Available online at www.sciencedirect.com
ScienceDirect
Journal of Second Language Writing 18 (2009) 191–208
JOURNAL OF SECOND LANGUAGE WRITING

Aspects of lexical proficiency in writing summaries in a foreign language

Kyoko Baba *

Kinjo Gakuin University, Department of English, 2-1723 Omori, Moriyama-ku, Nagoya 463-8521, Japan

Abstract

This study investigated the impact of aspects of the lexical proficiency of EFL students on their summary writing in English (L2) by controlling for the impact of a range of linguistic abilities in English and Japanese (L1). Sixty-eight Japanese undergraduate students wrote two summaries of English texts in English. Their English lexical proficiency, English reading comprehension, English proficiency, knowledge of Japanese vocabulary, and writing proficiency in Japanese as well as the length of summaries were assessed. Multiple regression analysis of the data showed that the effect of L2 lexical proficiency as a whole on summary writing performance was not pronounced compared to the effect of reading comprehension and the length of summaries. However, the ability to write definitions made a unique contribution over and above the other variables including reading comprehension and the length of summaries. It is suggested that different aspects of L2 lexical proficiency have a differential impact on EFL learners' summary writing, and that two factors in particular (structure of semantic network of words, and the ability to metalinguistically manipulate words) may constitute the construct of summary writing in L2.

Keywords: Lexical proficiency; Summary writing; EFL writers; Multiple regression; Word definition; Reading and writing; Writing in academic contexts; Second language writing

Introduction

One of the most typical and critical academic writing skills for L2 learners is summary writing, for ''it is impossible to assign academic writing tasks that don't require preliminary reading'' (Johns, 1993, p. 277). Some empirical studies have shown that reading-based writing is vital in academic training across disciplines (Carson, 2001; Cumming, Kantor, Powers, Santos, & Taylor, 2000; Horowitz, 1986). Reading-based writing tasks are, therefore, real-life challenges for L2 learners in academic contexts. Previous research on summary writing in L2 has mainly investigated how L2 writers summarize texts by examining the discourse features of summarized texts (e.g., Cumming, Rebuffot, & Ledwell, 1989; Johns & Mayes, 1990; Keck, 2006; Kim, 2001) and writing processes and strategies used during summarization (e.g., Sarig, 1993; Yang & Shi, 2003). For instance, Keck (2006) studied the paraphrasing strategies of L1 and L2 university students that were used to write a summary. The L1 students modified the source text to a greater extent than the L2 students did, and the L2 students relied on the source text more heavily. On the basis of an analysis of the paraphrased sentences, Keck speculated that students' language proficiency might have affected the degree of

* Tel.: +81 52 798 4854; fax: +81 52 799 2089.
E-mail address: kbaba@kinjo-u.ac.jp.

1060-3743/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jslw.2009.05.003

their dependence on the source text in the sense that some students copied from sources possibly ''because they lacked the linguistic resources'' (p. 275) needed to paraphrase. In fact, Keck showed that paraphrasing necessitated knowledge of synonyms and clause structures.

However, there has been little research into what linguistic knowledge and abilities are needed to write L2 summaries. The present study examined the language abilities and knowledge that may play a prominent role in how L2 learners write summaries in an attempt to identify the component language abilities that are needed to explain the construct of summary writing.

Relations between lexical proficiency and writing in L2

There may be various language abilities and knowledge that are called on when L2 learners write a summary, but the central focus of the present study is lexical proficiency in L2. Grabe and Kaplan (1996) underscored the significance of vocabulary for basic writers, maintaining that ''vocabulary development not only supports reading and writing, it also promotes syntactic flexibility and creates a foundation for further learning'' (p. 275). Language learners are aware of the significance of vocabulary, and they often consider a lack of knowledge of vocabulary to be the cause of deficiencies in their writing skills. In a survey of ESL students' perceptions of training in writing English for Academic Purposes (EAP), Leki and Carson (1994) found that knowledge of vocabulary was the most frequently mentioned L2 knowledge that students wanted. Protocol studies have also demonstrated that L2 writers often face word-related problems (e.g., Cumming, 1990; Raimes, 1985) and devote considerable attention to lexis during writing (for a review see Manchón, Murphy, and Roca de Larios, 2007). Other studies have analyzed L2 writers' texts and suggested that the size of vocabulary available to the writer plays an important role in L2 writing. For instance, more proficient L2 users (or native speakers) use a wider variety of words and more sophisticated (e.g., low-frequency) words in their writing than do less proficient L2 users (Harley & King, 1989; Laufer & Nation, 1995; Linnarud, 1986). Likewise, an analysis of ESL learners' compositions (Engber, 1995) showed that L2 compositions that use a wider variety of words tend to be rated more highly than those with a narrower one.

In addition to breadth of vocabulary knowledge, previous research has also suggested that the semantic network of words in a learner's mind (also called depth of vocabulary knowledge, Haastrup & Henriksen, 2000; Meara, 1996) impacts on fluent topic progression in writing (Schneider & Connor, 1990). For example, Reynolds (1995) showed that proficient L1 writers elaborate and develop the key concept (*special event*) of an essay by employing hypernyms (e.g., *parties*) and a variety of related verbs (*take place*, *played*, *held*), but texts written by L2 writers tend to include redundant content by repeating the same words and concepts. Reynolds' finding supports Hoey's (1991) claim that word relations in texts are the key factor in the development of coherence. Hoey describes how the use of word relations such as *complex lexical repetition* (in which two words share a lexical morpheme or have the same form with a different part of speech) and *simple paraphrase* (in which two words are exchangeable in meaning or at least one word can replace the other word) may enhance coherence in writing and avoid the redundancy generated by *simple repetition* (in which the same word is repeated without any change or with only a minor change).

There are, however, two limitations in the studies reviewed so far on the role of lexical proficiency in L2 writing. First, they analyzed texts written and protocols produced by L2 learners, but did not directly assess L2 writers' vocabulary knowledge, which might not have been displayed in their writing and protocols. Engber's (1995) study, for instance, indicated that compositions that exhibited a wider lexical diversity tended to be rated more favorably, but this finding did not show how well, or how fully, L2 writers utilized their vocabulary knowledge when writing. Yet, vocabulary knowledge itself and the ability to use it can be different (Chapelle, 1994) in that all vocabulary knowledge is not fully available for use. For example, Uzawa (1996) observed that Japanese adult EFL writers produced sentences at an i-1 level (a linguistically lower level than the level they were able to understand) because they used ''only words, expressions, and syntax readily accessible to them'' (p. 288). Furthermore, the aforementioned studies did not examine the extent of the effect of learners' vocabulary on their writing (but see Koda, 1993; Schoonen et al., 2003 for exceptions).

Koda (1993) investigated the impact of linguistic knowledge on L2 writing ability of English-L1 college students learning Japanese as a foreign language. In her study, the correlation between students' vocabulary knowledge (assessed with a word definition task in their native language) and the quality of their essays was around $r = .7$. Regression analysis showed that the single strongest predictor of their writing was lexical diversity, which explained

roughly half of the variance in L2 writing performance.[1] Koda also conducted a topical structure analysis of the compositions written by students. She found that overall ratings of compositions strongly correlated with the degree to which the students developed subtopics in their writing. On the basis of the finding that the quality of the composition had a high correlation with vocabulary knowledge, Koda suggested that "knowledge of content-word meanings underlies the ability to expand and elaborate preceding subtopics in discourse" (1993, p. 343). This suggestion is consistent with Reynolds' (1995) finding mentioned above that skilled writers utilized the knowledge of semantically related words to develop a topic in their writing.

While Koda (1993) emphasized the significance of vocabulary knowledge in writing, the results are not conclusive when we consider its impact on L2 writing in relation to other language abilities. This is concerned with the second limitation in the studies that have investigated the relationship between lexical proficiency and second language writing proficiency. That is, previous research has largely disregarded the fact that the effects of lexical proficiency are relative to those of other abilities. For example, Schoonen et al. (2002, 2003) took a componential approach to this issue and examined the relative importance of L2 linguistic knowledge, metacognitive knowledge, fluency of linguistic knowledge processing, and L1 writing proficiency to L2 writing performance by eighth-grade Dutch students learning English. As in Koda's study, there was a clear correlation between L2 vocabulary knowledge (breadth of vocabulary knowledge) and L2 writing proficiency (about $r = .6$). However, when the effect of vocabulary knowledge was assessed in comparison to other language abilities (L1 writing proficiency, grammatical knowledge, orthographic knowledge, metacognitive knowledge, speed of word retrieval, and speed of sentence building in L2) in their structural equation modeling analysis, vocabulary knowledge did not make a unique contribution to L2 writing proficiency. In particular, its effect was suppressed due to a high correlation with metacogntive knowledge. This result implies that although vocabulary knowledge is important to L2 writing, its effect is overshadowed by other abilities that require similar common cognitive knowledge or skills. In addition, L1 writing proficiency had a stronger impact on L2 writing proficiency than both L2 linguistic knowledge and how fluently this knowledge could be processed. Heeding the results of Schoonen et al.'s studies, the present study assesses the importance of lexical proficiency in relation to other relevant abilities (see below). What is more, so far no study has examined the relationship between L2 writers' lexical proficiency and summary writing performance. In summary writing tasks, the content to be expressed is given and thus, the effects of differences in writers' background knowledge are neutralized to some extent (Watanabe, 2001). Therefore, if lexical proficiency has any impact on summary writing performance, it is likely that the impact of writers' lexical proficiency on other types of writing is nonnegligible. The present study aims to extend the componential analyses of L2 writing and contribute an understanding of the role of lexical proficiency in L2 summary writing.

Theoretical distinctions of L2 lexical proficiency

It is worth noting here that what the aforementioned studies refer to as "vocabulary knowledge" is not always identical, which makes it difficult to compare the results from these studies. For example, breadth and depth of vocabulary knowledge and lexical diversity in writing are all different constructs in a strict sense. The multiple-choice vocabulary test in Schoonen et al.'s (2002, 2003) studies might have elicited a different construct from the lexical diversity assessed in Koda's (1993) study. The difference in these explanatory factors may have resulted in the different findings of these studies on the contribution of lexical proficiency to L2 writing. In the present study, *lexical proficiency* is used as an umbrella term to cover distinct constructs of vocabulary knowledge.

On the basis of theories of L2 lexis, this study concentrated on the two fundamental distinctions between aspects of L2 lexical proficiency: breadth versus depth of vocabulary knowledge, and receptive versus productive aspects of vocabulary use. Simply put, *breadth* of vocabulary knowledge (also called vocabulary size) refers to the number of words one knows, while *depth* of vocabulary knowledge is how much one knows about each word and also how this knowledge of words is structured in one's mind (Meara, 1996). *Receptive* (or passive) and *productive* (or active) vocabulary are distinguished depending on their usage in either reading/listening or writing/speaking modes (Read, 2000). In an assessment setting, receptive lexical proficiency is elicited with a task in which a test taker chooses answers from given alternatives, as in multiple-choice tests, for instance. Productive lexical proficiency, in contrast, is

[1] Yet, this result should be interpreted with caution, for she defined lexical diversity as the "total number of different words," which incorporates another factor, text length. In fact, lexical diversity was strongly correlated with text length ($r = .92$) in her study.

evaluated by examining what a test taker has said or written in response to stimuli or a task. Each construct of lexical proficiency is a combination of these four aspects. For example, the lexical diversity assessed in Engber's (1995) study is ''breadth of'' lexical proficiency, and at the same time it is ''productive'' vocabulary knowledge, while vocabulary knowledge assessed in Schoonen et al's (2002, 2003) studies is ''breadth of'' and ''receptive'' lexical proficiency.

To cover these theoretical distinctions of aspects of L2 lexical proficiency, four aspects of students' English lexical proficiency were assessed in this study – vocabulary size, depth of vocabulary knowledge, word definition ability, and lexical diversity. Previous research on the relationship between lexical proficiency and writing in L2 has largely focused on the ''breadth'' aspect of vocabulary knowledge such as vocabulary size (receptive) and lexical diversity in writing (productive). In contrast, there are still relatively few studies that have investigated the influence of the ''depth'' aspect of lexical proficiency on L2 writing in detail, while the protocol studies referred to above (cf. Cumming, 1990; Raimes, 1985) have emphasized it. The present study, therefore, assessed this aspect of lexical proficiency with two measures: depth of vocabulary knowledge (receptive) and word definition ability (productive). Depth of vocabulary knowledge was assessed receptively by eliciting the test taker's knowledge of word associations (see Method section).

A few remarks should be made concerning word definition ability, for this aspect of lexical proficiency has rarely been studied in L2 writing research. Defining words typically involves explicit knowledge of ''*genus*—the class of concepts to which [an entity] belongs, and *differentia*—features that distinguish the entity from all other members of the class'' (McKeown, 1993, p. 18). For example, *dog* is defined as ''a four-legged furry canine mammal related to the wolf and fox, commonly used for a pet'' (Wordsmyth dictionary, June 1, 2006). In this definition, *canine*, *mammal* and *a pet* are *genus* and superordinate terms of *dog*. The other words function as *differentia*, in which *wolf* and *fox* are co-hyponyms of *dog*. Thus, the knowledge of word relations that is used to define words includes formal sense relations such as hyponymous, synonymous, and antonymous. This structure of the semantic network of words in one's mind may contribute to writing fluency, as Reynolds (1995) demonstrated. Word definition also requires metalinguistic manipulation of concepts through controlled language processing. People may know superordinate words to define words, but may not utilize them because of their unfamiliarity with definition forms (e.g., Scott & Nagy, 1997; Skwarchuk & Anglin, 1997), or they may not be able to ''retrieve these words rapidly and integrate them into the information structure being conveyed'' (Snow, 1990, p. 709). These two factors in the ability to define words (structure of the semantic network of words and the ability to metalinguistically manipulate words) can be required when writing summaries, in which writers express the given content with different grammatical structures and words that are closely related to the phrases in the source text.

Controlling variables

The impact of L2 lexical indices on L2 writing was assessed after five L1 and L2 abilities were controlled. The selection of the controlling variables was made on the basis of their effects on L2 writing shown in previous research. It was assumed that by controlling for these variables, it is possible to assess the relative effect of lexical proficiency.

The first controlling variable is general L2 proficiency. While the relationship between L2 proficiency and L2 writing proficiency is controversial (e.g., Raimes, 1985), some studies have suggested its explanatory power for L2 writing. For example, Cumming (1989) investigated the effects of L2 proficiency and writing expertise itself on L2 writing proficiency of Francophone writers learning English as L2. He found that L2 proficiency had an impact on the overall quality of the students' writing, while their writing expertise mainly explained the writing processes and strategies used during writing. Similarly, Sasaki and Hirose (1996) assessed the effects of L2 proficiency, L1 writing proficiency, and metaknowledge about writing on L2 writing proficiency of Japanese EFL students. Their analyses showed that L2 proficiency made the largest unique contribution to L2 writing proficiency, but the contribution of the other two factors (L1 writing proficiency and metaknowledge about writing) was negligible.

A second controlling variable is L1 writing proficiency. In Schoonen et al.'s (2003) study, L1 writing proficiency was the strongest predictor of L2 writing proficiency of young EFL students. Similarly, in Hirose and Sasaki's (1994) study, which was a pilot study of Sasaki and Hirose (1996), the contribution of L1 writing proficiency to L2 writing proficiency was slightly stronger than that of L2 proficiency. A third controlling variable is L1 vocabulary knowledge. This variable was included on the basis of Schoonen et al. (2002, 2003) though the present study is the first to use it as an explanatory variable for L2 writing.

A fourth controlling variable is reading comprehension in L2. Writing models have come to incorporate reading as a component of the writing processes and to emphasize the interdependency of reading and writing (Grabe & Kaplan, 1996; Hayes, 1996). As summary writing is generally considered to be a composite skill of reading and writing, the role of reading is even more important to summary writing than to independent writing. This factor also controls for the difference in participants' understanding of the source texts of the summary writing tasks.

A fifth controlling variable is the length of summaries. Text length, which is often used as an index of writing fluency, is an important factor in the quality of writing in general (e.g., Grant & Ginther, 2000). Jarvis, Grant, Bikowski, and Ferris (2003) analyzed the text characteristics of highly rated timed compositions by ESL writers. They showed that highly rated compositions may display different profiles with strengths in different linguistic characteristics, but one characteristic the highly rated ones shared was that the text was long. The importance of text length seems applicable to reading-based writing. Cumming et al. (2006), who used the same summary tasks as in the present study and the other two types of writing tasks, showed that as a whole more effective L2 writers wrote longer compositions than did less effective L2 writers.

Thus, the research questions for this study were:

(1) What are the relationships between different aspects of English lexical proficiency and the ability of EFL Japanese-background university students to write summaries in English?
(2) Does English lexical proficiency as a whole account for the EFL students' summary writing performance when the effects of English reading comprehension, the length of summaries, English proficiency, knowledge of Japanese vocabulary, and Japanese writing proficiency are controlled?
(3) How important is each aspect of EFL students' lexical proficiency to their summary writing performance?

Method

Participants

Participants were 68 full-time students at a university in Japan who were native Japanese speakers learning English as a foreign language. They were aged from 18 to 25; 47 students were female (69.1%) and 21 were male (30.9%). About one-third of them were majoring in English literature, another third had not yet chosen their majors (because they did not have to until their third year), and the other third of the students were majoring in various fields such as philosophy, psychology, and sociology. All of them had at least 6 years (an average of 8.36 years) experience learning English in junior high school and high school in Japan. The participants shared similar educational and cultural backgrounds, so these intervening variables were controlled in the study. Their level of English proficiency varied, but most were in the intermediate level judging from the scores they reported for some standardized tests (for example, about a third reported their TOEIC scores, and their average was 697).

Tasks

The participants completed eight tasks: (1) writing two summaries in English, (2) a test of English vocabulary size, (3) a test of depth of English vocabulary knowledge, (4) a word definition test in English, (5) reading comprehension tests in English, (6) a self-assessment questionnaire of communicative English ability, (7) writing a composition in Japanese, and (8) a test of Japanese vocabulary knowledge.

Summary writing in English

The participants performed integrated reading-writing tasks that had been field-tested for the new TOEFL Internet-based test (TOEFL-iBT) (e.g., Cumming et al., 2006, Cumming et al., 2000). These were selected from *LanguEdge Courseware* (Educational Testing Service, 2002, and reprinted with permission). The students wrote summaries in English of two English texts: one about two political parties in the 19th century United States, and the other about the early development of cinema. The source text and the writing prompt for one summary writing task are given in Appendix A. Before they started to write, the task was explained in Japanese, including details of the topics of the reading passages, the questions being asked, and the evaluation criteria. Because it was a reading-based writing task, the students were encouraged to use their own words if possible. They were shown a model summary that had been

written on a different topic. They were allowed to use dictionaries, and most students carried an electronic dictionary, which usually includes English–Japanese and Japanese–English dictionaries. A summary writing task and a reading comprehension test for each reading passage were administered together, and the approximate time limit for these tasks was one hour (45 minutes for the summary writing task and 15 minutes for the reading comprehension test). Students could spend more time completing the tasks if necessary, but they were instructed not to spend more than one hour on each summary writing task. On average, students took 66 minutes to complete each combination of a summary writing task and a reading comprehension test. As shown in the writing prompt, they were told that their summary should be between 175 and 200 words, but most summaries were shorter than 175 words. The average number of words per summary was 140.85 (SD = 37.34); the minimum was 75, while the maximum was 268 (since each student wrote two summaries, the figures reported here are the average of both for each student).

Test of English vocabulary size

The Vocabulary Levels Test (VLT) estimates learners' receptive vocabulary size (Schmitt, Schmitt, & Clapham, 2001). It is a multiple-choice test and consists of 50 questions. In each question a test taker chooses three words from six words and matches them with the three given meanings. For example, when three meanings (*written agreement*, *way of doing something*, *reason for believing something is or is not true*) are given, one should choose *contract*, *method*, and *evidence*, respectively from six words (*area*, *contract*, *definition*, *evidence*, *method*, *role*). The test words are taken from five different vocabulary lists (different word frequency lists and the Academic Vocabulary list). On average, it took 27 minutes for participants to complete the VLT. The reliability (Cronbach's alpha) for the VLT was .77.

Test of depth of English vocabulary

The Word Associates Test (WAT) assesses the depth of a person's receptive vocabulary knowledge in English (Read, 1998). This test contains 40 test words, all of which are adjectives. Under each test word, there are eight words from which a test taker chooses four words that are associated with the test word. The four words have either paradigmatic or syntagmatic relations to the test word. For example, for a test word *beautiful*, one should choose four words (*enjoyable*, *face*, *music*, *weather*) from eight alternatives (*enjoyable*, *expensive*, *free*, *loud*, *education*, *face*, *music*, and *weather*). On average, participants took 21 minutes to complete the WAT. The internal consistency (Cronbach's alpha) for the WAT was .83.

Word definition test in English

This study used a specially devised word definition test based on Schwanenflugel, Stahl, and McFalls (1997). For the word definition test in the present study, 10 words were chosen from the two reading texts used in the reading–writing task: *cinema*, *mass-consumption*, *projection*, *reproduction*, *exhibitor*, *market*, *presidency*, *economy*, *responsibility*, and *welfare*. The participants wrote a definition *and* a sample sentence for each word in English. (Participants wrote as many different definitions and sample sentences as possible if they knew more than one meaning for a word.) The words were chosen by asking three native English-speaking graduate students specializing in Second Language Education (SLE) to identify the five most important words in each reading passage necessary to understand the text. Seven words were identified as key words by more than two students, and three were recognized as important to understanding the texts by one student and myself. The word definition test was always conducted before the summary writing and reading comprehension test, so that the participants would not guess their meanings from the texts. On average, participants took 29 minutes to complete the word definition test.

Reading comprehension tests in English

Either before or after writing the summary, each participant answered a reading comprehension test for each text. The task order of summary writing and the reading comprehension test was reversed for each text, so all the participants experienced both task orders. The two reading comprehension tests were adapted from *LanguEdge Courseware* (Educational Testing Service, 2002). Since each test consisted of 11 questions with the same multiple-choice format, the scores from the two tests (22 items in total) were combined for the data analyses. Their internal consistency (Cronbach's alpha) was .58. This was not as high as desirable, perhaps because these reading tests are only a part of the overall reading tests in the new TOEFL-iBT.

Self-assessment questionnaire of communicative English ability

The participants' English proficiency was estimated using a self-assessment questionnaire of communicative English ability (Bachman & Palmer, 1989). A self-assessment questionnaire was used in this present study because it was not feasible to have all participants undergo the same standardized English proficiency test like TOEFL, and also because Bachman and Palmer (1989) demonstrated the validity of this questionnaire. It consists of 21 questions, for each of which the students evaluated their English ability on a four-point Likert scale. The tool was translated into Japanese to ensure comprehension and practicality (Oskarsson, 1978). After analyses of the items, four items were dropped, which resulted in the high internal consistency of the self-assessment questionnaire (Cronbach's alpha = .91). Twenty-three students reported TOEIC scores, and there was a high correlation between their TOEIC and self-assessment scores ($r = .87$, $p < .01$), which supports the validity of this questionnaire as a measure of English proficiency. In addition to self-assessment, students were also asked about their biodata and English learning experience.

Test of Japanese vocabulary knowledge

To measure the students' Japanese vocabulary knowledge, Hattori's (1990) personal-computerized adaptive test of Japanese verbal ability was adapted by extracting 50 items from a pool of 346 items.[2] Hattori's verbal ability test has a multiple-choice format, in which the test taker selects one of five possible meanings that corresponds to each test word. On average, participants spent 10 minutes on the Japanese vocabulary test. The internal consistency (Cronbach's alpha) for this test was .82.

Writing a composition in Japanese

To assess the students' Japanese writing proficiency, they were asked to write an argumentative composition in Japanese on whether they agree or disagree with non-Christian Japanese celebrating Christmas. A writing prompt was taken from Sasaki (2000, p. 267) and given in Japanese. On average, participants spent 14 minutes on this task.

Procedure

The experiment was conducted in a quiet room of the university department. Participants attended two experimental sessions individually; the sessions were held on two different days with a one-week interval between. Each session took about 2–2.5 hours (i.e., each student spent a total of 4–5 hours in two sessions). Prior to each task, individual participants were given instructions in how to perform it. I sat silently in the room while the tasks were carried out. Every task had an approximate time limit, which was based on a pilot study. However, the participants were allowed to extend the time limit. They were asked not to spend too much time on any one task, in particular, the summary writing task (e.g., no more than one hour). The order of the tasks was randomized to reduce task-order effects. Ten different orders of the tasks were devised and randomly assigned to the participants.

Rating of data

Summary writing performance in English

The summaries that the students wrote were holistically scored on a five-point scale based on the reading/writing task scoring guidelines for the new TOEFL-iBT. The guidelines involve description of text characteristics at each level concerning (1) ideas included, (2) the organization, (3) language forms, and (4) the use of words from the source text (Appendix B).[3] In addition, key points of the two source texts that should be included in an effective summary are indicated in the guidelines together with sample responses at each level. In principle, the guidelines are lenient with

[2] Hattori (personal communication, January 18, 2004) recommended that 40 items would suffice for a rough estimation of Japanese students' vocabulary knowledge. Because this assessment tool was developed for a wide range of educational levels (from junior high to university and graduate levels), the 50 most difficult items were extracted. The adapted tool was then piloted, and problematic items were replaced with easier ones.

[3] LanguEdge Courseware materials are reprinted by permission of Educational Testing Service, the copyright owner. However, the test questions and any other testing information are provided in their entirety by Elsevier. No endorsement of this publication by Educational Testing Service should be inferred.

accuracy of information included in a summary. Irrelevant or incorrect information in a summary is disregarded as long as it does not disturb the intelligibility of the summary.

Two native English-speaking graduate students specializing in SLE with ESL teaching experience scored all the summaries independently. Before the actual rating, I met them twice to explain how to rate the summaries, let them rate sample summaries, and talk over any discrepancies. They practiced rating 10 summaries that were written for the pilot study, and discussed the evaluation criteria through e-mails as well. After all of us felt that the two raters had developed a similar way of rating, they individually rated all the 136 summaries. The inter-rater reliability was 0.67.[4] This inter-rater reliability was not as high as desirable, so the raters were asked to reach agreement on all of the discrepancies in their ratings and to provide one final score for each summary. These agreed-on scores were used for the data analyses. Each student wrote two summaries, so the average of the two scores represented the student's summary writing performance.

Word definition ability

The word definition test was holistically scored according to Schwanenflugel et al.'s (1997) evaluation scale. Each definition and sample sentence was evaluated on the scale of 0 to 2. A correct definition and sentence scored two points each. A ''domain-related but essentially incorrect definition'' (Schwanenflugel et al., 1997, p. 537) (e.g., *an essential function to start the computer* for PROJECTOR) and a sentence that was grammatically correct but semantically awkward or too vague (e.g., *You have to think and act more presidency* for PRESIDENCY) scored one point each. A totally incorrect definition and sentence scored zero points. If the students wrote more than one definition and sentence for a word, they could receive more than four points for the word. In the analysis, *word definition ability* refers to the combined scores for definitions and sample sentences. Both of the two subcomponents of word definition ability assess depth of productive vocabulary knowledge, but it is possible that they have different constructs. Therefore, the effect of each subcomponent was also examined: ability to write word definitions (henceforth called *definition writing ability*), and ability to write sample sentences (*sentence writing ability*).

Two native English-speaking graduate students specializing in SLE with ESL teaching experience (different from those who rated the students' summaries) independently scored 20% of the data. The inter-rater reliability was 0.97. Once this reliability was established, one of the two raters scored the rest of the data.

Lexical diversity

The traditional measure of lexical diversity, the type–token ratio (TTR), is generally recognized as unsatisfactory because of its sensitivity to the variable length of different texts. For instance, TTRs for texts of 100 and 500 words are not comparable, because the TTR tends to decrease as texts become longer. This is a serious problem when using TTR as an index of language development, for learners tend to produce longer texts as their proficiency increases. Therefore, an alternative measure, *D*, which is designed to control for text length, was used in this study. *D* is calculated from a stochastic mathematical model in which the decline of the TTR with the increase in the length of the text produces a curved line. Every text produces this TTR line, but the position of the line differs according to lexical diversity. A text with a wide lexical diversity produces a curve in a higher position. To calculate *D* values, the TTR line is drawn by plotting TTR values against different numbers of tokens in a text. One hundred samples of 16 different numbers of tokens (from 35 to 50 tokens) from each text were collected, and the average TTR for each token size was plotted. The actual TTR curve is then matched against the theoretical curves that are formed based on the mathematical model. The reliability and validity of *D* have been demonstrated in several studies (e.g., Durán, Malvern, Richards, & Chipere, 2004; Read, 2005).[5]

Text length

In this study, text length was operationalized as the total number of words in a summary, following Polio's (1997) definition for word count.

[4] All the inter-rater reliabilities in the study were calculated according to the following formula given by Hatch and Lazaraton (1991) inter-rater reliability $= \frac{2\times\text{correlation}}{1+\text{correlation}}$.

[5] Recently, however, McCarthy and Jarvis (2007) showed that even *D* is more or less influenced by text length. They proposed some ranges of text length in which indices of lexical diversity, including *D*, are safely compared.

Writing proficiency in Japanese

Sasaki and Hirose's (1999) analytic rating scale was used to score the Japanese L1 writing. For inter-rater reliability, a Japanese native-speaking graduate student with a Master's degree in Teaching Japanese as a Second Language and I rated 20% of the summaries. The inter-rater reliability was 0.83. Once this reliability was established, I rated the rest of the data.

Data analysis

The study involved one dependent variable (summary writing performance in English) and nine predictor variables: (1) vocabulary size in English, (2) depth of English vocabulary knowledge, (3) word definition ability, (4) lexical diversity, (5) reading comprehension, (6) text length, (7) English proficiency, (8) Japanese vocabulary knowledge, and (9) writing proficiency in Japanese.

Descriptive statistics were used to answer Research Question 1, while multiple regression analysis was used for Research Questions 2 and 3. Prior to regression analysis, data screening was carried out. The distribution of the dependent variable (holistic scores of the students' summaries) was positively skewed, so a square root transformation was carried out. No univariate or multivariate outliers were found. Assumptions of normality, linearity, and homoscedasticity were met. Neither multicollinearity nor singularity were signaled by SPSS; however, if any correlation between two variables is over .7, it may indicate that the two variables are redundant and weaken an analysis (Tabachnick & Fidell, 2007). In fact, a high correlation was found between vocabulary size and depth of vocabulary knowledge: $r(68) = .71, p < .001$. Therefore, depth of vocabulary knowledge was dropped from regression analysis. Vocabulary size, instead of depth of vocabulary knowledge, was retained as a predictor variable because Schoonen et al. (2002, 2003) assessed vocabulary knowledge using a similar tool to the VLT, and thus, the results of the present study would be comparable with theirs. After removing depth of vocabulary knowledge, no collinearity was observed based on collinearity diagnostics.

For Research Question 2, a multiple regression analysis examined whether the three lexical proficiency variables (vocabulary size in English, lexical diversity, and word definition ability) in combination predict the EFL students' summary writing performance by controlling for the five language abilities (reading comprehension, text length, English proficiency, Japanese vocabulary knowledge, and writing proficiency in Japanese). Another multiple regression analysis was conducted for Research Question 3 to assess the contribution of each aspect of lexical proficiency to the participants' summary writing performance after the five language abilities were controlled. In this analysis, all orders of entry of the three predictor variables were tried, and the effect of each lexical proficiency variable was compared. However, since vocabulary size and lexical diversity made little contribution to the dependent variable, the results largely remained the same regardless of order of entry. Therefore, one representative analysis is presented in this paper.

Results

In this section, the results are presented according to the three research questions guiding the study.

Research Question 1 What are the relationships between different aspects of English lexical proficiency and the ability of EFL Japanese-background university students to write summaries in English?

Table 1 displays the means and standard deviations of all the variables and the correlations of each predictor variable with the students' summary writing performance. All the predictor variables except for lexical diversity, Japanese vocabulary knowledge, and Japanese writing proficiency correlated with summary writing performance in English.

Two lexical proficiency indices – vocabulary size and word definition ability – had moderate correlations of .40 and .51, respectively, with summary writing performance. In contrast to vocabulary size and word definition ability, lexical diversity did not have a linear correlation with summary writing performance. Scatterplots of lexical diversity and summary writing performance exhibited a pyramid-like relationship, as shown in Fig. 1. The *D* values of score-1 summaries diverged widely from 29.26 to 75.19, whereas those of score-3 summaries converged between 47.02 and 57.34. It is likely that lexical diversity tended to increase as summary writing

Table 1
Means and standard deviations for all variables, and correlations of predictor variables with summary writing performance ($N = 68$).

Variable	*M*	SD	*r* with summary
1. Summary[a]	1.50	0.58	–
2. Vocabulary size[b]	104.03	16.33	.40**
3. Depth of vocabulary knowledge[c]	111.13	13.34	.34**
4. Word definition[d]	27.13	7.08	.51***
4a. Definition writing[e]	13.47	3.66	.53***
4b. Sentence writing[f]	13.66	3.74	.44***
5. Lexical diversity (*D*)	50.13	9.10	−.13
6. Reading[g]	16.13	2.83	.48***
7. Text length	140.09	38.06	.41**
8. English proficiency[h]	35.18	7.35	.28*
9. Japanese vocabulary[i]	31.62	7.19	.16
10. Japanese writing[j]	45.03	3.19	.12

Note: The mean and standard deviation for summary writing performance are the figures before transformation, and correlations are calculated after the transformation of the dependent variable (summary writing performance).
[a]Maximum score = 5; [b]Maximum score = 150; [c]Maximum score = 160; [d]There was in principle no maximum score, but the highest score was 40; [e]The highest score was 22; [f]The highest score was 20; [g]Maximum score = 22; [h]Maximum score = 68; [i]Maximum score = 50; [j]Maximum score = 60; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

performance improved, but also that some summaries showed high lexical diversity if the writers relied heavily on the source text.

Research Question 2 Does English lexical proficiency as a whole account for the EFL students' summary writing performance when the effects of English reading comprehension, the length of summaries, English proficiency, knowledge of Japanese vocabulary, and Japanese writing proficiency are controlled?

The results of a multiple regression analysis indicated that the five language abilities significantly accounted for the EFL students' summary writing performance, $R^2 = .40$, $F(5, 62) = 8.08$, $p < .001$ (adjusted $R^2 = .35$). However, the three lexical proficiency variables did not predict their summary writing performance over and above the five language abilities, R^2 change = .05, $F(3, 59) = 1.69$, $p > .05$. These results suggest that the effects of lexical proficiency were relatively weak when they were compared with basic language abilities. In particular, reading comprehension and text length were the two strongest predictors, uniquely accounting for 5% and 8% (squared semipartial correlation) of the variance in summary writing performance respectively.

Research Question 3 How important is each aspect of EFL students' lexical proficiency to their summary writing performance?

Next, after controlling for the five language abilities, the unique effect of each aspect of L2 lexical proficiency on summary writing performance was examined with hierarchical regression analysis (Table 2). Lexical diversity was entered in Step 2, vocabulary size in Step 3, and word definition ability in Step 4. Two additional hierarchical regression analyses were conducted for the two components of word definition ability by substituting the variable in Step 4 (this is why there are three Step 4s in Table 2).

The point to observe in Table 2 is the change in R^2 in the left column. In Step 2, vocabulary size did not predict summary writing performance over and above the five controlling language measures (R^2 change = .01, $F(1, 61) = 1.03$, $p = .31$). This result was probably due to the significant correlations of vocabulary size with reading comprehension ($r = .51, p < .001$) and with English proficiency ($r = .31, p < .05$). Although vocabulary size itself had a moderately high correlation with summary writing performance, its effect was largely accounted for by the controlling variables, especially reading comprehension. Similarly, lexical diversity did not contribute significantly to summary writing performance over and above the variables entered in Steps 1 and 2 (R^2 change = .00, $F(1, 60) = 0.24$, $p = .63$). Unlike vocabulary size, this result was expected because of its nonlinear relationship with summary writing performance, as shown above.

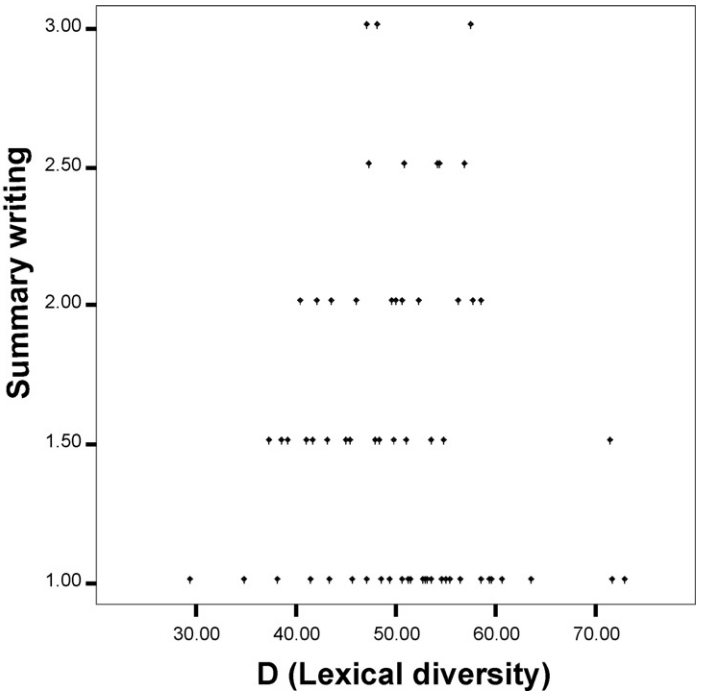

Fig. 1. Scatterplots between *D* (lexical diversity) and summary writing performance.

In contrast, word definition ability accounted for 4% of the variance in summary writing performance ($F(1, 59) = 3.75$, $p = .06$), which was close to a significance level of $p = .05$. More noteworthy is the difference between the two components of word definition ability. Definition writing ability explained 5% of the summary writing variance over and above all the other variables (R^2 change = .05, $F(1, 59) = 5.32$, $p < .05$). However, sentence writing ability did not predict summary writing performance after the other variables were controlled (R^2 change = .02, $F(1, 59) = 1.72$, $p = .20$). As with vocabulary size, sentence writing ability had a moderately high correlation with reading comprehension ($r = .49$, $p < .001$), and this is a plausible reason for the nonsignificant contribution of this variable.

Discussion

The results of this study demonstrated that different aspects of lexical proficiency associate differently with L2 learners' summary writing performance. It is unreasonable to assess the impact of one particular aspect of lexical

Table 2
Hierarchical regression analysis predicting summary writing performance with basic language abilities and L2 lexical indices ($N = 68$).

Step	Predictor variable	β	R^2	ΔR^2
1	Reading	.42**	.40**	.40**
	Text length	.34**		
	English proficiency	.18		
	Japanese vocabulary	.03		
	Japanese writing	.01		
2	Vocabulary size	.12	.41**	.01
3	Lexical diversity (*D*)	−.05	.41**	.00
4	Word definition	.26†	.44**	.04†
4	Definition writing	.29*	.46**	.05*
4	Sentence writing	.18	.42**	.02

†$p < .10$; *$p < .05$; **$p < .01$.

proficiency on writing and generalize the result to lexical proficiency in general. L2 vocabulary size, for instance, did not make a unique contribution to writing in L2 in the study by Schoonen et al. (2003), and the same result was obtained in the present study regardless of the difference in writing tasks. However, this study found that another aspect of lexical proficiency did significantly predict students' summary writing performance. Thus, it is possible to interpret the data as suggesting that a mere increase in vocabulary size will not immediately lead to better writing performance in L2, but other aspects of L2 proficiency may associate with L2 writing in a more intricate way. The distinction between the two receptive lexical indices (size and depth of L2 vocabulary knowledge) was not obvious (at least between the constructs elicited by the two vocabulary tests used in the present study), and their effect on the students' summary writing seemed to be dwarfed by the other predictors. In contrast, the two indices of productive L2 lexical proficiency (*D* and word definition ability) showed unique relationships with summary writing performance, which was not completely explained by other L2 abilities like English proficiency, reading and writing fluency.[6]

The study also showed that the effect of lexical proficiency on the summary writing task was not pronounced compared to the L2 abilities (English proficiency, reading comprehension, and writing fluency). This result seems to support the general findings from Schoonen et al. (2003) in which the effect of L2 vocabulary knowledge on L2 writing proficiency was smaller than that of other language abilities such as metacognitive knowledge about writing and L1 writing proficiency. In the present study, L2 proficiency indices correlated with L2 language skills and all these variables largely overlapped in the prediction of summary writing performance. L2 lexical proficiency (particularly vocabulary size and depth of vocabulary knowledge) is therefore likely to provide a basis for writing and reading, but may not in itself independently explain L2 summary writing. This makes sense for such complicated tasks as summary writing, which entails a wider variety of cognitive skills than simply writing or reading (Sarig, 1993). The effects of lexical proficiency may have declined in saliency due to the composite nature of summary writing performance.

The most notable finding of the present study was that the ability to define words made a unique contribution to the students' summary writing performance even after the effects of reading comprehension and writing fluency were controlled for. The ability to write sample sentences, the other component of word definition ability, did not uniquely predict summary writing performance, and this difference between the effects of the two components merits closer examination. In theory, writing definitions and sample sentences might have tapped into different abilities. As indicated in the introductory section of this paper, the ability to define words calls for two factors (i.e., structure of semantic network of words, and the ability to metalinguistically manipulate words). In contrast, writing a sample sentence using a word may be possible with implicit and contextual knowledge concerning the word with relatively less controlled processing. These two factors appear to be crucial in explaining why the ability to write definitions, but not the ability to write sample sentences, uniquely predicted the students' summary writing performance.

The students in this study tended to use more words with paradigmatic relations in writing definitions than in writing sample sentences. Some such examples were synonyms of the test items, e.g., *movie* for *cinema*, *show* or *display* for *exhibit*, *shop* or *mall* for *market*. The participants also tried to give superordinate words for test words (though not always successfully), e.g., *leadership* for *presidency*, and *machine* for *projector*. More typical definitions were extended explanations of the test words that involved words in a class such as synonyms, hyponyms (or hypernyms), antonyms (e.g., *produce* and *consume*) and co-hyponyms (e.g., *the old*, *the poor*, *the weak*). In contrast to definitions, most sample sentences contained words that had associative or syntagmatic word relations with the test words (e.g., *watch* and *cinema*, *serious* and *problem*).[7]

Words with paradigmatic relations were found in relatively effective summaries. For example, in the Sample Summary (Appendix C), the first sentence of the first paragraph states, ''Democrats and Whigs held different views on the economy.'' This sentence is a brief paraphrase of a sentence, ''Whigs and Democrats held different attitudes toward the changes brought about by the market, banks, and commerce.'' (ll. 8–9) in the source text (Appendix A). In the paraphrased sentence, *views* is used as a synonym for *attitudes*. Likewise, the first sentence of the third paragraph in the Sample Summary, ''Democrats and Whigs also differed in their perspectives on the role of government,'' comes from a

[6] One possible reason why word definition ability was more significant to summary writing performance than were other aspects of lexical proficiency is that the 10 test items in the word definition test were taken directly from the reading texts, while those of the Vocabulary Levels Test and Word Associates Test were not. However, this explanation is somewhat doubtful because the two components of the word definition test did not make the same contribution to the students' summary writing performance even when the test items were identical.

[7] For the distinction between paradigmatic and syntagmatic word associations in the development of depth of receptive vocabulary knowledge tests, see for example Greidanus and Nienhuis (2001).

source sentence, ''Whigs and Democrats differed not only in their attitudes toward the market but also about how active the central government should be in people's lives'' (ll. 23–24). In this paraphrased sentence, *perspectives* serves as a synonym of *attitudes*. Within the summary, *views* and *perspectives* are also synonyms, and this may contribute to coherence in the summary.

This observation supports Hoey's (1991) suggestion that specific word relations (e.g., complex lexical repetition and simple paraphrase) are important to the enhancement of coherence in text, and also Reynolds' (1995) observation that topic development in writing hinges on words that are semantically related to key concepts in essays. At the same time, a well-structured semantic network of words (especially knowledge of paradigmatically related words) may help L2 writers successfully paraphrase or generalize source texts in summary writing as Keck (2006) pointed out. The significance of network building in the L2 mental lexicon has been emphasized in L2 vocabulary research (Haastrup & Henriksen, 2000). However, the results of this study imply that the words in such a network are not evenly interconnected, and what is germane to summary writing is how these words are *clustered* and *stratified* in the network.

With respect to the ability to metalinguistically manipulate words, writing definitions seem to require more controlled and complicated cognitive endeavor than writing sample sentences. Generally speaking, the students used more complex syntactic structures (e.g., relative clauses) and constructed more sophisticated semantic relationships in writing definitions than in writing sample sentences. Here is an example:

responsibility
Definition: The idea or fact that we must be blamed of our action and its consequence.
Sample sentence: The responsibility of that accident is discussed now.

Regardless of whether their answers were correct or not, the students gave more detailed and even philosophical descriptions in more complex forms when writing definitions than when writing sample sentences. If they had been instructed to write definitions in Japanese, they would have produced Japanese translations of the English test words. However, defining words in the same language is by far the more challenging task and much more so in an L2 (e.g., Fukkink, Block, & de Glopper, 2001 for L1 children, and Ordoñez, Carlo, Snow, & McLaughlin, 2002 for L2 children). In contrast, the sample sentences were usually short and syntactically simple.

Writing summaries might also have called on this ability to metalinguistically manipulate words. The grammatical structures of source sentences were often and more or less successfully altered in more effective summaries. For instance, a source sentence ''Nor did the Whigs envision any conflict in society between farmers and workers on the one hand and businesspeople and bankers on the other.'' (ll. 18–20) was paraphrased as ''So they [Whigs] did not believe the conflict between the poor and the rich.'' (l. 6) in the Sample Summary. In another more effective summary, a source sentence ''Religion and politics they [Democrates] believed, should be kept clearly separate,. . .'' (ll. 28–29) was rephrased as ''Democrats (. . .) wanted the separation of religion and politics.'' In these two examples the writers gave paraphrases by replacing source words with synonym-like words (e.g., *envision* with *believe*, *believed* with *wanted*) and changing grammatical structures of the source sentences.

The observation above corroborates the findings of Cumming's (1990) protocol-analysis study. Cumming found that learners with a higher level of writing expertise did more metalinguistic thinking about words than less-skilled writers, by searching for ''right'' words and phrases and semantically and pragmatically evaluating them against alternative words in L2 and cross-linguistic equivalents in L1. It could be concluded that L2 writing may be influenced not only by explicit word knowledge but also by awareness of word usage and the ability to productively manipulate the knowledge (including syntactic operation).

The findings of this study should not be easily generalized to other types of writing. If the students in this study had performed independent writing tasks, we might have found a stronger impact of lexical proficiency, or different aspects of lexical proficiency (such as vocabulary size and *D*) might have made a unique contribution to their writing. In a sense, summary writing is a special type of writing in which writers inevitably use words from a source text. Of course, heavy dependence on the source text is problematic, as shown in the relationship between lexical diversity and the quality of summary writing that was found in this study. It is beyond the scope of this paper to analyze how the L2 writers' lexical proficiency affects their summary writing strategies (including how they copy verbatim from the text), but this issue may be further investigated in different studies (see Baba, 2007).

Conclusion and suggestions for future research

This study has shown that the ability to define words made a unique contribution to the students' summary writing performance even if they drew on the source text, that is, when their knowledge of words as well as background knowledge were controlled for to some extent. Moreover, the effect of this ability on summary writing was not fully explained by those of reading comprehension and writing fluency. Thus, the present study has empirically shown that the well-structured semantic network of words and the ability to productively use this network as well as the L2 writer's metalinguistic knowledge constitute the construct of summary writing. The study is thus pioneer in investigating the construct of summary writing with componential analyses, although further research is necessary to explore not only how the writers summarize texts in L2, but also what language abilities and knowledge they need to write a summary. This line of investigation will throw light on the mechanism of summary writing from an additional angle.

The study was not without limitations. For example, the summary writing task seemed to be rather challenging for participants, which resulted in the positive skew of its score distribution. Easier reading texts might be more suitable for unskilled L2 writers. In addition, the reliability coefficients of some tasks were rather moderate (i.e., overall quality of summaries, reading comprehension test), and future studies should address these problems. Furthermore, sample size was relatively small for the number of predictor variables, which might have weakened statistical power and resulted in failure to find significant effects.

With these limitations in mind, there are some possible avenues for future research. First, in order to increase the external validity of the findings from this study, a wider variety of writing tasks in various academic writing contexts need to be employed. For example, there are various ways even to summarize texts. Factors such as referencing or not referencing source texts (Stein & Kirby, 1992), or the structure and difficulty of the source texts will no doubt make a difference to learners' performances. It may be also useful to include independent writing tasks (e.g., traditional essays with a short writing prompt) and to compare them with reading-based writing tasks.

Secondly, the contribution of L2 lexical proficiency is relative to other factors included in a regression model. Therefore, if different variables had been controlled, this study might have produced different results. As far as the participants in this study are concerned, knowledge of vocabulary and writing expertise in L1 were not good predictors of their summary writing performance in L2. These L1 variables can be replaced in future research with other variables such as grammatical competence and metacognitive knowledge about summary writing when we involve participants with similar L2 proficiency levels and educational backgrounds.

It is not my intention to claim in this paper that teaching how to write definitions in L2 will directly improve L2 writers' summary writing performance. Since the data of this study are correlational, such a causal interpretation is not warranted. However, on the basis of the results from this study, it is possible to infer that L2 learners may develop their summary writing skills by enhancing the ability to write definitions in L2 or by paying attention to their own semantic structure of words and metalinguistic knowledge and applying them to writing summaries. Longitudinal classroom studies are in order to investigate these possibilities.

Acknowledgements

I thank Alister Cumming, Kumiko Inutsuka, Philip M. McCarthy, Wataru Suzuki, and *Journal of Second Language Writing*'s anonymous reviewers for their insightful comments on earlier drafts of this article. This article is based on my doctoral research with slightly different analyses.

Appendix A. Sample source text and writing prompt for the summary writing task

Directions: You have 45 minutes to answer the question below by writing a response based on information from the passage. Typically an effective response will be between 175 and 250 words.

Question: Explain the different viewpoints held by nineteenth-century Whigs and Democrats on the economy and the role of government and explain why different people supported each of the two parties.

The development of the modern presidency in the United States began with Andrew Jackson who swept to power in 1829 at the head of the Democratic Party and served until 1837. During his administration he immeasurably enlarged

the power of the presidency. ''The President is the direct representative of the American people,'' he lectured the Senate when it opposed him. ''He was elected by the people, and is responsible to them.'' With this declaration, Jackson redefined the character of the presidential office and its relationship to the people.

During Jackson's second term, his opponents had gradually come together to form the Whig Party. Whigs and Democrats held different attitudes toward the changes brought about by the market, banks, and commerce. The Democrats tended to view society as a continuing conflict between ''the people''—farmers, planters, and workers—and a set of greedy aristocrats. This ''paper money aristocracy'' of bankers and investors manipulated the banking system for their own profit, Democrats claimed, and sapped the nation's virtue by encouraging speculation and the desire for sudden, unearned wealth. The Democrats wanted the rewards of the market without sacrificing the features of a simple agrarian republic. They wanted the wealth that the market offered without the competitive, changing society; the complex dealing; the dominance of urban centers; and the loss of independence that came with it.

Whigs, on the other hand, were more comfortable with the market. For them, commerce and economic development were agents of civilization. Nor did the Whigs envision any conflict in society between farmers and workers on the one hand and businesspeople and bankers on the other. Economic growth would benefit everyone by raising national income and expanding opportunity. The government's responsibility was to provide a well-regulated economy that guaranteed opportunity for citizens of ability.

Whigs and Democrats differed not only in their attitudes toward the market but also about how active the central government should be in people's lives. Despite Andrew Jackson's inclination to be a strong President, Democrats as a rule believed in limited government. Government's role in the economy was to promote competition by destroying monopolies and special privileges. In keeping with this philosophy of limited government, Democrats also rejected the idea that moral beliefs were the proper sphere of government action. Religion and politics, they believed, should be kept clearly separate, and they generally opposed humanitarian legislation.

The Whigs, in contrast, viewed government power positively. They believed that it should be used to protect individual rights and public liberty, and that it had a special role where individual effort was ineffective. By regulating the economy and competition, the government could ensure equal opportunity. Indeed, for Whigs the concept of government promoting the general welfare went beyond the economy. In particular, Whigs in the northern sections of the United States also believed that government power should be used to foster the moral welfare of the country. They were much more likely to favor social-reform legislation and aid to education.

In some ways the social makeup of the two parties was similar. To be competitive in winning votes, Whigs and Democrats both had to have significant support among farmers, the largest group in society, and workers. Neither party could win an election by appealing exclusively to the rich or the poor. The Whigs, however, enjoyed disproportionate strength among the business and commercial classes. Whigs appealed to planters who needed credit to finance their cotton and rice trade in the world market, to farmers who were eager to sell their surpluses, and to workers who wished to improve themselves. Democrats attracted farmers isolated from the market or uncomfortable with it, workers alienated from the emerging industrial system, and rising entrepreneurs who wanted to break monopolies and open the economy to newcomers like themselves. The Whigs were strongest in the towns, cities, and those rural areas that were fully integrated into the market economy, whereas Democrats dominated areas of semisubsistence farming that were more isolated and languishing economically.

Glossary: Monopolies: companies or individuals that exclusively own or control commercial enterprises with no competitors

Appendix B. Scoring guidelines for the summary writing task

5 A response at this level has all of the following qualities:
- principal ideas presented accurately with ample and accurately connected key supporting points/elaboration as required to fulfill the task effectively
- organization effective in response to the task
- sentence formation and word forms accurate and appropriate; response may have occasional minor grammatical or lexical errors
- appropriate use of own language and language from source text

4 A response at this level has all of the following qualities:

- principal ideas presented accurately as required by the task, though one or two key supporting points/details/elaboration may be omitted, misrepresented, or somewhat unclear, inexplicit, or inexplicitly connected
- organization generally effective in response to the task
- sentence formation and word choice generally accurate and appropriate; response may have noticeable minor errors and some imprecision and/or unidiomatic language use and/or imprecise connections among ideas; however, these do not obscure meaning
- generally appropriate use of own language and language from the source text

3 A response at this level is marked by inconsistency:
- principal ideas inconsistently presented: some are discussed accurately with key supporting points/elaboration; other support/elaboration may be absent, incorrect or unclear/obscured by weaknesses in language; *or*
- inconsistent facility in sentence formation and word choice present (meaning may be unclear and may be occasionally obscured); *or*
- efforts at paraphrasing may result in a number of sentence and word form errors, but meaning is not usually obscured, or there are efforts at paraphrasing, but they do not move sufficiently away from exact wordings and/or structures in the source text; *or*
- inconsistent facility in expressing connections between and among ideas (connections exist but are not effective)

2 A response at this level is marked by flaws in presentation of information or language:
- significantly incomplete, or unclear presentation of principal ideas and key supporting points; *or*
- consistent lack of facility in sentence formation, word choice, word forms and/or connection between and among ideas; *or*
- efforts at paraphrase usually unsuccessful or very limited attempts at paraphrase

1 A response at this level exhibits one or more major flaws:
- little or no comprehensible presentation of principal ideas and key supporting points required by the task
- failure to connect points to the required task
- pervasive language errors that make it difficult for the reader to derive meaning
- text too brief or too borrowed to allow for judgment of writing proficiency

Appendix C. Sample summary with a rating of 3

Democrats and Whigs held different views on the economy. Democrats' viewpoint stood on the side of ''the people''—farmers and workers, not privileged aristocrats. They thought there was always a conflict between those two kinds of people. They wanted to bring wealth to ''the people'' and maintain a simple agrarian republic.

Whigs' one was that commerce and economic growth would benefit everyone, farmers and aristocrats. So they did not believe the conflict between the poor and the rich.

Democrats and Whigs also differed in their perspectives on the role of government. In their view, government's role was a limited one. In the economic sphere, it should promote competition by giving all the people equal opportunities. But they did not believe that moral beliefs were within the government's sphere.

On the other hand, Whigs believed that government should foster the moral welfare of the country. So their concept of government was not limited in the economic sphere.

Why different people supported each of the two parties is because each party had different colour which we have seen. Democrats appealed to those who felt alienated from the market. Democrats' will to remove a barrier from the privileged people was compatible with the poor people's situation. Whigs especially appealed to those who had power in the market. Those people were comfortable with Whigs' idea that economic growth was for everyone's benefit, with little heed to the poor's predicament.

References

Baba, K. (2007). *Dimensions of lexical proficiency in writing summaries for an English as a foreign language test.* Unpublished doctoral dissertation, Ontario Institute for Studies in Education of the University of Toronto.

Bachman, L. F., & Palmer, A. S. (1989). The construct validation of self-ratings of communicative language ability. *Language Testing, 6*, 14–29.

Carson, J. G. (2001). A task analysis of reading and writing in academic contexts. In D. Belcher & A. Hirvela (Eds.), *Linking literacies: Perspectives on L2 reading-writing connections* (pp. 48–83). Ann Arbor: University of Michigan Press.

Chapelle, C. A. (1994). Are C-tests valid measures for L2 vocabulary research? *Second Language Research, 10*, 157–187.
Cumming, A. (1989). Writing expertise and second language proficiency. *Language Learning, 39*, 81–141.
Cumming, A. (1990). Metalinguistic and ideational thinking in second language composing. *Written Communication, 7*, 482–511.
Cumming, A., Kantor, R., Baba, K., Erdosy, U., Eouanzoui, K., & James, M. (2006). *Analysis of discourse features and verification of scoring levels for independent and integrated prototype written tasks for Next Generation TOEFL* (TOEFL Monograph 30). Princeton, NJ: Educational Testing Service.
Cumming, A., Kantor, R., Powers, D., Santos, T., & Taylor, C. (2000). *TOEFL 2000 writing framework: A working paper* (TOEFL monograph 18). Princeton, NJ: Educational Testing Service.
Cumming, A., Rebuffot, J., & Ledwell, M. (1989). Reading and summarizing challenging texts in first and second languages. *Reading and Writing: An Interdisciplinary Journal, 2*, 201–219.
Durán, P., Malvern, D., Richards, B., & Chipere, N. (2004). Developmental trends in lexical diversity. *Applied Linguistics, 25*, 220–242.
Educational Testing Service. (2002). *LanguEdge courseware: Handbook for scoring speaking and writing*. Princeton, NJ: Educational Testing Service.
Engber, C. A. (1995). The relationship of lexical proficiency to the quality of ESL compositions. *Journal of Second Language Writing, 4*, 139–155.
Fukkink, R. G., Blok, H., & de Glopper, K. (2001). Deriving word meaning from written context: A multicomponential skill. *Language Learning, 51*, 477–496.
Grabe, W., & Kaplan, R. B. (1996). *Theory and practice of writing: An applied linguistic perspective*. Harlow, England: Longman.
Grant, L., & Ginther, A. (2000). Using computer-tagged linguistic features to describe L2 writing differences. *Journal of Second Language Writing, 9*, 123–145.
Greidanus, T., & Nienhuis, L. (2001). Testing the quality of word knowledge in a second language by means of word associations: Types of distractors and types of associations. *The Modern Language Journal, 85*, 567–577.
Haastrup, K., & Henriksen, B. (2000). Vocabulary acquisition: Acquiring depth of knowledge through network building. *International Journal of Applied Linguistics, 10*, 221–240.
Harley, B., & King, M. L. (1989). Verb lexis in the written compositions of young L2 learners. *Studies in Second Language Acquisition, 11*, 415–439.
Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. New York: Newbury House.
Hattori, T. (1990). Kojinsa ni oujita tesuto houshiki niyoru goirikairyoku no sokutei [An individual-difference-sensitive method to assess vocabulary comprehension ability]. *Kyoiku Shinrigaku Kenkyu, 38*, 445–454.
Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1–27). Mahwah, NJ: Laurence Erlbaum.
Hirose, K., & Sasaki, M. (1994). Explanatory variables for Japanese students' expository writing in English: An exploratory study. *Journal of Second Language Writing, 3*, 203–229.
Hoey, M. (1991). *Patterns of lexis in text*. Oxford: Oxford University Press.
Horowitz, D. M. (1986). What professors actually require: Academic tasks for the ESL classroom. *TESOL Quarterly, 20*, 445–462.
Jarvis, S., Grant, L., Bikowski, D., & Ferris, D. (2003). Exploring multiple profiles of highly rated learner compositions. *Journal of Second Language Writing, 12*, 377–403.
Johns, A. M. (1993). Reading and writing tasks in English for academic purposes classes: Products, processes and resources. In J. G. Carson & I. Leki (Eds.), *Reading in the composition classroom: Second language perspectives* (pp. 274–289). Boston: Heinle & Heinle.
Johns, A. M., & Mayes, P. (1990). An analysis of summary protocols of university ESL students. *Applied Linguistics, 11*, 253–271.
Keck, C. (2006). The use of paraphrase in summary writing: A comparison of L1 and L2 writers. *Journal of Second Language Writing, 15*, 261–268.
Kim, S. A. (2001). Characteristics of EFL readers' summary writing: A study with Korean university students. *Foreign Language Annals, 34*, 569–580.
Koda, K. (1993). Task-induced variability in FL composition: Language-specific perspectives. *Foreign Language Annals, 26*, 332–346.
Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics, 16*, 307–322.
Leki, I., & Carson, J. (1994). Students' perceptions of EAP writing instruction and writing needs across the discipline. *TESOL Quarterly, 28*, 81–101.
Linnarud, M. (1986). *Lexis in composition: A performance analysis of Swedish learners' written English*. Malmo, Sweden: Liber Forlag Malmo.
Manchón, R. M., Murphy, L., & Roca de Larios, J. (2007). Lexical retrieval processes and strategies in second language writing: A synthesis of empirical research. *International Journal of English Studies, 7*, 147–172.
McCarthy, P. M., & Jarvis, S. (2007). *vocd*: A theoretical and empirical evaluation. *Language Testing, 24*, 459–488.
McKeown, M. G. (1993). Creating effective definitions for young word learners. *Reading Research Quarterly, 28*, 17–31.
Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjaer, & J. Williams (Eds.), *Performance and competence in second language acquisition* (pp. 35–53). Cambridge: Cambridge University Press.
Ordoñez, C. L., Carlo, M. S., Snow, C. E., & McLaughlin, B. (2002). Depth of breadth and vocabulary in two languages: Which vocabulary skills transfer? *Journal of Educational Psychology, 94*, 719–728.
Oskarsson, M. (1978). *Approaches to self-assessment in foreign language learning*. Strasbourg, France: Council of Europe.
Polio, C. G. (1997). Measures of linguistic accuracy in second language writing research. *Language Learning, 47*, 101–143.
Raimes, A. (1985). What unskilled ESL students do as they write: A classroom study of composing. *TESOL Quarterly, 19*, 229–258.
Read, J. (1998). Validating a test to measure depth of vocabulary knowledge. In A. J. Kunnan (Ed.), *Validation in language assessment* (pp. 41–60). Mahwah, NJ: Lawrence Erlbaum.
Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
Read, J. (2005). Applying lexical statistics to the IELTS speaking test. *Research Notes, 20*, 12–16.
Reynolds, D. W. (1995). Repetition in non-native speaker writing. *Studies in Second Language Acquisition, 17*, 185–209.

Sarig, G. (1993). Composing a study-summary: A reading/writing encounter. In J. G. Carson & I. Leki (Eds.), *Reading in the composition classroom: Second language perspectives* (pp. 161–182). Boston: Heinle & Heinle.

Sasaki, M. (2000). Toward an empirical model of EFL writing processes: An exploratory study. *Journal of Second Language Writing, 9*, 259–291.

Sasaki, M., & Hirose, K. (1996). Explanatory variables for EFL students' expository writing. *Language Learning, 46*, 137–174.

Sasaki, M., & Hirose, K. (1999). Development of an analytic rating scale for Japanese L1 writing. *Language Testing, 16*, 457–478.

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing, 18*, 55–88.

Schneider, M., & Connor, U. (1990). Analyzing topical structure in ESL essays: Not all topics are equal. *Studies in Second Language Acquisition, 12*, 411–427.

Schoonen, R., van Gelderen, A., de Glopper, K., Hulstijn, J., Simis, A., Snellings, P., et al. (2003). First language and second language writing: The role of linguistic knowledge, speed of processing, and metacognitive knowledge. *Language Learning, 53*, 165–202.

Schoonen, R., van Gelderen, A., de Glopper, K., Hulstijn, J., Snellings, P., Simis, A., et al. (2002). Linguistic knowledge, metacognitive knowledge and retrieval speed in L1, L2, and EFL writing. In S. Ransdell M.-L. Barbier (Eds.), *New directions for research in L2 writing* (pp. 101–122). Dordrecht, Netherlands: Kluwer Academic.

Schwanenflugel, P. J., Stahl, S. A., & McFalls, E. L. (1997). Partial word knowledge and vocabulary growth during reading comprehension. *Journal of Literacy Research, 29*, 531–553.

Scott, J. A., & Nagy, W. E. (1997). Understanding the definitions of unfamiliar verbs. *Reading Research Quarterly, 32*, 184–200.

Skwarchuk, S. L., & Anglin, J. M. (1997). Expression of superordinates in children's word definitions. *Journal of Educational Psychology, 89*, 298–308.

Snow, C. E. (1990). The development of definitional skill. *Journal of Child Language, 17*, 697–710.

Stein, B. L., & Kirby, J. R. (1992). The effects of text absent and text present conditions on summarization and recall of text. *Journal of Reading Behavior, 24*, 217–232.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson.

Uzawa, K. (1996). Second language learners' process of L1 writing, L2 writing, and translation from L1 into L2. *Journal of Second Language Writing, 5*, 271–294.

Watanabe, Y. (2001). *Read-to-write tasks for the assessment of second language academic writing skills: Investigating text features and rater reactions*. Unpublished doctoral dissertation, The University of Hawai'i, Hawai'i.

Yang, L., & Shi, L. (2003). Exploring six MBA students' summary writing by introspection. *Journal of English for Academic Purposes, 2*, 165–192.

Kyoko Baba is assistant professor in the Department of English at Kinjo Gakuin University. Her current research interests are learning and assessment of vocabulary and academic writing in a second and foreign language.