Second Language Vocabulary Research: 2007

Yuji Tanaka

This is a review article on second language vocabulary research. Articles published in leading international research journals in 2007 are the scope of this investigation. The present review comprises the following key themes: the use of vocabulary by second language learners, formulaic sequences, second language learners' lexical organization, a comparison of the effects of image-schema-based instruction and translation-based instruction, the effects of word frequency and word intervention tasks on second language vocabulary learning and text comprehension, and the development of radical awareness.

The Use of Vocabulary by Second Language Learners

The Use of Multi-Word vs. One-Word Verbs

Previous research on multi-word verbs suggests that English learners often have problems with these verbs, and that they may even avoid their use. A study by
Siyanova and Schmitt (2007) explores this issue by comparing the likelihood of using multi-word vs. one-word verbs by both native speakers of English and advanced nonnatives.

The participants of this study were 65 native speakers of English and 65 advanced nonnative speakers of English. The native speakers were either undergraduate or postgraduate students at the University of Nottingham, or young professionals who graduated in the previous six years. Their age range was 20-35 years old, with 17 males and 48 females (p. 123).

As for the nonnative speakers of English, they were chosen on the basis of their advanced English language proficiency. Among the 65 nonnatives, forty participants were international undergraduate and postgraduate students who were studying a variety of subjects at the University of Nottingham. Their age range was 20-35 years old, with 8 males and 32 females. The other 25 participants were young professionals who studied English at the university level at overseas or UK universities and had graduated in the previous six years. Most of them were working as translators or interpreters, teachers of English, or held managerial and administrative posts. They all worked with English on a daily basis, and their proficiency could be considered as advanced, although no direct measures were taken of this. Some of them lived in the UK, while others resided in their home countries. Their age range was 22-35 years old, with 5 males and 20 females (p. 124).

In order to directly compare native and nonnative usage of multi-word/one-word verb pairs, a questionnaire was developed by the researchers. The pilot version of the questionnaire contained 26 verb pairs, and was given to 14 native speakers (undergraduates and postgraduates at British universities/professionals), who were asked to choose which of the two verb forms they were more likely to use in colloquial English. Because the focus of their study is on multi-word verbs, the
researchers analyzed the questionnaires to find pairs of verbs in which the majority (≥60%) of the natives preferred the multi-word form. Nineteen verb pairs showed such a multi-word preference, and only these were included in the final questionnaire used in the main study (pp. 122-123, pp. 135-138).

The 19 verb pairs were as follows: put off, postpone; tell off, reproach; put up with, stand; turn down, decrease; run into, meet; come round, come; tidy up, organize; go on, continue; pull over, stop; work out, train; set up, start; get back, return; figure out, understand; walk off, leave; mess around, misbehave; come up, arise; come up with, suggest; hold on, wait; and show off, boast (p. 135).

In addition to the aforementioned questionnaire, the researchers used the following three corpora: the Cambridge and Nottingham Corpus of Discourse in English (CANCODE), the British National Corpus (BNC), and the International Corpus of Learner English (ICLE). The researchers examined native usage of spoken multi-word verbs and their one-word counterparts by referring to the CANCODE corpus, a collection of five million words of mainly informal spoken English. The same verbs were compared in written discourse by consulting the 90 million word written component of the BNC. Learner use of the verbs in written discourse was determined by checking the ICLE, a 2.5 million word collection of 3,640 essays written by learners from eleven European languages (p. 122).

The data obtained through the questionnaire indicated that the nonnative participants were less likely to use multi-word verbs than the native speakers in informal spoken contexts (p. 119, pp. 128-129). In addition, a correlation analysis between the number of months spent in an L2 country and learners' one-word and multi-word verb preferences suggested that the amount of exposure to native-speaking environments did not increase the likelihood of using the multi-word verbs (pp. 129-130). Furthermore, the results of the corpus analysis of the same verb pairs show that
the one-word verbs are often more frequent in both written and spoken discourse (p. 119, pp. 124-125).

The learning of multi-word verbs is undoubtedly one of the issues second language vocabulary research should deal with, and I am sure that this article by Siyanova and Schmitt (2007) would be one of the basic readings in the field of multi-word-verb research.

**The Use of Emotion Vocabulary**

The purpose of Pavlenko and Driagina's (2007) study is to examine how advanced American learners of Russian use Russian emotion vocabulary to describe others' feelings. To this end, the study compared the uses of emotion vocabulary in narratives elicited from (a) monolingual speakers of Russian \( n = 49 \), (b) monolingual speakers of English \( n = 50 \), and (c) advanced American learners of Russian \( n = 30 \).

As a narrative elicitation stimulus, the researchers used a 3-minute film, *the Letter*, with a musical soundtrack but no verbal exchanges, created specifically for this study. In the film, a young woman comes home, gets her mail, opens a letter, reads it, and becomes visibly upset. Her roommate comes in, tries to talk to her unsuccessfully, sees the letter, and begins reading it without permission. The first woman notices that the roommate is reading her letter, grabs the letter, and stomps out of the room (pp. 217-218).

After watching the film, each participant was asked to describe it orally. The oral narratives were recorded with a tape recorder. The researchers transcribed all the tapes, and the transcriptions were analyzed both quantitatively and qualitatively.

The results showed that, first, monolingual speakers of English favored an adjectival pattern of emotion description. Second, monolingual speakers of Russian
favored a verbal one. Third, advanced American learners of Russian began approximating the usage of native speakers of Russian. In other words, it was revealed that “advanced American learners of Russian . . . were beginning to approximate native speakers of Russian in terms of the lexical richness of their emotion vocabulary and morphosyntactic and lexical choices, shifting from the preference for an adjectival pattern of emotion description in English to a verbal one in Russian” (p. 228).

At the same time, the data revealed 6 areas where the usage by the learners differed from that by the monolingual speakers of Russian: “morphosyntactic transfer from the first language (L1), semantic transfer from the L1, greater use of adverbal constructions, absence of a language-specific verb frequently used by native Russian speakers, violations of appropriateness of sociolinguistic register, and a significantly lower proportion of emotion word tokens” (p. 213).

In second language vocabulary research, only a few studies have been made that deal with the use and acquisition of emotion vocabulary. In my view, this article by Pavlenko and Driagina (2007) is a significant contribution to the advancement of this line of research (i.e., research that pertains to the use and acquisition of emotion vocabulary). I am sure that this article would be a valuable text for those who intend to tackle the issues concerned with the use and acquisition of emotion vocabulary by second language learners.

**Formulaic Sequences**

Formulaic sequences, or formulas, are “multiword expressions that occur as phrases and as coherent semantic units at a relatively high frequency” (Jiang & Nekrasova, 2007, p. 433). Examples of formulaic sequences include the following:
“as soon as, in any case, to begin with, take a look at, to tell the truth, on the contrary, at the moment, on the whole, as a result, at any rate, to sum up, in other words, take care of, look forward to, on the other hand, in the end, at the same time, take part in, in order to, in front of, out of work, and in the first place” (p. 445).

A focus of Jiang and Nekrasova's (2007) study is on the representation and processing of formulaic sequences. They conducted two online grammaticality judgment experiments with a view to clarifying how formulaic sequences were represented and processed.

As for Experiment 1, 20 native speakers (NSs) of English and 20 nonnative speakers (NNSs) of English participated in the study. All participants (except for 1 NS) were students at Georgia State University (2 undergraduate and 37 graduate students). There were 25 females and 15 males, and their ages ranged from 20 to 40 years (p. 436).

As regards test materials, there were 26 formulaic sequences, 26 nonformulaic sequences, and 26 ungrammatical sequences. From these 78 items two counterbalanced test lists were created. Each list included 13 formulaic sequences, 13 nonformulaic sequences, and 26 ungrammatical sequences. Experiment 1 employed a 2 x 2 x 2 mixed design, with the sequence grammaticality (grammatical vs. ungrammatical) and formulaicity (formulaic vs. nonformulaic sequences) as within-participant variables, and participant group (NS vs. NNS participants) as a between-participant variable. The participants, who were tested individually, each responded to one of the two test lists, for which their assignment was random (pp. 436-437).

The results of Experiment 1 showed that both NSs and NNSs of English responded to the formulaic sequences significantly faster and with fewer errors than they did to nonformulaic sequences.

Following Experiment 1, Experiment 2 was conducted; 22 NSs and 22 NNSs
of English participated in the study. In the case of Experiment 2, all test materials were presented in uppercase letters (not in lowercase letters as in the case of Experiment 1). This was done to rule out the possibility that the relatively rapid reaction times and fewer error rates observed for formulaic sequences in Experiment 1 were a result of the participants' familiarity with the visual shape of the test materials (pp. 439-440).

The results of Experiment 2 replicated those of Experiment 1. That is, the results indicated, as in the case of Experiment 1, that both NSs and NNSs of English responded to formulaic sequences significantly faster and with fewer errors than they did to nonformulaic sequences.

Researchers have long been interested in the acquisition and use of formulaic sequences by second language learners. I am sure that this article by Jiang and Nekrasova (2007) would be a valuable addition to the existing literature.

**Second Language Learners' Lexical Organization**

Zareva (2007) is interested in issues related to the organization of second language (L2) mental lexicon and its qualitative and quantitative features.

First, Zareva (2007) reviewed the theoretical framework underlying the use of word association (WA) tests in lexical research as developed and used in first language (L1) psychological and cognitive research. Specifically, the researcher focused on some traditional distinctions L1 researchers made between qualitative and quantitative features of WA domains which were frequently not taken into account in L2 WA research (pp. 124-127).

Second, the researcher gave a brief overview of L2 word association research. To be concrete, Zareva (2007) made a review of the application of word association
tests in L2 experimental work and some of the most significant findings concerning L2 learners' vocabulary structure (pp. 127-130).

After reviewing L1 word association research and L2 word association research, Zareva (2007) conducted an empirical study involving 29 native speakers (NSs) of English, 29 L2 advanced learners of English, and 29 L2 intermediate learners of English.¹

The analysis of the obtained data reveals that “differences in the organization of lexical knowledge between L2 speakers and NSs are quantitative rather than qualitative” (Zareva, 2007, p. 123). In addition, because there were no statistically significant differences among the three groups in the proportion of paradigmatic and syntagmatic responses;² the researcher states that “L2 advanced and intermediate learners' qualitative characteristics of lexical organization are not different from the qualitative features of NSs' mental lexicons” (p. 141).

**A Comparison of the Effects of Image-Schema-Based Instruction and Translation-Based Instruction**

A study by Morimoto and Loewen (2007) examined the effects of image-schema-based instruction and translation-based instruction on the acquisition of second language (L2) polysemous words.

The participants of their study were 58 Japanese high school learners of English studying in Japan. All of them spoke Japanese as their first language (L1) and had studied English for at least three years as a school subject. The age of the students ranged from 16 to 17 years old. The researchers used the following three intact classes. Two served as treatment groups, with the first group (n = 17) receiving image-schema-based instruction (ISBI) and the second group (n = 18) receiving translation-
based instruction (TBI). The remaining third group served as a control ($n = 23$) and received no instruction on the target words (pp. 353-354).

Each treatment group received 20 minutes of instruction on the target words, i.e., break (verb) and over (preposition) (p. 354). As stated above, the control group did not receive any instruction on the target words.

In order to examine the effects of the three conditions, an acceptability judgment test and a production test were administered prior to instruction (pre-test), two days after instruction (post-test 1) and two weeks after instruction (post-test 2) (p. 347, pp. 355-356).

The results of their study showed that image-schema-based instruction was as effective as translation-based instruction in terms of both the acceptability judgment test and the production test. In other words, contrary to the researchers' expectations, the results did not indicate that image-schema-based instruction was more effective than translation-based instruction. Except for one case where the scores of the ISBI group were significantly higher than those of the TBI group (i.e., the acceptability judgment test, over), no statistically significant differences were found between the ISBI and TBI groups.

The Effects of Word Frequency and Word Intervention Tasks on Second Language Vocabulary Learning and Text Comprehension

Rott (2007) examined the effects of word frequency and word intervention tasks on second language vocabulary learning and text comprehension.

The participants of the study were 38 learners of German. They were native speakers of English in four intact fourth-semester language classes (pp. 173-174).

As for word frequency, the following two were compared: (a) words that
occurred once, and (b) words that occurred four times in a text.

As regards word intervention tasks, the researcher compared the following three conditions: (a) Target words were glossed four times in the text (four-gloss); (b) Target words were first glossed, then retrieved in the first language, and bolded twice (gloss-retrieval); and (c) Target words were first glossed and then bolded three times (gloss-bolding) (p. 165, p. 171).

Findings revealed that the gloss-retrieval and four-gloss reading conditions resulted in more productive word gain than the gloss-bolding condition or when readers encountered a target word only once. The comprehension of main ideas was highest when the target word was glossed four times followed by the gloss-bolding reading condition and the gloss-retrieval condition (p. 165, pp. 179-187).

**The Development of Radical Awareness**

Shen and Ke (2007) examined developmental trends in acquiring knowledge of radicals, radical perception skills, and radical knowledge application skills among nonnative learners of Chinese. In addition, the researchers examined how the development of radical knowledge application skills was associated with Chinese word acquisition.

Let me give a concise description of the technical word, “radicals.” They are orthographic components of a character, and can be classified into “semantic radicals” and “phonetic radicals.” Take 河 (river) as an example. The left hand side of this character 河 is a semantic radical, and suggests that the meaning of this character has something to do with water. As for the right hand side of this character 可, it is a phonetic radical, and indicates the pronunciation of 河. However, “[o]wing to the historical evolution of Chinese phonology, in modern Chinese, only about 26% of
these phonetic radicals can be considered reliable cues to pronunciation” (p. 98).

The results of this study suggest that radical knowledge, radical perception skills, and radical knowledge application skills do not develop synchronously and that there is a moderate positive association between the development of radical knowledge application skills and Chinese word acquisition.

**Conclusion**

Notes

1 The 29 native speakers of English were undergraduate students enrolled in an introductory course in linguistics at a US university. As for the L2 learners of English, there were 58 participants in total. They were adult learners of English (with different L1 backgrounds). The L2 participants consisted of those who attended an ESL program or university courses at two US universities and those who took EFL certificate preparation courses in Europe. They were divided into two groups (i.e., advanced [n = 29] and intermediate [n = 29]) according to their proficiency in English (Zareva, 2007, p. 131).

2 To be concrete, the participants of all the three groups (i.e., native speakers of English, advanced-level learners of English, and intermediate-level learners of English) gave “more paradigmatic than syntagmatic responses to the words they were familiar with” (Zareva, 2007, p. 141).

3 Alderson’s (2007) study investigated whether subjective frequency estimates of English words might be a reliable substitute for objective word frequency counts derived from corpora.

4 Barcroft (2007a) explored the benefits of providing opportunities for target-word retrieval during picture-based intentional second language vocabulary learning.

The participants of the study were 24 L2 Spanish students (second semester) at a university in the Midwestern United States (p. 42, p. 46).

5 A study by Barcroft (2007b) examined the effects of (a) writing (copying) target words and (b) writing (copying) word fragments on intentional second language vocabulary learning.

6 Barcroft’s (2007c) study examines how the presence of real versus unreal words in sentences affects the ability of native English speakers to make accurate grammaticality judgements.

7 Barker (2007) attempted to propose a new approach to the analysis of cost and benefit in vocabulary selection.

8 A cross-linguistic study by Bartning and Hammarberg (2007) investigates the functions of the following high-frequency collocations: French c’est and Swedish det är. (Both of them mean “it is.”)

9 A study by Boers, Eyckmans, and Stengers (2007) explored whether etymological knowledge of unfamiliar idioms could help learners understand their meanings and whether the etymological knowledge could offer learners cues for usage restrictions with regard to register.

10 Borer (2007) examined the effects of vocalizations involving three cognitive processing depths (repetition, manipulation, and generation) on the retention of word knowledge.
The participants of this study were eight adult EAP learners in a continuing education class. Their ages ranged from 31 to 48 years, and they came from China, Pakistan, Bulgaria, Egypt, and Iran (pp. 279-280).

The following three research questions were addressed in the study (p. 279): (a) Are word-focused utterances that exhibit elaborative processing associated with better short- and long-term retention on written tests?; (b) How might deeper processing facilitate matches between task-completion utterances and test responses?; and (c) Which learning condition, solitary or collaborative, results in deeper processing and greater retention?

11Bruton (2007a) attempts to examine and evaluate vocabulary testing instruments that have been used in second language vocabulary research.

12Bruton's (2007b) study assessed the amount of vocabulary items that were acquired by looking them up in a short text translation task (from L1 Spanish into L2 English). This task was completed collaboratively by an EFL class in a Spanish secondary school.

13What is the average vocabulary gain per student from an L1-to-L2 translation task? This is the main research question addressed in a study by Bruton (2007c). The participants of this study were thirteen intermediate-level EFL students in a state secondary school in Spain.


15Bush (2007) discusses the integration of culture with the teaching of vocabulary, and describes an approach for the categorization and use of pictures in the classroom.

16Chang (2007) attempted to examine the effect of “learning vocabulary items before taking listening comprehension tests” on listening comprehension test performance.

17Corrigan (2007) examines the affective dimensions of deep vocabulary knowledge to determine whether they influence the vocabulary choices of native speakers of American-English in describing interpersonal interactions.


19Gardner (2007) argues that vocabulary researchers should endeavor to improve the validity of the construct of “word” in applied corpus-based vocabulary research.

20Gardner and Davies (2007) analyzed the British National Corpus with a view to identifying the highest frequency phrasal verb constructions in English.

21Grant (2007) compares the frequency of spoken figurative idioms identified in two sources of spoken American English with their frequency in the spoken part of the British National Corpus (BNC). In addition, the researcher also compares the frequency of figurative
idioms identified as frequent in two British idiom dictionaries with their frequency in the BNC.

22 Hyland and Tse (2007) examine the notion of an academic vocabulary by analyzing the distribution of 570 word families in the widely used Academic Word List (Coxhead, 2000) in a corpus of 3.3 million words from a variety of genres and disciplines.

23 McCarthy and Jarvis (2007) are interested in issues concerned with lexical diversity (LD), which can be described as “the range and variety of vocabulary deployed in a text by either a speaker or a writer” (p. 459). LD is “a phenomenon that is measured by a wide variety of indices, each of which offers a specific, verifiable, and objective score of lexical deployment” (p. 460).

Although a reliable index of LD “has remained stubbornly elusive for over 60 years” (p. 459), an LD measurement instrument known as vocd “appears to be steadily becoming the LD index of choice for researchers and students alike” (p. 461).

In their paper McCarthy and Jarvis (2007) evaluate this new LD measurement instrument (i.e., vocd) both theoretically and empirically.

24 Mori, Sato, and Shimizu (2007) examined the relationship between how learners of Japanese as a second language perceive the learning of kanji and their ability to learn novel kanji words.


27 Pulido (2007) examined the effects of topic familiarity and passage sight vocabulary on second language lexical inferencing and retention through reading. The participants of the study were thirty-five adult learners of Spanish as a second language.

28 Shiotsu and Weir (2007) were interested in clarifying the relative significance of syntactic knowledge and vocabulary breadth knowledge in the prediction of reading comprehension test performance. In their paper the results of three structural-equation-modelling studies were described.

29 Shirato and Stapleton (2007) attempted to compare English vocabulary in a spoken corpus provided by 117 Japanese EFL learners with that in the British National Corpus.

30 Steinel, Hulstijn, and Steinel (2007) examined the effects of the following four independent variables on second language idiom learning in a paired-associate paradigm: (a)
direction of learning, (b) direction of testing, (c) idiom imageability, and (d) idiom transparency.

31 Webb (2007a) examined the effects of a single glossed sentence context on five aspects of vocabulary knowledge: (a) orthography, (b) paradigmatic association, (c) syntagmatic association, (d) grammatical functions, and (e) meaning and form.

32 Webb (2007b) examined the effects of repetition on second language vocabulary learning. Specifically, his study was designed to investigate the effects of repetition (1, 3, 7, and 10 encounters) on knowledge of orthography, syntax, meaning and form, association, and grammatical functions. The participants of this study were 121 learners of English as a foreign language in Fukuoka, Japan. They were selected from four second-year EFL classes at Kyushu University.

33 Webb (2007c) examined the effects of synonymy on second language vocabulary learning. The participants of this study were 84 Japanese learners of English as a foreign language from two first-year classes at a university in Fukuoka, Japan.

34 Winke and Abuhl (2007) conducted a case study with a view to identifying the vocabulary learning strategies used by nine learners of Chinese as a foreign language. When analyzing the obtained data, the researchers devised and used the following three categories: (a) input-based strategies, (b) output-based strategies, and (c) cognition-based strategies.

35 Xing and Fulcher (2007) were interested in Version A and Version B of the Vocabulary Levels Test at the 5000 word level, and attempted to examine whether the two versions were parallel.

References


Borer, L. (2007). Depth of processing in private and social speech: Its role in the retention of word knowledge by adult EAP learners. The Canadian Modern Language Review, 64, 273-300. (NB: In the case of the online version of this article, the page numbers are as follows: pp. 269-296. In other words, the page numbers of the online version are different from those of the print version. [Retrieved on September 15, 2008, from http://utpjournals.metapress.com/content/93w061626p6001l1/fulltext.pdf.])


