Administration Volume 22 Number 1 (2015) ISSN 2187-378X

Second Language Vocabulary Research: 2009

Yuji Tanaka

This is a review article on second language vocabulary research. Articles published in leading international research journals in 2009 are the scope of this investigation. The present review is composed of the following key themes: the relationship between vocabulary knowledge and listening comprehension; English L1 and L2 speakers' knowledge of lexical bundles; measuring second language lexical growth using hypernymic relationships; the effects of vocabulary test announcement, word relevance, and vocabulary task on second language vocabulary learning; the effects of learning method and cognate status on second language vocabulary learning by children; the effect of keeping vocabulary notebooks on second language vocabulary learning; and the effects of pedagogic modifications on second language learners' reading comprehension and vocabulary recognition.

The Relationship Between Vocabulary Knowledge and Listening Comprehension

Stæhr (2009) investigated the relationship between vocabulary knowledge and listening comprehension.

After reviewing previous studies on "vocabulary knowledge" (pp. 578-580), "listening comprehension" (pp. 580-582), and "the role of vocabulary knowledge in reading and listening" (pp. 582-585), Stæhr (2009) presented the following three research questions (p. 585): (a) To what extent are vocabulary size and depth of vocabulary knowledge associated with listening comprehension?; (b) To what extent does depth of vocabulary

knowledge, in addition to vocabulary size, contribute to successful listening comprehension?; and (c) How[1] much vocabulary is needed for adequate listening comprehension?

This study comprised 115 Danish learners of English as a foreign language; there were 90 females and 25 males. They were first-year students of English at the Copenhagen Business School. They had similar educational backgrounds, and had received at least 8 years of formal English instruction in primary and secondary schools. At the start of their English studies at the Copenhagen Business School, these participants were expected to be at least at B2 (i.e., upper intermediate) level of the Common European Framework of Reference for Languages (CEFR) (pp. 585-586).

As for testing instruments, the participants were given three paper-and-pencil tests: (a) a listening comprehension test, (b) a vocabulary size test, and (c) a depth of vocabulary knowledge test (pp. 586-589).

In what follows, let me give a concise description of the above (a), (b), and (c). First, in respect of the above (a), in order to assess the participants' listening comprehension abilities, a standardized listening test from the Cambridge Certificate of Proficiency in English (CPE) was employed. The CPE listening test is designed for the C2 (i.e., very advanced) level of the CEFR and consists of different target-language situations that address a wide range of skills involved in listening comprehension at the most advanced level. The test, which takes approximately 40 minutes to complete, contains 28 questions and consists of four parts that cover a variety of different text types (p. 586).

Second, regarding the aforementioned (b), in order to assess the learners' vocabulary size, the researcher used the Vocabulary Levels Test (version 2; developed by Schmitt, Schmitt, & Clapham, 2001). This test is composed of five separate sections, which represent four levels of word frequency and one level of academic words. In other words, this test measures learners' receptive knowledge of word meaning at the following four distinct vocabulary frequency levels and one level of academic words: the 2,000 level, the 3,000 level, the 5,000 level, and the 10,000 level, as well as academic words from the Academic Word List (Coxhead, 2000). Each vocabulary level contains 60 words and 30 definitions presented in groups of 6 words and 3 definitions. For each of these groups, test takers are required to match the words to the definitions. Results are said to form an approximate estimate of how many words a learner knows (p. 587).[2]

Third, with respect to the aforementioned (c), Stæhr (2009) developed a depth of vocabulary knowledge test based on the Word Associates Test (Read, 1993, 1998). The Word Associates Test format proposed by Read (1993) has been adopted, for example, by Greidanus and Nienhuis (2001), and Stæhr's (2009) depth of vocabulary knowledge test is also based on Read's (1993) version of the Word Associates Test (pp. 587-589).

Next, Stæhr (2009) showed the means, standard deviations, and Cronbach's alpha

reliability coefficients of the three tests (i.e., the listening comprehension test, the vocabulary size test, and the depth of vocabulary knowledge test) (pp. 589-590).

As mentioned above, in this study, the three research questions (i.e., [a], [b], and [c]) were addressed, and the researcher dealt with them in order of mention. First, Stæhr (2009) examined the following research question: (a) To what extent are vocabulary size and depth of vocabulary knowledge associated with listening comprehension? With a view to clarifying the relationship between the two dimensions of vocabulary knowledge (i.e., vocabulary size and depth of vocabulary knowledge) and listening comprehension, Pearson product-moment correlation coefficients were computed. The results indicated that both vocabulary size and depth of vocabulary knowledge were significantly correlated with listening comprehension. Specifically, the correlation between vocabulary size and listening comprehension is $r = .70$, $p < .01$ (two-tailed); that between depth of vocabulary knowledge and listening comprehension is $r = .65$, $p < .01$ (two-tailed). In sum, it was shown that both vocabulary size and depth of vocabulary knowledge had strong associations with listening comprehension. Additionally, a multiple regression analysis was also performed to determine the extent to which the independent variables (i.e., vocabulary size and depth of vocabulary knowledge) could explain the variance of the dependent variable (i.e., listening comprehension). The findings showed that vocabulary size and depth of vocabulary knowledge together accounted for 51% of the variance of the dependent variable (i.e., listening comprehension). Overall, it was empirically demonstrated that there was a strong relationship between the two dimensions of vocabulary knowledge (i.e., vocabulary size and depth of vocabulary knowledge) and listening comprehension (pp. 590-591).

Second, the following research question was addressed: (b) To what extent does depth of vocabulary knowledge, in addition to vocabulary size, contribute to successful listening comprehension? In order to determine the unique contribution of each of the two independent variables (i.e., vocabulary size and depth of vocabulary knowledge), Stæhr (2009) conducted a stepwise multiple regression analysis. The results indicated that depth of vocabulary knowledge explained 43% of the variance of the dependent variable (i.e., listening comprehension scores) and that vocabulary size explained 49% of the variance of the dependent variable (pp. 591-592).

Third, Stæhr (2009) examined the following research question: (c) How much vocabulary is needed for adequate listening comprehension? It is needless to say that the answer to this question will, of course, depend on what adequate listening comprehension means. As Stæhr (2009) lucidly puts it, “the vocabulary size or lexical coverage required for comprehension will always depend on the degree of comprehension needed” (p. 592).
In this study, the obtained data suggested that if adequate listening comprehension was defined as scoring higher than 70% on an advanced listening comprehension test, “5,000

word families might be a useful vocabulary size target for advanced foreign language listeners" (p. 596). Additionally, the findings also suggested that if adequate listening comprehension was defined as scoring higher than 80% on an advanced listening comprehension test, a receptive vocabulary knowledge of 10,000 word families might be necessary (pp. 592-596).

The relationship between vocabulary knowledge (i.e., vocabulary size and depth of vocabulary knowledge) and listening comprehension is one of the issues second language vocabulary research should deal with, and, in my view, this article by Stæhr (2009) is one of the basic readings for those wishing to examine and substantiate the relationship.

English L1 and L2 Speakers' Knowledge of Lexical Bundles

Nekrasova (2009) investigated English L1 and L2 speakers' knowledge of lexical bundles.

After providing a brief overview of research conducted on formulaic sequences in "phraseology" (pp. 648-649), "pragmatics" (p. 649), and "corpus linguistics" (pp. 649-650), Nekrasova (2009) gave a description of *lexical bundles,* a type of formulaic sequences (pp. 650-651).

Lexical bundles are defined as the most frequently occurring sequences of three or more words that show a statistical tendency to co-occur. Lexical bundles were initially identified in two major registers of *Longman Grammar of Spoken and Written English* (Biber, Johansson, Leech, Conrad, & Finegan, 1999) — conversation and academic prose — as units that occurred at least 10 times per million words (Nekrasova, 2009, p. 650).

Examples of lexical bundles include the following (p. 680): (a) referential bundles[3] (e.g., *is one of the, one of the most, the rest of the, the top of the, in the form of, in the middle of, in the case of, in terms of the, as a result of, on the basis of, in the absence of, the way in which, in the presence of, at the same time, the beginning of the, at the end of,* and *or something like that*) and (b) discourse-organizing bundles[4] (e.g., *what do you think, if you look at, if you have a, to look at the, what I want to, want to do is, want to talk about, I would like to, take a look at, a little bit about, to do with the, nothing to do with, on the other hand, as well as the,* and *know what I mean*).

After reviewing "L1 and L2 research on lexical bundles" (pp. 651-652) and "research on psychological reality of lexical bundles" (pp. 652-654), Nekrasova (2009) described two experiments (i.e., Experiment 1 and Experiment 2) (pp. 655-669). Both of them were designed to examine whether lexical bundles were recognized by L1 and L2 English speakers as holistic units. Participants' recognition of lexical bundles as holistic units was operationalized as (a) their ability to produce them as fixed units in both short and extended pieces of discourse, (b) their ability to produce lexical bundles in a contextually appropriate

manner, and (c) their use of lexical bundles to ease the processing burden during text comprehension and subsequent production (p. 654).

As for Experiment 1, the participants were L1 English speakers (n = 20), advanced L2 English speakers (n = 18), and intermediate L2 English speakers (n = 23), all of whom were undergraduate and graduate students at a regional university in the western United States (p. 655).

In Experiment 1, a gap-filling task was used. This task was designed to measure "whether [the] participants were able to recognize and produce the missing parts of the target lexical bundles based on the surrounding context" (p. 656). The test materials used in Experiment 1 consisted of 32 sentences with embedded lexical bundles that performed two different discourse functions: *discourse-organizing* bundles (n = 15) and *referential* bundles (n = 17). The two sets of lexical bundles were matched for frequency (p. 656).

The gap-filling task was designed in the following way. First, 32 lexical bundles and their functions were identified from Biber, Conrad, and Cortes's (2004) corpus-based study of university discourse. The study was based on an analysis of texts from university registers in the TOEFL 2000 Spoken and Written Academic Language Corpus and focused on four-word bundles that occurred 40 or more times per million words. Second, using the academic subcorpus of *Longman Grammar of Spoken and Written English* (Biber, Johansson, Leech, Conrad, & Finegan, 1999), Nekrasova (2009) embedded each lexical bundle within an attested context in which the function of the bundles (i.e., discourse-organizing or referential) would remain the same as initially identified. Third, one content word within a bundle was deleted with space provided to be filled in by the participants. Finally, all test items were randomly ordered and presented as a list of sentences (pp. 656-657).

Experiment 1 employed a 2 x 3 mixed design, with discourse function (discourse-organizing vs. referential) serving as a within-participant variable and participant group (L1 vs. advanced L2 vs. intermediate L2 speakers) serving as a between-participant variable (p. 657).

The results of Experiment 1 showed that there was a significant main effect for group, $F(2, 112.91) = 132.90$, $p < .05$, omega^2 = .39, that there was a significant main effect for bundle type, $F(1, 112.91) = 59.97$, $p < .05$, Cohen's d = .80, but that there was no significant interaction between group and bundle type. Furthermore, a pairwise comparison of the test scores for the three participant groups using the Bonferroni procedure indicated that there was a significant difference in the scores between L1 and intermediate L2 speakers, $p < .001$, Cohen's d = 3.88, that there was a significant difference in the scores between advanced and intermediate L2 speakers, $p < .001$, Cohen's d = 3.22, but that there was no statistically significant difference in the scores between L1 speakers and advanced L2 speakers. Additionally, all the three participant groups displayed greater familiarity

with the targeted discourse-organizing bundles than referential bundles (pp. 658-660).

With regard to Experiment 2, the L1 participants were 21 native speakers of American English who were students at a regional university in the western United States. None of them took part in Experiment 1. The L2 participants were 40 English as a foreign language learners enrolled in a public university in western Siberia, Russia. All of them were native speakers of Russian. The L2 learners were divided into the following two groups based on the results of a cloze test which was used to assess their general proficiency in English: (a) the higher proficiency group (n = 23) and (b) the lower proficiency group (n = 17) (pp. 661-662).

What follows is the lexical bundles tested in Experiment 2 (p. 683): (a) referential bundles (*was one of the, the rest of the, the top of the, in terms of the, than or equal to,* and *the nature of the*) and (b) discourse-organizing bundles (*to do with the, nothing to do with, on the other hand, as well as the, has to do with,* and *in this chapter we*). Nekrasova (2009) selected these 12 lexical bundles[5] based on Biber, Conrad, and Cortes's (2004) corpus-based study of university discourse (Nekrasova, 2009, p. 662).

In Experiment 2, a dictation activity was employed to examine the participants' immediate recall of lexical bundles. In this dictation activity, the participants listened to a recorded text divided into sections and recalled the text section by section. The text used for the dictation was a section from an introductory sociology textbook; it was adapted (a) in consideration of the proficiency level of the lower proficiency L2 group and (b) to provide context for the lexical bundles. The adapted text consisted of 14 sentences that included the 12 target lexical bundles. After the text was modified, it was divided into 13 sections (with 1 of the 13 sections containing two sentences), so that the participants could process and recall the text section by section, rather than the whole text at once. The mean length of a section was 25 words. Some sections contained more than one lexical bundle, and four sections did not include any lexical bundles. A native speaker of American English read aloud each section of the text twice at a normal speed, and this audio recording was used in Experiment 2 (pp. 662-663).

In addition to the above dictation, a questionnaire and a cloze test were also employed in Experiment 2. The questionnaire was used in order to "collect information about linguistic choices that [the] participants made during the dictation" (p. 663). The cloze test was employed with a view to obtaining the L2 participants' integrative measure of proficiency in English (pp. 663-664).

The main results of Experiment 2 can be summarized as follows.

The descriptive statistics showed that the higher proficiency L2 speakers recalled more lexical bundles (M = 6.83, SD = 2.29) than both the L1 speakers (M = 5.14, SD = 1.85) and the lower proficiency L2 speakers (M = 3.65, SD = 1.90). Additionally, all the three participant groups recalled the discourse-organizing bundles more frequently than the

referential bundles (pp. 666-667).

Next, the obtained data were analyzed using a 3 x 2 analysis of variance, with participant group serving as a between-participant variable (the L1 speakers vs. the higher proficiency L2 speakers vs. the lower proficiency L2 speakers) and function of lexical bundles serving as a within-participant variable (discourse-organizing vs. referential bundles).

The results indicated that, first, the effect of participant group was significant, $F(2, 111.112) = 17.84$, $p < .05$, omega^2 = .13. A pairwise comparison of the participant groups using the Bonferroni procedure showed that the higher proficiency L2 group recalled significantly more lexical bundles than both the L1 group ($p < .05$, Cohen's $d = .81$) and the lower proficiency L2 group ($p < .001$, Cohen's $d = 1.51$). Additionally, the L1 group recalled significantly more bundles than the lower proficiency L2 group ($p < .05$, Cohen's $d = .79$). Second, in respect of the effect of function of lexical bundles, it was also significant, $F(1, 111.112) = 95.34$, $p < .05$, Cohen's $d = 1.56$, which corroborates the aforementioned finding that the discourse-organizing bundles were recalled more frequently than the referential bundles (p. 667).[6]

The learning of lexical bundles is undoubtedly one of the issues second language vocabulary research should deal with, and I think that this article by Nekrasova (2009) is a valuable addition to the existing literature on lexical bundles.

Measuring Second Language Lexical Growth Using Hypernymic Relationships

Crossley, Salsbury, and McNamara (2009) are interested in measuring second language (L2) lexical growth using hypernymic relationships.

Hypernymic relationships are, according to Crossley et al.[7] (2009), "semantic links between conceptually related words such as *animal* and *dog*" (pp. 308-309). "In this example, *animal* is semantically linked to *dog* but functions as a superordinate term, as it is more abstract than the concrete, subordinate term *dog*" (p. 309). The researchers offer an additional explanation of hypernymic relationships[8] as follows:

> Hypernymic relations are hierarchical associations between hypernyms (superordinate words) and hyponyms (subordinate words). A hypernym is defined as a word that is more general than a related word (*animal* compared to *dog*) and a hyponym is more specific than a related word (*dog* as compared to *animal*). Another example of hypernymy is the association between *car* and *vehicle*. In this case, *car* is the hyponym of the hypernym *vehicle* because *car* has a narrower and more specific denotative scope than *vehicle,* which would also include *trucks, go-carts, golf carts,* and *hearses.* (p. 309)

The purpose of Crossley et al.'s (2009) study is "to explore whether L2 learners develop hypernymic lexical relationships as their lexicon grows and, if so, how such growth relates to the development of lexical networks" (p. 315). To this end, the researchers use WordNet to assess the spoken utterances of L2 learners. (WordNet is a computational, lexical database meant to emulate lexical networks.) They examine whether WordNet indexes of hypernymic relations increase or decrease as learners acquire an L2, and make the prediction that as L2 learners develop lexical knowledge, there will be a corresponding increase in lexical hypernymy levels in a manner similar to the growth of hypernymic relations in an L1 (p. 315).

The participants of this study were six English as a second language learners enrolled in an intensive English program at a large American university. They ranged in age from 18 to 29 years, and their first languages were as follows: Spanish (n = 1), Japanese (n = 1), Korean (n = 1), and Arabic (n = 3). The participants were interviewed every 2 weeks (not including program and university breaks) over a 1-year period. The elicitation sessions were scheduled such that they coincided with the students' regular speaking class. Their proficiency levels were tested upon arrival to the program, and they were placed into the lowest proficiency level, Level 1, of a six-level program (p. 316).

The spoken data collected from the six learners were transcribed, and a total of 99 transcriptions were collected. In preparation for the analysis of the learner corpora, the transcriptions were modified in the following ways (p. 317): (a) Interjections were deleted; (b) Any words that were clearly non-English words were deleted as well; and (c) All[9] punctuation marks except the period and question mark were eliminated from the transcriptions.

In this study, Crossley et al. (2009) used Coh-Metrix (Graesser, McNamara, Louwerse, & Cai, 2004) to examine the lexical development of the six learners. Coh-Metrix is a computational tool which "integrates lexicons, pattern classifiers, part-of-speech taggers, syntactic parsers, shallow semantic interpreters, and other components that have been developed in the field of computational linguistics" (p. 318). Coh-Metrix also provides WordNet measures for hypernymy, and Crossley et al. (2009) used these WordNet values to evaluate the growth of hypernymic relations (p. 318).

Besides, with a view to examining whether the L2 learners exhibit signs of lexical growth over the course of a year, Crossley et al. (2009) employed Measure of Textual and Lexical Diversity (MTLD) that was also reported by Coh-Metrix. MTLD is a lexical diversity (LD) measure that helps alleviate potential problems pertaining to text length. MTLD was selected to link the WordNet hypernymy data to lexical growth because LD measures are commonly used to analyze lexical growth. MTLD is also similar to other measures of LD such as D or TTR, but, unlike other LD measures, MTLD does not vary as a function of text length for text segments whose length is in the 100-2,000-word range.

MTLD also allows for comparisons between text segments of largely different lengths (up to 2,000 words) and produces reliable results over a wide range of genres while strongly correlating with other LD measures (pp. 318-319).

The results of their study showed that, first, hypernymic relations developed over time. In other words, as time spent learning English increased, so did hypernymy levels. This was verified by analyzing the obtained data with the use of the following four: a repeated measures analysis of variance, a growth modeling analysis, a correlation test between time spent learning English and WordNet hypernymy values, and a correlation test between WordNet hypernymy values and MRC concreteness values[10] (pp. 319-322). Second, lexical diversity grew as time spent learning English increased. The researchers verified this with the use of the following three: a repeated measures analysis of variance with MTLD values, a growth modeling analysis, and a correlation test between time spent learning English and MTLD values (pp. 322-323). Third, it was also found that the L2 learners' hypernymy values given by WordNet and their lexical diversity values as expressed by MTLD were correlated significantly, suggesting that the development of hypernymic relations in L2 learners was related to that of lexical diversity (pp. 322-323).

The Effects of Vocabulary Test Announcement, Word Relevance, and Vocabulary Task on Second Language Vocabulary Learning

Peters, Hulstijn, Sercu, and Lutjeharms (2009) investigated the effects of three vocabulary teaching techniques on second language vocabulary learning.

Specifically, the three techniques are the following (p. 115): (a) telling students that the reading task will be followed by a vocabulary test (vocabulary test announcement), (b) forcing students to pay attention to unfamiliar words in the reading text via comprehension questions (task-induced word relevance), and (c) requiring students to pay attention to the unfamiliar words again, after completion of the reading task (vocabulary task).

These researchers' specific focus was on how the three variables (i.e., vocabulary test announcement, task-induced word relevance, and vocabulary task) and the interaction among them would affect students' look-up behavior (i.e., online dictionary use) and word retention (p. 115).

The participants of this study were 137 college students in Belgium (L1 = Dutch, L2 = German). Their proficiency level in German was considered to be upper intermediate to advanced (p. 120).

After reviewing the relevant literature on “test announcement” (p. 116), “word relevance” (pp. 116-117), and “word-focused activities” (pp. 117-118), Peters et al. (2009) presented the following research questions (pp. 118-119): (a) What is the effect of

vocabulary test announcement on students' look-up behavior?; (b) What is the effect of word relevance on students' look-up behavior?; (c) Is there a significant interaction effect between vocabulary test announcement and word relevance on students' look-up behavior while they read a text?; (d) Is there a significant interaction effect between vocabulary test announcement and word relevance on students' look-up behavior while they carry out a vocabulary task?; (e) What is the effect of vocabulary test announcement on word retention?; (f) What is the effect of vocabulary task on word retention?; (g) What is the effect of word relevance on word retention?; and (h) Is[11] there a significant interaction effect among vocabulary test announcement, word relevance, and vocabulary task on word retention?

As for research design, this study adopted a factorial design with two between-subject independent variables and one within-subject independent variable. To be concrete, the between-subject variables consisted of *vocabulary test announcement (presence or absence)* and *vocabulary task (presence or absence)*; the within-subject variable was *word relevance (plus-relevant or minus-relevant target words)* (p. 119).

It should be noted that the within-subject variable (i.e., word relevance) was operationalized by asking text comprehension questions in the following way. "Half of the target words appeared in text passages whose content was focused on by the comprehension questions; the other target words appeared in passages whose content was not targeted by the comprehension questions" (p. 120). Peters et al. (2009) call the former plus-relevant target words; they call the latter minus-relevant target words. That is, in order to answer the comprehension questions, the participants had to process half of the target words (i.e., the plus-relevant target words). In contrast, with regard to the minus-relevant target words, they did not have to process these words as elaborately as they had to process the plus-relevant ones (p. 120).

All the participants had to read a text in German; the German text contained 16 pseudowords. These pseudowords (the target words of this study) appeared only once in the text. The participants were allowed to look up the meanings of the 16 target words and those of all low-frequency words (those not belonging to the 2,000 most frequent words in German). If they clicked on these words, a window popped up — a window which showed the translation of a clicked word in L1 as well as its meaning paraphrase in L2 (p. 119).

As regards the dependent variables of this study, they are the following: (a) look-up behavior and (b) word retention. In respect of the former (i.e., the participants' look-up behavior), it refers to such things as whether a target word was looked up or not and the frequency of clicks on the target words. The latter (i.e., word retention) was measured in three immediate and two delayed receptive vocabulary tests. Specifically, the three immediate vocabulary tests consisted of (a) Word Form Recognition Test, (b) Isolation

Test, and (c) Context Test. In the first test (i.e., the Word Form Recognition Test), the participants were shown 40 words comprising the 16 target words and 24 distractor words, and required to tick off whether the word had appeared in the text. In the second test (i.e., the Isolation Test), the participants had to provide a German definition or Dutch translation for the 16 target words, which were presented in isolation. In the third test (i.e., the Context Test), the target words were presented in the original sentences as they appeared in the text. With respect to the delayed vocabulary tests, the aforementioned Isolation Test and Context Test were used to examine the participants' word retention two weeks after the immediate tests were administered (p. 120, pp. 124-125).

An analysis of the obtained data indicated the following. As for the participants' look-up behavior (pp. 128-131), the results showed that, first, both test announcement and word relevance had a moderate to large effect on the participants' look-up behavior. Second, when forewarned of a vocabulary test, the participants clicked more frequently on the target words than when not forewarned. Third, plus-relevant target words were looked up more often than minus-relevant ones. Fourth, in the case of the vocabulary task, exactly the opposite occurred; that is, the participants clicked more often on the minus-relevant words than on the plus-relevant words. Fifth, there was only one statistically significant interaction between vocabulary test announcement and word relevance; that is, the participants forewarned of the vocabulary test looked up more minus-relevant words than those not forewarned (p. 131).

In respect of the participants' word retention (pp. 131-139), the main findings showed that, first, the effect of vocabulary test announcement on word form recognition was statistically significant, $F(1, 131) = 5.48$, $p = .02$, $\text{eta}^2 = .02$.[12] In other words, the participants who had been forewarned of an upcoming vocabulary test recognized the target word forms better than the participants who had not been forewarned.

Second, the effect of vocabulary task on word form recognition was also statistically significant, $F(1, 131) = 101.28$, $p < .0001$, $\text{eta}^2 = .42$.[13] What this means is that the participants who accomplished a vocabulary task performed better on the word form recognition test than those who did not do this task.

Third, regarding word retention as measured by the Isolation Test, the effect of vocabulary task was significant, $F(1, 96) = 27.37$, $p < .0001$, $\text{eta}^2 = .21$.[14] Additionally, the effect of word relevance was also significant, $F(1, 96) = 11.47$, $p = .001$, $\text{eta}^2 = .10$.[15] In other words, the two independent variables (i.e., both vocabulary task and word relevance) yielded significant effects on word retention.

Fourth, with regard to word retention as measured by the Context Test, the effect of vocabulary task was significant, $F(1, 103) = 54.64$, $p < .0001$, $\text{eta}^2 = .34$.[16] Besides that, the effect of word relevance was also significant, $F(1, 103) = 47.93$, $p < .0001$, $\text{eta}^2 = .26$.[17] In other words, as in the case of the Isolation Test, word retention as measured by the

Context Test was affected by both vocabulary task and word relevance.

I think that this article by Peters et al. (2009) is one of the basic readings for those interested in the effects of vocabulary test announcement, task-induced word relevance, and vocabulary task on second language learners' look-up behavior and word retention; to my knowledge, it is one of the first, if not the first, studies that investigated these independent variables' effects on second language vocabulary learning.

The Effects of Learning Method and Cognate Status on Second Language Vocabulary Learning by Children

Tonzar, Lotto, and Job (2009) investigated the effects of learning method and cognate status on second language vocabulary learning by children.

Tonzar et al. (2009) conducted two pretests before administering their main experiment. The first one was a picture-naming-in-L1 task, and this task was used to make sure that the pictures they selected were consistently named with a single name. This pretest involved 142 pictures, some of which were different versions of the same object. Specifically, objects belonging to 12 semantic categories (buildings, vehicles, weapons, musical instruments, animals, plants, fish, kitchen utensils, fruits, tools, clothes, and vegetables) were depicted. The participants of this pretest were 20 Italian fourth graders (nine-year-olds). Each picture was presented to all the participants in the classroom for 15 seconds, and they were asked to write down the names of the depicted objects (p. 627).

As the second pretest, the researchers made a cognate-rating study; the participants were 20 third-year middle school Italian children (13-year-olds). These participants were presented with 63 Italian-English, Italian-German, and English-German word pairs corresponding to the names of the pictures used in the first pretest, and were asked to evaluate the orthographic similarity within each word pair on a 7-point scale (p. 627).

Tonzar et al. (2009) obtained two measures from the above pretests: an index of name agreement and an index of orthographic similarity. They used these indexes to select the experimental material, which consisted of a list of 40 pictures with the corresponding names in Italian, English, and German. Specifically, the 40 items consisted of 20 cognates (e.g., *Piramide* [Italian], *Pyramid* [English], and *Die Pyramide* [German]) and 20 noncognates (e.g., *Freccia* [Italian], *Arrow* [English], and *Der Pfeil* [German]) (p. 628, pp. 644-646).

Next, let me move on to Tonzar, Lotto, and Job's (2009) main study, which comprised 123 fourth graders and 106 eighth graders. These participants were taken from eight schools in Treviso, Italy, and all of them were native speakers of Italian. At the time of the study, the fourth graders had not learned any foreign language; in contrast, the eighth graders had learned some English in the sixth and seventh grades (p. 629).

As for research design, Tonzar et al. (2009) adopted a between-subject design. That is, half of the participants were asked to learn the experimental material using a word-learning method, and the other half were asked to learn using a picture-learning method. In the former method (i.e., the word-learning method), "each Italian word was paired with the corresponding words in English (Italian-English set) and in German (Italian-German set)" (p. 629). In the latter method (i.e., the picture-learning method), "each picture was paired with the corresponding English and German words" (p. 629).

Tonzar et al. (2009) had their participants take two immediate tests (i.e., Immediate Test 1 and Immediate Test 2) and two delayed tests (i.e., Delayed Test 1 and Delayed Test 2). Immediate Test 1 was administered at the end of the first learning session, and Immediate Test 2 was administered at the end of the second learning session.[18] Delayed Test 1 was administered one week after Immediate Test 1, and Delayed Test 2 was administered one month after Immediate Test 2 (p. 630).

Overall, the results suggested that the picture-learning method led to better performance than the word-learning method in the cases of both the fourth and eighth graders. Additionally, this tendency manifested itself in both of the two second languages (i.e., English and German). Furthermore, the results also suggested that cognates were easier to learn than noncognates for both the fourth and eighth graders (pp. 630-636).

The Effect of Keeping Vocabulary Notebooks on Second Language Vocabulary Learning

Walters and Bozkurt (2009) investigated the effect of keeping vocabulary notebooks on second language vocabulary learning.

After reviewing previous studies on vocabulary learning strategies and vocabulary notebooks (pp. 404-405), Walters and Bozkurt (2009) presented the following two research questions (p. 405): (a) How does the use of vocabulary notebooks affect students' vocabulary acquisition (receptive, controlled productive, and free productive)?; and (b) What[19] are students' and a teacher's attitudes towards the use of vocabulary notebooks?

This study comprised 60 learners of English as a foreign language; their ages ranged from 17 to 20 years.[20] Specifically, there were three intact lower intermediate level classes. One of them served as a treatment group, and the other two served as control groups (Group A and Group B). The treatment group (n = 20) consisted of 12 males and 8 females. Regarding the control groups, Group A (n = 20) was composed of 13 males and 7 females, and Group B (n = 20) comprised 10 males and 10 females (p. 406).

A vocabulary notebook program was drawn up and implemented in the treatment group over a 4-week period. The other two groups (i.e., Group A and Group B) followed the same curriculum with the same materials as the treatment group without keeping

vocabulary notebooks. The participants of all the three groups received a receptive vocabulary test and a controlled productive vocabulary test as pretests three weeks before the implementation of the vocabulary notebook program. Besides, the above two tests were also used as posttests, and administered to all the participants at the beginning of the week that followed the end of the 4-week program. Additionally, the participants of all the three groups were asked to write free compositions. Specifically, with a view to examining the effect of the vocabulary notebook program on free use of target words,[21] they were assigned a topic at the end of each week and asked to write a composition on a topic consistent with the theme of the week's lessons (pp. 406-409).

The main results of Walters and Bozkurt's (2009) study can be summarized as follows. First, in respect of the receptive vocabulary test, the obtained data revealed that there was a significant difference in gain scores (i.e., improvement from pretest to posttest) between the treatment group and Group A (in favor of the treatment group), that there was a significant difference in gain scores between the treatment group and Group B (in favor of the treatment group), but that there was no significant difference in gain scores between Group A and Group B. In other words, the results indicated that the treatment group demonstrated greater gains in receptive knowledge of the target words than the control groups (i.e., Group A and Group B) (pp. 409-411).

Second, with respect to the controlled productive vocabulary test, the findings showed that there was a significant difference in gain scores (i.e., improvement from pretest to posttest) between the treatment group and Group A (in favor of the treatment group), that there was a significant difference in gain scores between the treatment group and Group B (in favor of the treatment group), but that there was no significant difference in gain scores between Group A and Group B. That is, the results pertaining to the controlled productive vocabulary test showed that as in the case of the receptive vocabulary test, the treatment group demonstrated greater gains in controlled productive knowledge of the target words than the control groups (pp. 410-412).

Third, the analysis of free composition data provided by the three groups demonstrated that the participants of the treatment group had a stronger tendency to use the target words in their free compositions than those of the control groups (pp. 411-412).[22]

Substantiating the effect of keeping vocabulary notebooks on vocabulary learning is one of the issues second language vocabulary researchers need to address, and I think that this empirical study by Walters and Bozkurt (2009) is a valuable addition to the existing literature on vocabulary notebooks.

The Effects of Pedagogic Modifications on Second Language Learners' Reading Comprehension and Vocabulary Recognition

O'Donnell (2009) investigated the effects of pedagogic modifications on second language learners' reading comprehension and vocabulary recognition.

Specifically, this study sought to examine the effects of one of the pedagogic modification techniques — textual elaboration — on Spanish as a second language learners' reading comprehension and vocabulary recognition immediately after reading short, authentic literary selections (p. 513).

After reviewing the relevant literature (pp. 512-517), O'Donnell (2009) presented the following research questions (p. 517): (a) Are comprehension scores of participants reading elaborated L2 literary texts different, quantitatively, from scores of the same readers reading unmodified versions of the same texts as measured by the amount of text information they produce immediately after reading?; (b) Do readers of elaborated and unmodified versions of the same literary texts differ in their ability to recognize, and therefore translate, words that appear in the original unmodified versions of these texts immediately after reading?; and (c) Do[23] readers of the unmodified versions of literary texts identify lexical items glossed in Spanish in the text margins differently from readers of the same texts presented without the gloss in an elaborated version immediately after reading?

As for reading materials, three texts were taken from an intermediate-level Spanish textbook. All the three texts were similar in narrative style. Moreover, the texts were approximately the same length and comparable in format and presentation. The topics treated in the three readings were (a) financial matters, (b) leisure time activities, and (c) the future (pp. 517-518).

First, O'Donnell (2009) conducted a pilot study with 63 students enrolled in six intact sections of a fourth-semester Spanish course; these students participated in think-aloud protocols while reading the three original Spanish language texts. This pilot study was conducted with a view to determining at which points in the texts the students encountered reading comprehension difficulties. Next, after analyzing the obtained data, the researcher modified the three original literary texts according to an elaborative modification procedure (p. 518).

Following the pilot study, O'Donnell (2009) made a main study with 197 undergraduate students enrolled in a fourth-semester Spanish as a second language course. These participants were classified as “late beginners or early intermediate language learners with 300 hours of Spanish instruction” (p. 518).

Each participant read two texts — one in elaborated form and one in unmodified form — from a potential text pool of six texts (i.e., two versions each of the three literary texts) (p. 519).

After finishing reading, the participants logged onto computers and were prompted to recall, in English, what they had just read. After finishing the recall protocol, they were asked to translate into English the 18 problematic words and phrases which appeared in the

texts they had just read. For clarification purposes, a sample testing item was included, taken from the first line of each of the three texts. The lexical items were presented in the chronological order in which they appeared in the texts (p. 519).

The obtained data demonstrated that, first, in respect of these participants' text comprehension, "the amount of information that the readers of the elaborated versions were able to recall proved significantly greater than the amount of information the readers of the unmodified versions were able to recall" (pp. 524-525). Second, those who read the elaborated versions could identify more vocabulary that appeared within the texts than those who read the unmodified versions (pp. 523-524, pp. 527-528).

In my view, an attempt to examine the effects of text elaboration on second language reading comprehension and vocabulary recognition is a subject of academic interest to those involved in second language reading research and vocabulary research. This is because, as O'Donnell (2009) lucidly states, "elaboration has the potential to increase comprehensibility and vocabulary recognition while preserving many of the features inherent in authentic texts" (p. 529). I think that this article by O'Donnell (2009) is a significant contribution to the advancement of this line of research (i.e., examining the effects of text elaboration on reading comprehension and vocabulary recognition), which will attract second language reading researchers' and vocabulary researchers' attention in the future.

Conclusion

In this article second language vocabulary research published in leading international research journals in 2009 was reviewed. In addition to the articles examined in the preceding sections, the following papers were also published in 2009: for example, Alcón-Soler (2009),[24] Barcroft (2009),[25] Boers, Píriz, Stengers, and Eyckmans (2009),[26] Boguslavsky, Cardeñosa, and Gallardo (2009),[27] Cheng and Good (2009),[28] Cheng, Greaves, Sinclair, and Warren (2009),[29] Durrant and Schmitt (2009),[30] Fan (2009),[31] Hall, Newbrand, Ecke, Sperr, Marchand, and Hayes (2009),[32] Hamada (2009),[33] Harrington and Carey (2009),[34] Mizumoto and Takeuchi (2009),[35] Myers and Chang (2009),[36] Pulido (2009),[37] Schwieter and Sunderman (2009),[38] Trebits (2009),[39] Ward and Chuenjundaeng (2009),[40] Webb (2009),[41] Webb and Kagimoto (2009),[42] Webb and Rodgers (2009a),[43] Webb and Rodgers (2009b),[44] Yamada (2009),[45] Yu (2009),[46] and Zyzik and Azevedo (2009).[47]

This is the eighth attempt to tackle the task of reviewing second language vocabulary research (Tanaka, 2008, 2009, 2010, 2011, 2012a, 2012b, 2014). Specifically, Tanaka (2008) examined articles published in 2006, Tanaka (2009) focused on articles published in 2007, Tanaka (2010) dealt with articles published in 2008, Tanaka (2011) investigated articles published in 2005, Tanaka (2012a) explored articles published in 2004, Tanaka (2012b) reviewed articles published in 2003, and Tanaka (2014) covered articles published

in 2002. I hope that the present review, together with the above seven (i.e., Tanaka, 2008, 2009, 2010, 2011, 2012a, 2012b, 2014), will be of help to those involved in second language vocabulary research.

Notes

[1]In the present review, there are several places in which sentences beginning with such words as (a), (b), and (c) are enumerated after a colon. It should be noted that in such cases, sentence-initial words begin with an uppercase letter (even if they are preceded by the conjunction *and*).

[2]Stæhr (2009) decided to "exclude the participants' scores on the academic word level from the final analysis" (p. 587) because of the following reason. The point is that the words at this level have frequencies that range from the 2,000 band to the 10,000 band. "These words thus occur within the other frequency bands in the Vocabulary Levels Test and should not be seen as adding to the overall vocabulary size" (p. 587).

[3]Referential bundles "make direct reference to physical or abstract entities, or to the textual context itself, either to identify the entity or to single out some particular attribute of the entity as especially important" (Biber, Conrad, & Cortes, 2004, p. 384). For examples of referential bundles, see Biber, Conrad, and Cortes (2004, pp. 387-388); for a detailed description of referential bundles, see Biber, Conrad, and Cortes (2004, pp. 393-396).

[4]Discourse-organizing bundles "reflect relationships between prior and coming discourse" (Biber, Conrad, & Cortes, 2004, p. 384). For examples of discourse-organizing bundles, see Biber, Conrad, and Cortes (2004, pp. 386-387); for a detailed description of discourse-organizing bundles, see Biber, Conrad, and Cortes (2004, pp. 391-393).

[5]Among these 12 lexical bundles, seven items appeared in both Experiment 1 and Experiment 2. The following five were used only in Experiment 2: *was one of the, than or equal to, the nature of the, has to do with,* and *in this chapter we* (Nekrasova, 2009, p. 662).

[6]The interaction between participant group and function of lexical bundles was not significant (Nekrasova, 2009, p. 667).

[7]In the present review, "et al." is not italicized (*Publication Manual of the American Psychological Association,* 2010, p. 105).

[8]In Crossley, Salsbury, and McNamara's (2009) article, they seem to use both *hypernymic relationships* and *hypernymic relations* interchangeably.

[9]See Note 1.

[10]MRC concreteness values are "based on the work of Paivio, Yuille, and Madigan (1968), Toglia and Battig (1978), and Gilhooly and Logie (1980), who asked participants to score nouns based on concreteness on a numerical scale (from 1 to 7)" (Crossley, Salsbury,

& McNamara, 2009, p. 318). Crossley et al. (2009) found a significant positive correlation (r = 0.62, p < .001, N = 99) between WordNet hypernymy values and MRC concreteness values (p. 322).

[11]See Note 1.

[12]These figures are taken from Table 9 (Peters et al., 2009, p. 132).

[13]As in the case of the preceding note (i.e., the note on the effect of vocabulary test announcement), these figures are taken from Table 9 (Peters et al., 2009, p. 132).

[14]These figures are taken from Table 11 (Peters et al., 2009, p. 135).

[15]As in the case of the preceding note (i.e., the note on the effect of vocabulary task), these figures are taken from Table 11 (Peters et al., 2009, p. 135).

[16]These figures are taken from Table 13 (Peters et al., 2009, p. 137).

[17]As in the case of the preceding note (i.e., the note on the effect of vocabulary task), these figures are taken from Table 13 (Peters et al., 2009, p. 137).

[18]The second learning session was administered one week after the first learning session (Tonzar, Lotto, & Job, 2009, p. 629).

[19]See Note 1.

[20]Walters and Bozkurt's (2009) study was conducted at English Language Preparatory School, Zonguldak Karaelmas University, Turkey (Walters & Bozkurt, 2009, p. 405).

[21]"Eighty target words were chosen from the 4 coursebook units that would be covered over the 4-week period" (Walters & Bozkurt, 2009, p. 406).

[22]In addition to the receptive vocabulary test, the controlled productive vocabulary test, and the free compositions, (a) interviews with the participants of the treatment group and (b) an interview with the teacher of the treatment group were conducted with a view to discovering "the students' and the teacher's attitudes towards the use of vocabulary notebooks" (Walters & Bozkurt, 2009, p. 408). As for the above (a), selected excerpts of the interviewees' comments are described in Section 4 of the Results section (Walters & Bozkurt, 2009, pp. 412-415); as regards the above (b), some of the teacher's comments are shown in Section 5 of the Results section (pp. 415-416).

[23]See Note 1.

[24]A study by Alcón-Soler (2009) investigated the relationship among focus on form, learner uptake, and subsequent lexical gains in learners' oral production.

This study comprised 12 Spanish learners of English as a foreign language aged 14-15 (pp. 350-351).

[25]Barcroft (2009) investigated the effect of synonym generation on second language vocabulary learning during reading in both incidental and intentional vocabulary learning contexts.

The participants of this study were 114 Spanish-speaking university students learning English as a second language at a large university in Mexico City. They consisted of

low-intermediate learners of English (*n* = 59) and high-intermediate learners of English (*n* = 55) (p. 90).

[26]Boers, Píriz, Stengers, and Eyckmans (2009) attempted to examine the effect of pictorial elucidation on second language idiom learning.

This study comprised 38 learners of English aged 19-21. They were language majors at a college in Brussels, Belgium (p. 372).

[27]Boguslavsky, Cardeñosa, and Gallardo's (2009) article aimed at proposing a new approach to creating disambiguated multilingual dictionaries.

[28]A study by Cheng and Good (2009) examined the effects on EFL learners' reading comprehension and vocabulary retention of the following three first language (L1) glosses: (a) L1 Chinese glosses plus L2 English example sentences, (b) L1 Chinese in-text glosses, and (c) L1 Chinese marginal glosses.

Their study comprised 135 non-English major undergraduates (business and engineering majors) at a national university of science and technology in southern Taiwan (p. 122).

[29]Cheng, Greaves, Sinclair, and Warren's (2009) article describes an analytical procedure for identifying phraseological variation within concgrams, which are sets of words that co-occur regardless of constituency variation, positional variation, or both (p. 236). For concgrams, see Cheng, Greaves, and Warren (2006).

[30]A study by Durrant and Schmitt (2009) investigated and described the extent to which nonnative writers of English made use of collocations in comparison to English native speaker norms.

[31]Fan (2009) attempted to examine the use of collocations by English as a second language (ESL) learners using two corpora based on the writings of Hong Kong ESL learners and those of native speakers of British English.

[32]A study by Hall, Newbrand, Ecke, Sperr, Marchand, and Hayes (2009) addresses the problem of "how adult learners establish a provisional hypothesis about the grammatical properties of new words in a third language (L3), paying particular attention to the role played by presumed translation equivalents in the first language (L1) and second language (L2) lexical network" (p. 154).

Their study is framed by the claims of the Parasitic Model of vocabulary acquisition (Hall, 2002; Hall & Ecke, 2003). The model posits that "the default cognitive processes involved in word learning are driven by automatic, unconscious detection of similarity between the new input and information already stored in the mental lexicon, be it from the L1, L2, or L3" (p. 154).

[33]A study by Hamada (2009) examined second language learners' word meaning inference behaviors.

The participants of this study were five Japanese ESL learners (two males and three

females) studying at a mid-size university in the United States (pp. 449-450).

[34]Harrington and Carey (2009) attempted to evaluate the validity of an on-line Yes/No test of recognition vocabulary as a placement tool at an Australian English language school.

A Yes/No vocabulary test uses a format in which "word and pseudoword items are presented sequentially, with the testee indicating whether or not s/he knows the word" (p. 614).

Harrington and Carey's (2009) study was conducted at Milton College, an established English language school located in Sydney, Australia. Specifically, newly entering students (n = 88) completed a Yes/No vocabulary test and a placement test battery used at Milton College. The ages of the participants ranged from 19 to 33 with many intending to continue on to university study in Australia or elsewhere at the end of language study at Milton College (pp. 615-617).

[35]Mizumoto and Takeuchi (2009) investigated the effectiveness of explicit instruction of vocabulary learning strategies with 146 Japanese learners of English as a foreign language aged 18-22.

[36]A study by Myers and Chang (2009) investigated the effect of a vocabulary teaching approach on second language word and collocation learning.

The participants of this study were 115 eleventh graders in a municipal senior high school in Taipei, Taiwan; they were taken from three intact EFL classes. Among the three, two served as experimental groups and the other as a control group (p. 181, p. 185).

[37]A study by Pulido (2009) investigated the effects of reader-based factors (i.e., second language reading proficiency and background knowledge) on second language lexical input processing while learners engaged in strategic tasks to promote deeper processing of new words.

The participants of this study were 35 English-speaking adult learners of Spanish. They were divided into the following: (a) 8 beginner (intensive course for false beginners), (b) 8 intermediate (fourth-semester language course), (c) 6 high-intermediate (fifth- and sixth-semester grammar and composition course), (d) 11 advanced (seventh- and eighth-semester literature), and (e) 2 high-advanced learners (MA graduate literature course) (p. 38).

[38]A study by Schwieter and Sunderman (2009) pertains to second language lexical processing, and examines differential predictions of La Heij's (2005) concept selection model (CSM) and Kroll and Stewart's (1994) revised hierarchical model (RHM).

[39]Trebits (2009) examined English language EU documents with a view to identifying the most frequent phrasal verb constructions.

[40]Ward and Chuenjundaeng (2009) attempted to examine second language learners' suffix knowledge.

[41]A study by Webb (2009) investigated the effect of pre-learning vocabulary on

reading comprehension and writing.

[42]Webb and Kagimoto (2009) investigated the effects of receptive and productive vocabulary learning tasks on learning collocation and meaning.

[43]Webb and Rodgers (2009a) investigated the lexical coverage of movies. Specifically, they sought to determine "the vocabulary size necessary to understand 95% and 98% of the words in movies" (p. 407).

[44]Webb and Rodgers (2009b) conducted a study aimed at determining the vocabulary demands of television programs.

[45]Yamada (2009) conducted a study aimed at proposing a way to assess textual gist identification by second language readers.

[46]Yu (2009) attempted to examine the effects of two teaching methods on the learning of a formulaic sequence, *despite the fact.*

[47]Zyzik and Azevedo (2009) investigated second language learners' knowledge of word class distinctions (e.g., noun vs. adjective) in a variety of syntactic contexts.

References

Alcón-Soler, E. (2009). Focus on form, learner uptake and subsequent lexical gains in learners' oral production. *IRAL, 47,* 347-365.

American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.

Barcroft, J. (2009). Effects of synonym generation on incidental and intentional L2 vocabulary learning during reading. *TESOL Quarterly, 43,* 79-103.

Biber, D., Conrad, S., & Cortes, V. (2004). *If you look at . . .* : Lexical bundles in university teaching and textbooks. *Applied Linguistics, 25,* 371-405. (NB: It should be noted that the title of this article [i.e., *If you look at . . .*] should be italicized.)

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

Boers, F., Píriz, A. M. P., Stengers, H., & Eyckmans, J. (2009). Does pictorial elucidation foster recollection of idioms? *Language Teaching Research, 13,* 367-382.

Boguslavsky, I., Cardeñosa, J., & Gallardo, C. (2009). A novel approach to creating disambiguated multilingual dictionaries. *Applied Linguistics, 30,* 70-92. (NB: In the case of the online version of this article, the publication year is 2008. [To be specific, the online version was published on September 11, 2008.] However, the March issue [i.e., Vol. 30, No. 1, print version] of *Applied Linguistics,* in which this article is included, was published in 2009. [To be concrete, the print version reached me on April 23, 2009.] What this means is

that the publication year of the online version is different from that of the print version. In the present review, the publication year of the print version [i.e., 2009] was adopted.)

Cheng, W., Greaves, C., & Warren, M. (2006). From n-gram to skipgram to concgram. *International Journal of Corpus Linguistics, 11,* 411-433.

Cheng, W., Greaves, C., Sinclair, J. M., & Warren, M. (2009). Uncovering the extent of the phraseological tendency: Towards a systematic analysis of concgrams. *Applied Linguistics, 30,* 236-252. (NB: In the case of the online version of this article, the publication year is 2008. [To be specific, the online version was published on November 11, 2008.] However, the June issue [i.e., Vol. 30, No. 2, print version] of *Applied Linguistics,* in which this article is included, was published in 2009. [To be concrete, the print version reached me on August 6, 2009.] What this means is that the publication year of the online version is different from that of the print version. In the present review, the publication year of the print version [i.e., 2009] was adopted.)

Cheng, Y. -H., & Good, R. L. (2009). L1 glosses: Effects on EFL learners' reading comprehension and vocabulary retention. *Reading in a Foreign Language, 21,* 119-142.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34,* 213-238.

Crossley, S., Salsbury, T., & McNamara, D. (2009). Measuring L2 lexical growth using hypernymic relationships. *Language Learning, 59,* 307-334.

Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *IRAL, 47,* 157-177.

Fan, M. (2009). An exploratory study of collocational use by ESL students: A task based approach. *System, 37,* 110-123.

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation, 12*, 395-427.

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers, 36,* 193-202.

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.

Hall, C. J. (2002). The automatic cognate form assumption: Evidence for the parasitic model of vocabulary development. *IRAL, 40,* 69-87.

Hall, C. J., & Ecke, P. (2003). Parasitism as a default mechanism in L3 vocabulary acquisition. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *The multilingual*

lexicon (pp. 71-85). Dordrecht: Kluwer.

Hall, C. J., Newbrand, D., Ecke, P., Sperr, U., Marchand, V., & Hayes, L. (2009). Learners' implicit assumptions about syntactic frames in new L3 words: The role of cognates, typological proximity, and L2 status. *Language Learning, 59,* 153-202.

Hamada, M. (2009). Development of L2 word-meaning inference while reading. *System, 37,* 447-460.

Harrington, M., & Carey, M. (2009). The on-line Yes/No test as a placement tool. *System, 37,* 614-626. (NB: It should be noted that both the first letter of "Yes" and that of "No" are writtern in uppercase letters in accordance with the way the title of this article is written [Harrington & Carey, 2009, p. 614].)

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33,* 149-174.

La Heij, W. (2005). Selection processes in monolingual and bilingual lexical access. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 289-307). New York: Oxford University Press.

Mizumoto, A., & Takeuchi, O. (2009). Examining the effectiveness of explicit instruction of vocabulary learning strategies with Japanese EFL university students. *Language Teaching Research, 13,* 425-449.

Myers, J. L., & Chang, S. -F. (2009). A multiple-strategy-based approach to word and collocation acquisition. *IRAL, 47,* 179-207.

Nekrasova, T. M. (2009). English L1 and L2 speakers' knowledge of lexical bundles. *Language Learning, 59,* 647-686.

O'Donnell, M. E. (2009). Finding middle ground in second language reading: Pedagogic modifications that increase comprehensibility and vocabulary acquisition while preserving authentic text features. *The Modern Language Journal, 93,* 512-533.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monograph Supplement, 76,* 1-25.

Peters, E., Hulstijn, J. H., Sercu, L., & Lutjeharms, M. (2009). Learning L2 German vocabulary through reading: The effect of three enhancement techniques compared. *Language Learning, 59,* 113-151.

Pulido, D. (2009). How involved are American L2 learners of Spanish in lexical input processing tasks during reading? *Studies in Second Language Acquisition, 31,* 31-58.

Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing, 10,* 355-371.

Read, J. (1998). Validating a test to measure depth of vocabulary knowledge. In A. J. Kunnan (Ed.), *Validation in language assessment: Selected papers from the 17th Language Testing Research Colloquium, Long Beach* (pp. 41-60). Mahwah, NJ: Erlbaum.

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.

Schwieter, J., & Sunderman, G. (2009). Concept selection and developmental effects in bilingual speech production. *Language Learning, 59,* 897-927.

Stæhr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition, 31,* 577-607.

Tanaka, Y. (2008). Second language vocabulary research: 2006. *Administration, 15*(1 & 2), 77-103.

Tanaka, Y. (2009). Second language vocabulary research: 2007. *Administration, 16*(1), 55-72.

Tanaka, Y. (2010). Second language vocabulary research: 2008. *Administration, 17*(1 & 2), 37-57.

Tanaka, Y. (2011). Second language vocabulary research: 2005. *Administration, 17*(3 & 4), 61-81.

Tanaka, Y. (2012a). Second language vocabulary research: 2004. *Administration, 18*(3 & 4), 259-284.

Tanaka, Y. (2012b). Second language vocabulary research: 2003. *Administration, 19*(1), 49-61.

Tanaka, Y. (2014). Second language vocabulary research: 2002. *Administration, 21*(1), 112-122.

Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms.* Hillsdale, NJ: Erlbaum.

Tonzar, C., Lotto, L., & Job, R. (2009). L2 vocabulary acquisition in children: Effects of learning method and cognate status. *Language Learning, 59,* 623-646.

Trebits, A. (2009). The most frequent phrasal verbs in English language EU documents: A corpus-based analysis and its implications. *System, 37,* 470-481.

Walters, J., & Bozkurt, N. (2009). The effect of keeping vocabulary notebooks on vocabulary acquisition. *Language Teaching Research, 13,* 403-423.

Ward, J., & Chuenjundaeng, J. (2009). Suffix knowledge: Acquisition and applications. *System, 37,* 461-469.

Webb, S. A. (2009). The effects of pre-learning vocabulary on reading comprehension and writing. *The Canadian Modern Language Review, 65,* 441-470.

Webb, S., & Kagimoto, E. (2009). The effects of vocabulary learning on collocation and

meaning. *TESOL Quarterly, 43,* 55-77.

Webb, S., & Rodgers, M. P. H. (2009a). The lexical coverage of movies. *Applied Linguistics, 30,* 407-427.

Webb, S., & Rodgers, M. P. H. (2009b). Vocabulary demands of television programs. *Language Learning, 59,* 335-366.

Yamada, K. (2009). Lexical patterns in L2 textual gist identification assessment. *Language Testing, 26,* 101-122.

Yu, X. (2009). A formal criterion for identifying lexical phrases: Implication from a classroom experiment. *System, 37,* 689-699.

Zyzik, E., & Azevedo, C. (2009). Word class distinctions in second language acquisition: An experimental study of L2 Spanish. *Studies in Second Language Acquisition, 31,* 1-29.

About the Author

Yuji Tanaka is an associate professor of applied linguistics at the Prefectural University of Kumamoto, Japan. His work mainly focuses on second language vocabulary acquisition.