FOCUS ON THE FORMS: FORM RECOGNITION PRACTICE IN CHINESE VOCABULARY LEARNING

Michael Harrington *University of Queensland*
   Michael Harrington[i] is a senior lecturer in second language acquisition in the Applied Linguistics program in the School of Languages Comparative Cultural Studies at the University of Queensland. He has research interests in applied psycholinguistics, lexical development and language processing.
   m.harrington@uq.edu.au
Wenying *Jiang University of Queensland*
   Wenying Jiang is a lecturer in Chinese in the Chinese program in the School of Languages Comparative Cultural Studies at the University of Queensland. Her research interests include second language acquisition of Chinese, Chinese pedagogy, technology–enhanced language learning and intercultural communication.
   w.jiang2@uq.edu.au

This study examines the effect of recognition-based retrieval practice on vocabulary learning in a university Chinese class. Students (N=26) were given practice retrieving new vocabulary (single or two-character words) in a series of simple form recognition tests administered over four weeks. The test sets consisted of target vocabulary that appeared in the previous week's lesson and distracter items drawn from upcoming vocabulary. Tests were group-administered via PowerPoint and students used a checklist response to indicate whether a given item had appeared in the previous week's material. Responses relied on episodic knowledge of previous exposure and required no processing of semantic information. Students were able to reliably identify the target items in the retrieval task with performance on these items being found superior to that for supplementary list control words on mid-term and final vocabulary tests. The findings indicate that a focus on word forms can have a measurable effect on vocabulary learning in the classroom and underscores the efficacy of retrieval-based testing (the *testing effect*, Barcroft, 2007; Roediger & Karpicke, 2006) in facilitating vocabulary learning. The implications for recognition-based retrieval practice in vocabulary instruction in the Chinese classroom are discussed.

KEY WORDS: Vocabulary learning, word retrieval, Chinese language, classroom practice, testing effect, generation effect

INTRODUCTION

Learning vocabulary involves learning a word form, its meaning and usage (Meara, 1996). Specific attention to the word form itself, phonological or orthographic, is typically limited to beginning learners at the initial stages of learning. Once the student is able to pronounce or write a new word the emphasis shifts to learning the meanings and conditions of use linking to that form. This is a complex and ultimately open-ended task (Jiang, 2002; Nation, 2001).

Fluent use requires the knowledge of form, its meaning, as well as the capacity to retrieve (or access) these elements quickly and appropriately for the specific context, all under the pressure of real time communication needs. The immediate accurate retrieval of the intended meaning is the gateway—for the learner, often a bottleneck—to fluent comprehension production (Koda, 2005).

This capacity, or what Chapelle (2006) refers to as *word ability,* emerges with use. Each time a word is encountered the form must first be recognised and then linked to the appropriate meaning in the learner's mental lexicon. The use of new words is usually effortful and tentative as the learner searches for the appropriate form-meaning match (Jiang, 2002; Nation, 2001). Repeated retrieval of the word in various contexts increases the strength of the mental representation, ideally to the point where the word can be retrieved automatically (Nassaji, 2003; Segalowitz, 2010). In this study we will focus on the development of this retrieval capacity in the classroom.

Retrieval practice has been shown to facilitate learning. Retrieval is the process of accessing the mental representation of a word or larger knowledge element (e.g. collocation) in response to some type of external cue or stimulus. The recall of an item from memory will strengthen memory of the item itself as well as the links with its associated elements, thus facilitating long term retention and easier future access (Karpicke & Roediger, 2007; Morris & Fritz, 2000). The effectiveness of the humble flashcard as a learning tool is due to the systematic repeated retrieval of the word forms as the learner works through the stack (Mondria & Mondria-De Vries, 1994). Research on what is termed the *testing effect* has shown that test-driven retrieval of material from memory can significantly affect how the material is stored and maintained in the learner's mind, especially at the initial test (Roediger & Karpicke, 2006), and can substantially boost the long-term retention of new material (Carpenter & DeLosh, 2005; Karpicke & Roediger, 2007). It has also been suggested that the mere attempt at retrieval may have a positive effect on learning, regardless of whether test performance was successful (Richland, Kornell & Kao, 2009). These findings have led some to advocate more widespread use in the classroom of tests involving retrieval as a tool for learning, rather than just as a means for assessment (Hinze & Wiley, 2011; McDaniel, Anderson, Derbish & Morrisette, 2007).

Although the role of retrieval on learning has been established in the cognitive psychology literature, its effect on second language (L2) vocabulary development has been addressed in only a handful of studies. Retrieval practice on learned words has been shown to promote greater lexical retrieval speed, which in turn facilitates written production (Snellings, Van Gelderen & De Glopper, 2002), as well as automaticity in word comprehension (Akamatsu, 2008; Fukkink, Hulstijn & Simis, 2005). By developing fluency in these low level lexical processes, the individual is able to devote more resources to higher level text processes. The effect of practice on learned material more generally has been well established (see Segalowitz,

2010), but less is understood about the potential contribution retrieval practice can make to learning new vocabulary in the L2. One of the few studies to address this issue is Barcroft (2007), who examined the effect of retrieval practice in teaching vocabulary in Spanish as a foreign language in the United States. In the study learners were taught new Spanish words using picture-word combinations. After an initial presentation phase the items were either studied as a list of picture-word combinations or in a generation-based retrieval condition. In the latter they were presented with the pictures of the newly learned items and asked to mentally retrieve the appropriate word. A 6-second delay was given between the presentation of the picture and the accompanying word to provide the learners time to generate the word mentally. Results showed that words in the retrieval condition were remembered better than those in the list study condition at both immediate and delayed testing. The results replicated earlier first language (L1) research that showed positive learning outcomes for the use of word retrieval practice involving the generation of the appropriate word in response to a cue, e.g., a translation counterpart (Royer, 1973) or a non-word associate (McNamara & Healy, 1995).

Examined here is the effectiveness of a recognition-based retrieval task in learning Chinese words that are composed of character or characters. Recognition-based retrieval involves accessing the mental entry based on the presentation of the visual form of the word itself, in contrast to other cues such as translation counterpart, non-word associate such as a picture, etc. The treatment consists of a series of simple form recognition tests in which learners are presented with newly introduced Chinese words and judge whether the presented word was introduced in the previous lesson. The learners will simply be required to match the presented word forms with those available in the learner's memory. There is no need for any semantic processing as the task is assumed to rely primarily on episodic memory links to the prior learning event itself, and not to any meaning representations. Learners are thus encouraged to focus on the word form (not the meaning) and respond as quickly as possible, unlike Barcroft (2007) where learners were given six seconds to actively process each target word.

The advantage of generation-based retrieval tasks in learning new words is attributed to the complexity of the encoding required to perform the task (McNamara & Healy, 1995; Slamecka & Graf, 1978). In Barcroft (2007) learners were presented with a picture cue and then required to come up with the appropriate word in the absence of linguistic or other contextual cues. The activation of the semantic networks responsible for accessing the target word requires more effort than needed for, say, simply studying a paired word list, and the retrieval experience itself will further strengthen those networks for subsequent access. Traditional 'depth of processing' approaches stipulate that the more demanding the task, the more likely it is that learning will occur (e.g., Hulstijn & Laufers' [2001] involvement load hypotheses). Accordingly, one might question whether the positive effects arising from generation-based retrieval practice will be evident in a a less-demanding recognition retrieval task like the one proposed here. Conversely, as Barcroft points out, the nature of the retrieval

task may be less important than the opportunity to access the word form as stored in memory (Barcroft, 2007, p. 49). In this study we will see if retrieval practice with no explicit meaning component can affect vocabulary learning in a coursework setting.

Whether this type of form recognition practice can aid vocabulary learning is an empirical question. We believe there are several reasons why the question is well worth examining. Learning L2 vocabulary has a significant form component as it involves mapping a novel L2 form (phoneme or grapheme) onto a meaning representation that draws heavily on existing L1 knowledge. The recognition-based retrieval processes examined here focus the learner's attention on the form at early stages of learning, that is, when the form is still novel for the learner. To what extent early focus on the form can affect vocabulary learning has yet to be examined. The focus on form may be particularly useful in learning Chinese characters, which involve more complex visual (form) learning than alphabet-based languages (Everson, 1998; Wang & Koda, 2005). The task requires a quick yes/no judgment which promotes direct retrieval and discourages reflective strategic processing by the learner. The short individual trials will also permit a much greater number of words to be tested. This is in contrast to the more time-consuming tasks used in generation-based retrieval tasks like Barcroft (2007). The use of larger item sets and/or multiple repetitions is a potential benefit for instruction in character-based languages like Chinese or Japanese, where the attainment of proficiency requires— ideally—the acquisition of thousands of characters while instructional resources are limited. Finally, given the relatively speeded nature of the test task, recognition-based retrieval tasks provide practice in lexical retrieval in initial learning (Snellings et al., 2002). We are aware of no other research that has addressed the effectiveness of form-focused retrieval processes on L2 learning outcomes, particularly in non-alphabet based languages.

In the study vocabulary learning was assessed using a translation test that required the learner to give the Chinese equivalent of an English word. Previous studies used the same format of form-meaning mapping in treatment and testing to assess learning outcomes as all involved some kind of meaning manipulation (Barcroft, 2002; McNamara & Healy, 1995; Royer, 1973). Here the form-based nature of the retrieval task precluded the use of the same format at test, as it allows no way to assess whether the learner has learned the meaning. The difference between treatment and test means that any learning effect will not be attributable to overlap between the two, or *transfer-appropriate processing* (Lightbown, 2007), but rather will implicate retrieval in other aspects of word learning.

THE STUDY

This is an exploratory study that assesses the effect of a simple recognition-based retrieval activity (the treatment) on Chinese vocabulary learning. Similar to Barcroft (2007) the research is being operationalized in an instructional technique that can be readily implemented in the L2 classroom. The treatment activities and the vocabulary tests were part

of the regular course activities and marks on the vocabulary tests counted toward the semester grade. The second author was the course instructor who administered all the treatments and tests.

THE SETTING AND PARTICIPANTS

The study was carried out in a second-year Chinese class at an Australian university. The students were in their third semester of enrolment. Data were collected from 26 students, all of whom had English as a first language with no Chinese family background. The class used the textbook *Contemporary Chinese* (Wu, 2003), which covers both spoken and written components of the language courses. The class met twice a week. The treatment was done in the Monday lecture in teaching Weeks 3-6 by the instructor (the second author).

THE TREATMENT

In the text each weekly lesson was accompanied by a new vocabulary list of approximately 30 new items. The items consisted primarily of one or two character words. In order to assess the effect of form retrieval on learning, the new words were divided into three word groups. The *target* word group consisted of ten new words that were presented in class via a weekly PowerPoint presentation. The procedure, described below, required the students to actively retrieve each of the ten words.The second group of words was a weekly *supplement list* that the instructor provided to students on a sheet of paper after the PowerPoint presentation. The words in the supplement list were introduced by the instructor as being particularly useful for the lesson and students were encouraged to study them on their own outside of class. No attempt was made to assess the degree to which the students studied the supplement list. By providing a supplement list we wished to exclude the possibility that any observed advantage for the target words was due merely to the instructor treating the items in a special way. Finally, the remaining 10 words from the weekly vocabulary list comprised the thirds group of *textbook* words. These words received no additional attention from the instructor outside of the normal class activities. Words in group 3 were the control condition. Items were randomly assigned to the three groups by the instructor with attention given to ensure there was a balance of single- and double-character words. The test items for the immediate (mid-term) and delayed (final) tests were drawn from these three word groups.

The weekly recognition task was presented by PowerPoint to the entire class. The words were presented one at a time, and students were directed to check 'yes' or 'no' on a paper form as to whether the word was taught in the previous week's lesson. They were told that they did not need to know what the word meant, and only needed to judge if it was part of the previous week's vocabulary. Each week's recognition task contained the ten target words of the previous lesson, as well as ten new words that would be taught in the subsequent week. Four sets were developed and presented over four consecutive weeks. The order of presentation for each list of 20 words was randomized. The instructor ran the PowerPoint

presentations and distributed the supplement lists on four consecutive Mondays in the first part of the first semester in 2011.

In the presentation session each slide contained one word, which was left up until all the students had answered, typically taking 3-5 seconds. Students were encouraged to work quickly but not to rush. At the end of the presentation the completed checklists were collected and scored.

THE TESTS

In the week following the fourth treatment students completed an immediate post-test as part of the course's mid-semester test. Seven weeks later a delayed test was given as part of the final exam. There was no treatment presentation between the immediate and delayed tests and different items were used in all the tests. Both tests required the Chinese translation of words presented in English. The immediate post-test consisted of two sections, one requiring translation responses in Pinyin (Latin transcriptions of Chinese words) and the other requiring the students to write the translation in Chinese characters. The Pinyin section was given to accommodate the small number of students (not included in the sample) who were only taking the spoken section, but was completed by all the students. Only the Pinyin results of the students in both the spoken and written sections (N=26) are presented. These students also completed the character-based section of the immediate and delayed tests reported below. Both the Pinyin and character response sections in the immediate post-test consisted of fifteen items, including five each from the target, supplement and textbook groups. The items in each group were drawn from across all four weeks of the treatment for all the tests. The final test consisted of 18 items, 6 each from the respective groups. In both tests English words were presented and the students asked to translate them into Chinese. Each item was scored as 2 points for a complete response, 1 for partial response and 0 for an incorrect or no response. A partial response was only rewarded for the two-character words when only one of the characters was correct. The second author scored the tests.

RESULTS

Of interest is the effect of word form retrieval on vocabulary learning. Unlike previous research (e.g., Barcroft, 2007; McNamara & Healy, 1995) the task made no explicit demand on the student for any kind of semantic processing; rather, it required them to accurately recall whether the form was part of the previous week's class. Prior to examining whether retrieval had an effect on learning, it was first necessary to show that the learners were actually retrieving the words during the treatment. Table 1 presents the results for the four weekly treatment sessions. Students were able to accurately discriminate between target and new word items over 85% of the time across the four weekly treatment sessions.

Table 1. Response accuracy for judging whether presented word was target word from the previous week or a new word.

Treatment		Target word		New word	
		M	*SD*	*M*	*SD*
Week 1		91%	13%	88%	16%
Week 2		84%	18%	86%	16%
Week 3		91%	12%	81%	12%
Week 4		83%	16%	87%	15%
	Average	87%	15%	86%	15%

The main question is whether this retrieval helps vocabulary learning. Three measures of vocabulary translation knowledge were used. The immediate test used two response formats, a Pinyin section[ii] (15 items) and a Chinese character section (15 items). The delayed test consisted of a single Chinese character section (18 items). All three tests were reliable measures of vocabulary, as reflected in Cronbach's alpha: immediate Pinyin, $\alpha = .83$; immediate character, $\alpha = .78$; delayed character, $\alpha = .85$.

In the immediate Pinyin test students scored a higher mean percentage of correct responses on the target items ($M = 86.88$, $SD = 12.30$) than on the supplement list ($M = 78.12$, $SD = 23.06$) or textbook items ($M = 75.31$, $SD = 22.57$). The mean differences by word group are given in Figure 1. A one-way ANOVA with word group (Target x Supplement x Textbook) as the within-group independent variable showed a statistically significant difference between the means, $F(2, 50) = 6.27$, $p = .001$, with a very modest effect size (partial eta^2) of .17. Pairwise comparisons showed that performance on the target items were better than both the supplement ($p = .017$) and the textbook items ($p = .005$). There was no statistically significant difference between the means for the supplement and textbook items.
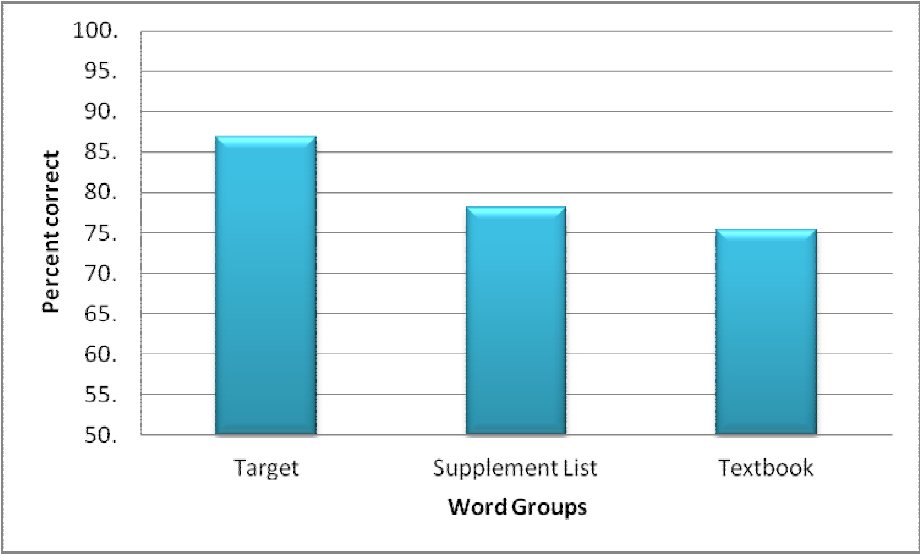
Figure 1. Performance on immediate vocabulary test (Pinyin response) by word list type: Target, Supplement List Textbook (N=26).

The Pinyin results support the notion that retrieval practice helps learning. Responses for the Chinese character section of the immediate test showed no advantage for the target items. See Table 2 for the means and standard deviations for the 22 students who completed both the immediate and delayed tests. However, in the delayed test, completed 7 weeks after the treatments, a positive effect for retrieval did emerge. Here the target words mean was 30% higher than either the supplement list or textbook item means. The mean differences were tested for statistical significance using a 2 x 3 ANOVA in which test time (Immediate x Delayed) and word group (Target x Supplement x Textbook) were within-group independent variables. The mean difference for test time was not statistically significant, $F(1,21) = 1.74$, not significant. The overall mean difference for word group was significant: $F(2,42) = 9.36$, $p = .000$, with an effect size (partial eta^2) of .32. The effect size is moderately strong, accounting for nearly a third of the variance explained (Field, 2009).

Table 2. Descriptive statistics (Means and Standard Deviations) for immediate delayed tests by list types for the Chinese character responses.

List Type	Immediate Test		Delayed Test		Overall Type Average	
	M	*SD*	*M*	*SD*	*M*	*SD*
Target	72.81	22.81	77.21	27.60	75.00	25.20
Supplement	66.36	23.61	57.25	23.07	61.80	23.34
Textbook	70.40	20.51	56.44	24.79	63.43	22.65
Overall average	69.85	22.31	63.63	25.15		

N = 22

There was also an interaction between test time and word group, $F(2, 42) = 18.22$, $p = .000$, partial $eta^2 = .46$. As indicated in Table 2, the interaction effect is the result of differences in word group performance on the delayed test. The interaction effect is illustrated in Figure 2. In contrast to the immediate character test, where there were no significant mean differences for the three word groups, at the delayed test the mean for the target items was markedly better than that of either the supplement list or textbook words. The effect is strong, as reflected in the large partial eta^2. Mean performance on the target items improved from immediate to delayed test, but a follow-up paired t-test showed the difference was not significant at $p < .05$, $t(21) = -1.056$, $p = .303$. In contrast, the decrement in performance for the supplement list approached significance, $t(21) = 1.81$ $p = .084$ and the mean difference in textbook items was significant, $t(21) = 2.38$, p = .024.
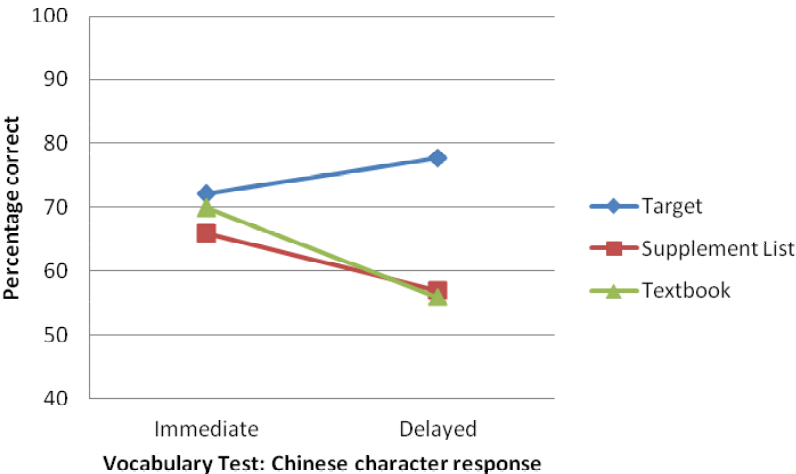

Figure 2. Performance on immediate and delayed Chinese character vocabulary tests by word group: Target, supplement list and textbook (N=22).

The findings support the notion that a simple recognition task can have a positive effect on vocabulary learning in a classroom setting. At the immediate test there was an advantage for the retrieval practice items in the Pinyin but not in the Chinese character responses. At the delayed test this advantage did become apparent. These results will be discussed in the following section, along with implications for teaching and future research.

DISCUSSION

Recognition-based retrieval practice had a positive effect on course vocabulary learning outcomes relative to the supplement list and textbook items. The effect was evident at both the immediate and delayed tests, although only for the Pinyin section in the immediate test. The advantage of Pinyin over Chinese characters as a response format may reflect the relatively lesser processing demands that Pinyin, as a Latin script, places on English-background learners. At the early stages of learning the task for the learner is to link the word form (phonological and orthographic) with its meaning, a process that is often facilitated for English background learners by the use of Pinyin (Lee & Kalyuga, 2011). The Pinyin results may suggest a staged process of Chinese vocabulary learning wherein knowledge of phonological form and meaning is learned first (as reflected in performance in the transcriptive Pinyin format), and the ability to produce the Chinese character later. This is a question for further research, as the items were not counter-balanced across the Pinyin and Chinese characters in the immediate test. Due to the course structure a Pinyin section was not administered at the delayed test.

This was an exploratory study carried out in the context of a university Chinese course and the results must be interpreted with caution. Obvious limitations arise from the small number of items used, as well as the limited treatment period. Items were randomly assigned to the three word groups but the possibility exists that the learners had some prior knowledge of the words used. Further research needs to test the effect across a larger number of items whose presentation (counterbalancing, etc.) is done more systematically. Likewise, there was no control for exposure to the target items outside of the treatment and these may have had more salience in the classroom activities. Also, there was no attempt to control for study of the supplement list. It is possible that the students largely ignored the supplement list. Thus the results cannot be interpreted in terms of retrieval versus list study (Roediger & Karpicke, 2006) but rather as evidence for the value of this type of retrieval activity in the context of a typical Chinese course. Future research should examine more directly how list-learning compares to retrieval stimulated by PowerPoint presentation.

Still, given the modest scope of the study, we believe the results show that form retrieval practice can facilitate Chinese vocabulary learning, and that more work in the area is clearly warranted. While successful vocabulary learning ultimately depends on linking form with meaning, there is a significant element of perceptual form learning involved in learning

Chinese (Tan & Perfetti, 1999). The task introduced here prompts the learner's retrieval of the form using as a retrieval cue the memory of recent exposure to the form. Thus the task is similar to first language (L1) retrieval studies (e.g., Slamecka & Graf, 1978) where "learning" is defined in terms of better task performance on known words due to a previous presentation condition. What is learned is the particular combination unique to the presentation episode. In this study episodic cues were used to retrieve newly-learned material, and the role these cues might play in bootstrapping learning of new L2 words deserves more attention. The superior performance for the target items may also be due to the PowerPoint presentation format, which involved large screen presentation to the class. It is possible that this type of input enhancement increased the salience of the items in such a way as to facilitate learning (Peters, Hulstijn, Sercu & Lutjeharms, 2009). Finally, while embedding the study in a classroom resulted in several compromises in research design, there was also an advantage in that the learning process could be examined in a context where participants had a higher stake in the outcomes (in the form of the course mark) than might otherwise be the case, e.g., when participating as volunteers or paid participants.

IMPLICATIONS FOR CHINESE VOCABULARY INSTRUCTION

The development of lexical retrieval skills has not been a focus in the L2 classroom, where the time available for vocabulary instruction is typically devoted to introducing new words and the development of fluency left to the individual learner (Nation, 2001). We believe the findings of this study, along with those of Royer (1973), McNamara and Healy (1995) and Barcroft (2007) indicate that the introduction of retrieval practice can have a positive effect on vocabulary learning in the classroom, without making undue demands on the instructor and learners' time. The PowerPoint activity introduced here takes a very limited amount of time to develop and administer in the classroom, and potentially allows a large number of items to be covered. This is particularly meaningful for instructors of Chinese as a foreign language because learning to recognize and write the script is widely considered the most difficult part in learning the language, due in part to the complex structure and perceptual diversity of Chinese characters (Everson, 1998). The process lends itself to the rapid presentation of a large number of words in a very short period of time, which in turn can be developed into a systematic program of presentation and review of these forms. Research is underway to further assess the pedagogical applications of the technique as part of regular classroom practice.

FUTURE RESEARCH

In closing, we strongly endorse Barcroft's (2007, p. 41) plea for more research in this still under-researched area. Further studies are needed to establish the positive effects of retrieval on vocabulary learning and to broaden our understanding of how these processes interact with different vocabulary teaching learning activities inside and outside the classroom.

The study here provides the basis for a more broadly scoped research program examining retrieval-based testing as a tool for Chinese vocabulary development. Future research should examine how the frequency spacing of word presentation affects learning outcomes (Rott, 2007), and whether unsuccessful performance at test can lead to subsequent learning (Richland et al., 2009). The relationship between the effectiveness of retrieval practice in vocabulary learning and level of learner proficiency also needs further attention. The use of converging evidence will provide a more complete picture of the role that retrieval practice can play in L2 vocabulary learning.

REFERENCES

Akamatsu, N. (2008). The effects of training on automatization of word recognition in English as a foreign language. *Applied Psycholinguistics, 29*, 2, 175-193.

Barcroft, J. (2002). Semantic and structural elaboration in L2 lexical acquisition. *Language Learning*, *52*(3), 323–363.

Barcroft, J. (2007). Effects of opportunities for word retrieval during second language vocabulary learning. *Language Learning, 57*(1), 35-56.

Carpenter, S. K. & DeLosh, E. L. (2005). Application of the testing spacing effect to name learning. *Applied Cognitive Psychology, 19*, 619-636.

Chapelle, C. (2006). L2 vocabulary acquisition theory: The role of inference, dependability and generalizability in assessment. In M. Chalhoub-Deville, C. A. Chapelle & P. Duff (Eds.) *Inference generalizability in applied linguistics: multiple perspectives* (pp. 47-64). Amsterdam: John Benjamins.

Everson, M. (1998). Word recognition among learners of Chinese as a foreign language: Investigating the relationship between naming and knowing. *Modern Language Journal, 82*(2), 194-204.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: Sage.

Fukkink, R. G., Hulstijn, J. & Simis, A. (2005). Does training in second-language word recognition affect reading comprehension? An experimental study. *Modern Language Journal, 89*(1), 54-75.

Hinze, S. R. & Wiley, J. (2011). Testing the limits of testing effects using completion tests. *Memory*, 19(3), 290-304.

Hulstijn, J. H. & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, *51*(4), 539-558.

Jiang, N. (2002). Form-meaning mapping in vocabulary acquisition in a second language. *Studies in Second Language Acquisition*, *24*(4), 617-637.

Karpicke, J. D. & Roediger, H. L., III. (2007). Repeated retrieval during learning is the key to long-term retention. J*ournal of Memory Language,* 57(2), 151-162.

Koda, K. (2005). *Insights into second language reading*: *A cross-linguistic approach*. NY: Cambridge University Press.

Lee, C. H. & Kalyuga, S. (2011). Effectiveness of different Pinyin presentation formats in learning Chinese characters: A cognitive load perspective. *Language Learning, 61*(4), 1099-1118.

Lightbown, P. (2007). Transfer appropriate processing as a model for second language acquisition. In Z. Han (Ed.), *Understanding second language process* (pp. 27-44). Clevedon: Multilingual Matters.

McDaniel, M. A., Anderson, J. L., Derbish, M. H. & Morrisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology, 19*(3), 371-385.

McNamara, D. S. & Healy, A. F. (1995). A generation advantage for multiplication skill training and nonword vocabulary acquisition. In A. F. Healy & L. E. Bourne, Jr. (Eds.), *Learning memory of knowledge skills: Durability and specificity* (pp. 132-169). Thousand Oaks, CA: Sage.

Meara, P. (1996*)*. The dimensions of lexical competence. In Brown, G., Malmkjaer, K., J. Williams, editors, *Performance competence in second language acquisition,* (pp. 35-53). Cambridge: Cambridge University Press*,*

Mondria, J. A. & Mondria-De Vries, S. (1994). Efficiently memorizing words with the help of word cards 'hand computer'. *System, 22*(1), 47-57.

Morris, P. E. & Fritz, C. O. (2000). The name game: Using retrieval practice to improve the learning of names. *Journal of Experimental Psychology: Applied*, *6*(2), 124-129.

Nassaji, H. (2003). Higher-level and lower-level text processing skills in advanced ESL reading comprehension. *Modern Language Journal, 87*(2), 261-276.

Nation, I. S. P. (2001). *Learning vocabulary in another language.* Cambridge: Cambridge University Press.

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*(2), 131-151.

Richland, L. E., Kornell, N. & Kao, L. S. (2009). The pretesting effect: do unsuccessful retrieval attempts enhance learning? *Journal of Experimental Psychology-Applied*, *15*(3), 243-257.

Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*(3), 249-255.

Rott, S. (2007). The effect of frequency of input-enhancements on word learning text comprehension. *Language Learning, 57*(2), 165-199.

Royer, J. M. (1973). Memory effects for test-like events during acquisition of foreign language vocabulary. *Psychological Reports*, *32*, 195-198.

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. New York: Routledge.

Slamecka, N. J. & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology*, *4*(6), 592–604.

Snellings, P., Van Gelderen, A. & De Glopper, K. (2002). Lexical retrieval: An aspect of fluent second language production that can be enhanced. *Language Learning, 52*(4), 723-754.

Tan, L. & Perfetti, C. (1999). Phonological activation in visual identification of Chinese two-character words. *Journal of Experimental Psychology: Learning, Memory, Cognition, 25*(2), 382-393.

Wang, M. & Koda, K. (2005). Commonalities and differences in word identification skills among learners of English as a second language. *Language Learning 55*(1), 71-98.

Wu, Z. (2003). *Contemporary Chinese*. Beijing: Sinolingua.

ENDNOTES

i   Contact: Michael Harrington, SLCCS 32-429, University of Queensland, St Lucia, QLD 4072, Australia m.harrington@uq.edu.au.

ii   Pinyin is the official phonetic system for transcribing the sound of Chinese characters into Latin script in China, Taiwan and Singapore.