems, think-aloud protocols, communication strategies