

7

Assessing Incidental Vocabulary Learning by Chinese EFL Learners: A Test of the Involvement Load Hypothesis

Chanchan Tang and Jeanine Treffers-Daller

7.1 Introduction

One of the most challenging tasks for learners of a Second Language (L2 learners) consists in developing a vocabulary large enough to be able to read and write fluently and take part in conversations on a range of topics. According to Adolphs and Schmitt (2003) learners need 2000–3000 of the most frequent English word families to be able to take part in everyday conversations, whilst they need 5000 word families to begin to read authentic texts (Schmitt, 2007). For unassisted comprehension of written texts it is assumed learners need around 8000–9000 word families, and a vocabulary of 6000–7000 word families for spoken text (Nation, 2006). Many researchers have indicated that L2 learners worry about the formidable task of learning thousands of words (see for example Jones, 1995; Kim, 2008; Lawson & Hogden, 1996), particularly in contexts where learners have few opportunities to go to the country of the target language and/or have little knowledge about the target language culture, as is the case for many Chinese learners of English (Shao, 2014). For teachers it is equally challenging to find ways to help students acquire a wide range of words within the limited class time. Researchers can help address this issue by providing evidence regarding the effectiveness of different approaches to vocabulary learning and teaching.

There are of course many different ways to build up a vocabulary. Hunt and Beglar (1998) outlined three approaches to enhance vocabulary learning—incidental learning, explicit instruction, and independent strategy development. Nation (1990) has shown that intentional vocabulary learning, in particular instruction, does aid in the learning

of words, especially in the earlier stages of language learning. However, because of the limited amount of time that is available in class, only a few words can be taught by direct instruction (Nagy & Herman, 1987). Instead, the large majority of words are assumed to be acquired while the learner is reading a text or listening to a message, and focuses on the content instead of on learning words, that is through incidental vocabulary learning (Hulstijn, 2003). However, more recent work in this field shows that vocabulary gains from reading or writing are very limited. According to Nation and Wang (1999) at least ten exposures are needed if learners are to be successful at learning unknown words. More recently, Pellicer-Sanchez (2012) has shown that effects of frequency of exposure of new words are significant from three to five repetitions onwards, whilst unknown words that are repeated eight times begin to be read like known words. Because L2 learners often have limited exposure to target language input, and their input is generally limited to classroom contexts, they are unlikely to make large vocabulary gains by repeated exposure alone.

As there is considerable evidence that vocabulary take-up from reading is rather limited, researchers need to focus their attention on how incidental vocabulary learning can be promoted in the process of reading, for example by encouraging learners to use dictionaries, or providing glosses or asking learners to engage in different post-reading tasks (Hulstijn, Hollander & Greidanus, 1996; Ko, 2005). More research needs to be done to explain the reasons for the different degrees of effectiveness of tasks (Anderson, 1995; Joe, 1995, 1998; Paribakht & Wesche, 1997). One possibility is that the effectiveness of each task is determined by the depth of processing of vocabulary items by learners, but operationalizing depth of processing is difficult. For this purpose, Laufer and Hulstijn (2001) put forward the Involvement Load Hypothesis (ILH), which is based on the idea that incidental vocabulary can be promoted by involving learners in different post-reading tasks which require learners to engage with the words in the text in a variety of ways. Tasks differ from each other with respect to the degree of processing depth needed to carry out the task (see Section 2 for details).

Laufer and Hulstijn (2001) and Hulstijn and Laufer (2001) call for further tests of the ILH, although few researchers (Kim, 2008, 2011; Yaqubi et al., 2010) have so far responded to their invitation. The current study aims to contribute to the discussion by focusing on two points which the authors mention in their papers as particularly relevant for tests of the ILH, namely the relative importance of different components of

the ILH and the differences between the effects of input-oriented and output-oriented tasks on vocabulary learning and retention.

The participants in our study are Chinese learners of English who are at A2 level in English. So far the ILH has not been tested with students who have a relatively low level of proficiency in English. The participants in Kim's (2011) study were adult L2 learners of English, divided over two groups: one group were enrolled on a pre-university intensive course and had TOEFL scores between 470 and 520 and the other group were university students who had a TOEFL score above 520. While Kim (2011) found no interaction effects between language proficiency and task type in their study of the ILH, they also call for including a wider range of proficiency in future studies testing the ILH, and the current study therefore fills an important gap in our knowledge by focusing on the lower end of English language ability.

7.2 Incidental vocabulary acquisition

In the domain of L1 and L2 pedagogy, the term incidental vocabulary acquisition is understood to mean "learning without an intent to learn, or as the learning of one thing, e.g. vocabulary, when the learners' primary objective is to do something else, e.g. to communicate" (Schmidt, 1994, cited in Laufer & Hulstijn, 2001, p. 10). In other words, incidental vocabulary acquisition means that learners focus on understanding the meaning of spoken or written information while reading or listening and not on vocabulary learning *per se*. In such a process, new words are acquired "as a by-product of other cognitive exercises involving comprehension" (Gass, 1999, p. 319). In practical terms, incidental vocabulary learning can be operationalized as a learning process with absence of any forewarning of subsequent retention tests (Hulstijn, 2003).

Paribakht and Wesche (1997) were among the first to show that vocabulary learning can be promoted through a combination of reading and enhancement activities. They found that words practised through exercises were retained better than words for which the meaning was inferred from the context. Hence, asking learners to carry out tasks could be an effective tool for vocabulary learning as this might stimulate learners to process words more deeply. Although depth of processing remains difficult to measure objectively, it is likely that the nature of processing activities in which learners engage affects their retention of information: more elaborate processing activities will lead to better retention. Laufer and Hulstijn (2001) were the first to apply Craik and

Lockhart's (1972) depth of processing hypothesis to vocabulary learning by proposing the Involvement Load Hypothesis (ILH), which claims that each task can induce a certain amount of "involvement load", and that the effectiveness of a task is determined by the "involvement load" it induces. Put simply, the more learners engage with the words they learn (for example by focusing on the spelling, the meaning or aspects of the way the words are used), the better they will retain them.

The motivational-cognitive construct of involvement consists of three basic components: *need*, *search* and *evaluation*. *Need* is a motivational construct, concerned with the "need to achieve" (Laufer & Hulstijn, 2001, p. 14), whilst *search* and *evaluation* belong to cognitive dimensions, concerned with noticing and attending to form-meaning relationship (Schmidt, 1994, cited in Laufer & Hulstijn, 2001). *Need* refers to the motivation of learning target words and the drive to comply with task requirements. *Search* is the attempt to find the meaning of an unknown L2 word or the attempt to find a suitable L2 word form for a particular L1 concept. *Evaluation* refers to whether or not learners are required to compare the target words with other words. Tasks can of course induce these involvement factors to different degrees. For the purpose of the ILH, the authors suggest that there are three possible levels of involvement for each: none, moderate and strong. All three involvement factors may not be at work simultaneously during a reading-based task, or in other words, a task can induce any one, two, or all three of the components of involvement for each word. The involvement load of a task is defined as the combination of the three involvement factors, which can be absent or present, moderate or strong (see Section 3 for more details).

So far, few researchers have attempted to directly test the ILH, although some researchers have tested aspects of it. Yaqubi et al. (2010) are among the few who tested ILH, but they did not find that tasks that induced a higher involvement load led to higher scores on the post-test. Instead, they claim that output-oriented tasks lead to better results regardless of the degree of involvement load of the tasks. Kim (2011), on the other hand, found moderate support for the ILH in initial vocabulary learning because learners acquired words more effectively through tasks that induced a higher degree of involvement, as tested in an immediate post-test. In addition, Kim found strong support for the ILH in retention of vocabulary in a delayed post-test. Importantly, the author suggests that the effects of different tasks may not be visible immediately but only at a later stage, and calls for further investigation of the long-term effects of tasks with different involvement loads.

We aim to first of all test the central claim of the ILH that tasks with a higher involvement load will be more effective than those with a lower involvement load. Second, we hope to contribute to the discussion about the relative contribution of different components of involvement to vocabulary acquisition. There is some evidence that the different components of involvement do not have an equal impact on students' vocabulary retention. Laufer and Hulstijn (2001) suggest that *search* may be less important than the other two components, and Kim (2011, p. 125) found some evidence that "strong evaluation induces much greater involvement in processing a word than the moderate evaluation and the other two components." It is therefore important to investigate the contribution of different components of involvement in greater detail. Third, we will look into differences between input-oriented tasks and output-oriented tasks, and test Laufer and Hulstijn's claim that these two task types are equivalent as long as the involvement load of the tasks remains constant. This is relevant because Yaqubi et al. (2010) found that output-oriented tasks and input-oriented tasks with the same involvement indices were not equally effective. Our fourth research question relates to the differences between initial vocabulary learning (which is measured with an immediate post-test) and vocabulary retention (to be measured with a delayed post-test), as we wanted to find out whether a higher involvement load leads to better vocabulary retention in the longer term, as Kim (2011) suggested. In Hulstijn and Laufer (2001), one of the few studies which directly tested the ILH, the superior results of the students who were required to write compositions may well have been due to the fact that they had more time to spend on the task, as Kim (2011) points out. This is an important point that Hill and Laufer (2003) explored in great depth in a follow-up study on incidental vocabulary learning. In the current study, we will carefully control for time-on-task, to ensure that differences in vocabulary learning and retention cannot be explained by differences in time-on-task.

For a variety of reasons it is particularly difficult for learners with Chinese as their L1 to learn English words. First of all, Chinese and English belong to different word families (Sino-Tibetan and Germanic), which means there are virtually no cognates between the languages which could facilitate vocabulary learning (Larrañaga, Treffers-Daller, Tidball & Gil Ortega, 2012). Second, while learning vocabulary in another language generally involves learning a new way to map meaning onto form, this is particularly complex for learners whose L1 uses a logographic script and who need to learn words in an alphabetic script (see also Cheng & Yang, 1989, who investigated differences in

processing of characters and words). Third, many Chinese learners are not very familiar with incidental vocabulary learning, because the key vocabulary learning strategy in EFL classrooms in China and Taiwan is rote learning of vocabulary lists (Li, 2004; Smith, Kilgariff & Sommers, 2008) and there are few opportunities for learning words from meaning-focused input, that is listening and reading in the classroom (Nation, 2007). For this reason a study into incidental vocabulary learning and assessment among Chinese learners is very much needed.

7.3 Methodology

7.3.1 Participants

This experiment was conducted in a secondary vocational school in China. The participants were 230 students (male and female) in six intact classes in the second term of Grade One. These six classes were used as six groups in the experiment and each carried out one of the six different tasks specified below. The majority (83 per cent) of participants were aged 18, 14 per cent were 17 years old and 3 per cent were 16 years old. At the time of data collection they had learned English for three years in junior school. The participants' proficiency level was roughly equivalent to the A2 level on the Common European Framework of Reference for Languages, which provides widely-used guidelines used to describe achievements of learners of foreign languages (Council of Europe, 2001). This was established by comparing students' performance on their high school examinations to the CEFR descriptors. Since they had been allocated to classes based on the results of the high school exams, the overall English level of the participants in these six classes was quite similar. The first author also obtained access to the students' most recent mid-term English examinations, the average scores of which were within a range of three points (the total score of the paper was 100 points), which lends support to our assumption that students were at similar levels of language proficiency. The experiment was carried out towards the end of the second term in Grade One, a period during which the participants were moving to Grade Two.

7.3.2 Choice of reading passage and target words

Since the participants had almost completed the second term of Grade One and would move on to Grade Two after the summer holiday, the text was taken from the textbook for the first term of Grade Two (see the appendix to this chapter). The first author selected the passage so that its level of difficulty was appropriate but still challenging for the

learners, to ensure the participants could understand the general meaning and at the same time acquire new vocabulary items incidentally. To examine the difficulty of the vocabulary in each text, several reading passages were analysed with the help of the VocabProfile program (<http://www.lex tutor.ca/vp/eng/>), which provides information about the frequency layers to which the words in a text belong. In addition, the first author consulted the teachers from the school as well as the Word Bank (vocabulary glosses) provided after the reading passage in the original textbook. On the basis of the information obtained from all these sources, the most suitable reading passage from among the first chapters of the book for Grade Two was chosen for the experiment. It is a narrative about an event that took place at an airport. The passage contains 222 words of which ten were unknown according to the students' teachers. Prior to the main data collection we also carried out a pilot study to investigate whether these were unknown. Ten students from the same level took part in this pilot. These students did not take part in the main study. The ten words were indeed unknown, except for *down*, because some students in the pilot had partial knowledge of this word: the students selected "antonym of up" as the meaning of *down*, which is possible, but not appropriate in the context. In other words, 95.5 per cent of the words in the selected passage were known, which made this a suitable but sufficiently challenging reading task.

We selected eight words as the target words in the experiment from among the ten originally chosen. The eight target words were: *airline*, *backup*, *frightening*, *kick*, *luggage*, *screen*, *spread*, and *stare at*. Two of the ten were excluded from the analysis: *down*, for reasons mentioned above and *point at* because the teachers told us that using two phrasal verbs in this study was too complex at this level. We decided to include words from different categories to ensure there would be some variation in level of difficulty among the target words. According to Ellis and Beaton (1993), part of speech is an important determinant of the learnability of words, with nouns being the easiest to learn, followed by adjectives, whilst verbs and adverbs are the most difficult to learn for L2 learners. The target words consisted of three nouns, two verbs, two adjectives and one phrasal verb, which means that the most and the least difficult word categories were included in the study. We also consulted VocabProfile to determine the frequency levels of these eight target words. Four of these (*frightening*, *kick*, *screen*, and *spread*) belonged to the 2000 word frequency band. The words *airline* and *stare at* were found to belong to the 3000 word frequency band, whilst *backup* and *luggage* were from the 4000 and the 7000 bands respectively. All ten new

words were printed in **bold** to increase the saliency of these words to participants. According to Sharwood Smith (1993), if a word is salient in the input to the learners, there is a greater chance for it to be selected and processed by the L2 learner. Input enhancement using bold typeface is one way in which input saliency can be increased.

7.3.3 Task design

We designed six tasks with different involvement loads to investigate the effects of the tasks, the amount of involvement load in each and the impact of the different components of involvement (*need*, *search* and *evaluation*) on vocabulary acquisition and retention. The degree of involvement required by the different tasks was expressed in an Involvement Index: absence of a factor was counted as 0, a modest involvement with a factor as 1, and strong involvement with factor as 2. The tasks carried out by the six groups are described below. As the aim of the study was to measure students' incidental acquisition of words from the reading passage (with or without additional activities), in line with the methodologies used in Hulstijn and Laufer (2001) and Kim (2011), students were not informed they would be tested on their knowledge of the words after reading the passages and undertaking the different tasks.

Group 1 reading only (-need, -search, -evaluation) (see appendix 7.1, Task 1)

The students involved in this task were only asked to read the given passage. The Chinese translation equivalents were provided in the text just after each new English word and there were no post-reading activities. This means the learners do not feel the need to learn the words, nor did they need to search for the translation equivalents or compare the meanings of the words to other words. This task was classified as scoring zero on the Involvement Index (0+0+0).

Group 2 reading + comprehension questions (-need, -search, -evaluation) (see appendix 7.1, Task 2)

The reading passage given to Group 2 was the same as the one used in Group 1, with the Chinese equivalent of the ten new words in the text. The difference from Task 1 was that there were comprehension questions which students had to answer after reading the text, but these were irrelevant to the target words. Since the new words were glossed in the text and they were irrelevant to the comprehension questions, the learners did not need to learn the words nor to search or evaluate

the words' meanings. Therefore, the Involvement Index of Task 2 is also 0 (0+0+0).

Group 3 reading + comprehension questions (+need, +search, -evaluation)
(see appendix 7.1, Task 3)

In this task, the comprehension questions were designed with relevance to the target words, so the participants needed to know the meaning of the target words in order to complete the task, and the factor *need* was clearly present. In addition, the ten new words were glossed in the end of the passage alphabetically, so the factor *search* was triggered. Since the glossary listed the word meaning that was relevant for the context, the *evaluation* factor is absent. So the Involvement Index of Task 3 is 2 (1+1+0).

Group 4 reading + comprehension questions (+need, -search, +evaluation)
(see appendix 7.1, Task 4)

The reading passage and comprehension questions were exactly the same as in Task 3. However, unlike Task 3, the new words were glossed in the margin rather than at the end, so there was no need to search for the meaning. Moreover, because different meanings of one word were provided, the participants needed to compare between them in order to choose one that was most suitable for the given context. For this reason, a modest *evaluation* was triggered. So the Involvement Index of Task 4 is 2 (1+0+1).

Group 5 reading + comprehension questions (+need, +search, +evaluation)
(see appendix 7.1, Task 5)

Task 5 was designed to involve all three involvement load components. It shared the same reading passage and comprehension questions as Tasks 3 and 4, so a moderate *need* was present. Students also had to evaluate the word meanings, as in Task 4, so the *evaluation* component was induced. The difference was that the glossary was located at the end of the passage according to alphabet instead of being in the margin, so the participants needed to search the word meaning. Therefore, the Involvement Index of Task 5 is higher, namely 3 (1+1+1).

Group 6 reading + sentence production (+need, -search, ++evaluation)
(see appendix 7.1, Task 6)

Task 6 shared the same reading passage with the previous tasks and also shared the same kind of glossary with Task 4 (with the word meaning

glossed in the margin and several options to choose from). The post-reading activity for Task 6 differed from the previous tasks in that students needed to create sentences rather than answer multiple-choice comprehension questions. In this task, in order to produce new sentences, participants were required to make a decision about additional words which would combine with the new word in an original text. Therefore, strong *evaluation* was induced. Therefore, the Involvement Index of Task 6 is 3 (1+0+2). Among all the tasks, Task 6 was the only one that was output-oriented.

The involvement load of each task is listed in Table 7.1. It shows that the involvement load is lowest for Tasks 1 and 2, moderate for Tasks 3 and 4, and highest for Tasks 5 and 6.

7.3.4 Procedure

The experiment was carried out in June 2011. Six classes were given six different reading tasks during their normal class time. They were only told to read the passage and complete the post-reading activities, except for Group 1, for which there were no post-reading activities. Students were not informed they would have to complete a vocabulary

Table 7.1 Involvement load of six tasks in the present study

Tasks	Involvement Components			Involvement Index
	Need	Search	Evaluation	
1. Reading with glosses in text but no comprehension questions afterwards	-	-	-	0
2. Reading with glosses in text but irrelevant to the comprehension questions	-	-	-	0
3. Reading with glossary in the end relevant to the comprehension questions	+	+	-	2
4. Reading with glosses in margin relevant to the comprehension questions	+	-	+	2
5. Reading with glossary in the end which consists of several options and relevant to comprehension questions	+	+	+	3
6. Reading with glosses in margin and make sentences afterwards	+	-	++	3

test afterwards. Each task took 17 minutes to complete. After the completion of the tasks, the task paper was collected and the participants were given a vocabulary test. They were required to provide the Chinese equivalents for these English words within eight minutes. Their answers were scored afterwards. The delayed post-test was held seven days after the immediate post-test, as in the studies of Hulstijn and Laufer (2001) and Yaqubi et al. (2010). All participants received the same vocabulary test again but the order of the items differed from that in the immediate post-test. This test also took eight minutes. The two post-tests were scored by the first author. The following scoring method was adopted in this study: Zero points were given for items which were not translated or were wrongly translated. One point was given for items which were semantically appropriate, such as the superordinate, synonym, but not the best possible translation for the target item. Two points were assigned to a complete correct translation. The maximum score that could be obtained on both tests was 16.

Forty-five of the 230 participants who took part in the study had to be excluded from the analysis because they either attended only one test or they did not provide consistent birthdays in two tests, which made it impossible for us to allocate the two test papers to the same student. Complete data sets were obtained from 185 students.

7.4 Results

Students obtained mean scores of 9.97 (SD 4.69) in the immediate post-test and to 6.17 (SD 4.13) in the delayed post-test (out of a maximum of 16 points). We used non-parametric tests to investigate whether these differences were significant because the scores from the immediate post-test and the delayed post-test were not normally distributed. The differences between the two post-tests were significant in a Wilcoxon's Signed Ranks test ($Z=-10.77$, $p < .001$).

The overall results of the study are displayed in Figures 7.1 and 7.2. It shows that students in groups 1 and 2 obtained the lowest scores, followed by those in groups 3 and 4, whilst students in groups 5 and 6 obtained the highest scores. The rank order of the groups is the same for the immediate and the delayed post-tests. The results of the Kruskal Wallis test reveal that the differences between six groups are significant in the immediate post-test ($\chi^2 = 56.02$, $df = 5$, $p < .001$) as well as the delayed post-test ($\chi^2 = 65.21$, $df = 5$, $p < .001$).

In order to get more evidence about the impact of the involvement load of different tasks on vocabulary retention we regrouped the

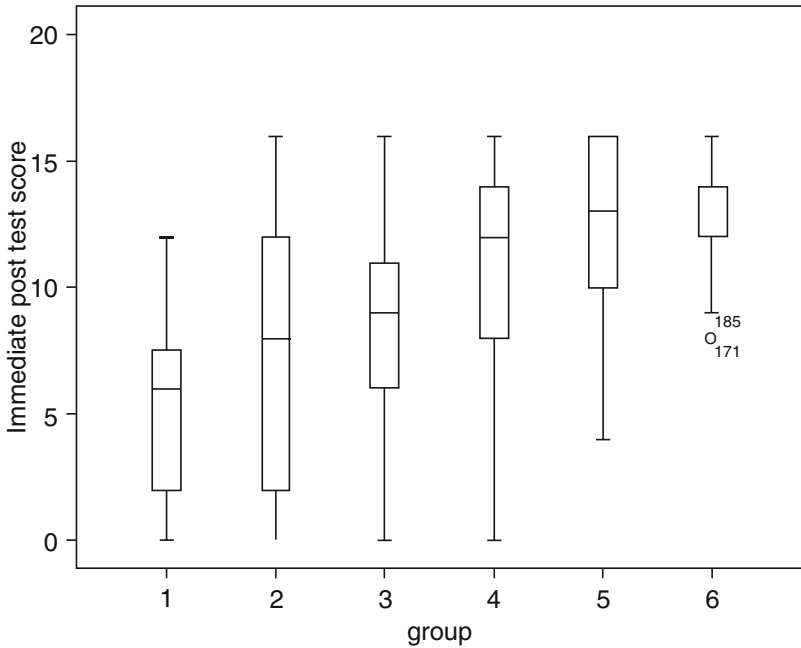


Figure 7.1 Immediate post test scores across groups

original six groups into three new groups according to their scores on the Involvement Index. The first two groups were combined to make Group A, as they both have an Involvement Index of 0. Group B consists of groups 3 and 4 because they share an Index of 2. The last two groups were combined to make Group C, since the Involvement Index of each is 3. As Table 7.2 shows, Group C performed best both in the immediate post-test and in the delayed post-test, whilst group A obtained the lowest mean score in two tests. The differences between these three new groups are significant in the immediate post-test (Kruskal Wallis, $\chi^2 = 41.61$, $df = 2$, $p < .001$) as well as the delayed post-test ($\chi^2 = 56.98$, $df = 2$, $p < .001$). In addition, post hoc comparisons indicate that all three groups are significantly different from each other (Table 7.3).

Further evidence regarding the importance of post-reading activities can be obtained from a comparison of the results of group 1 with those of all the other groups, because all groups except group 1 engaged in additional post-reading activities. After correcting for multiple

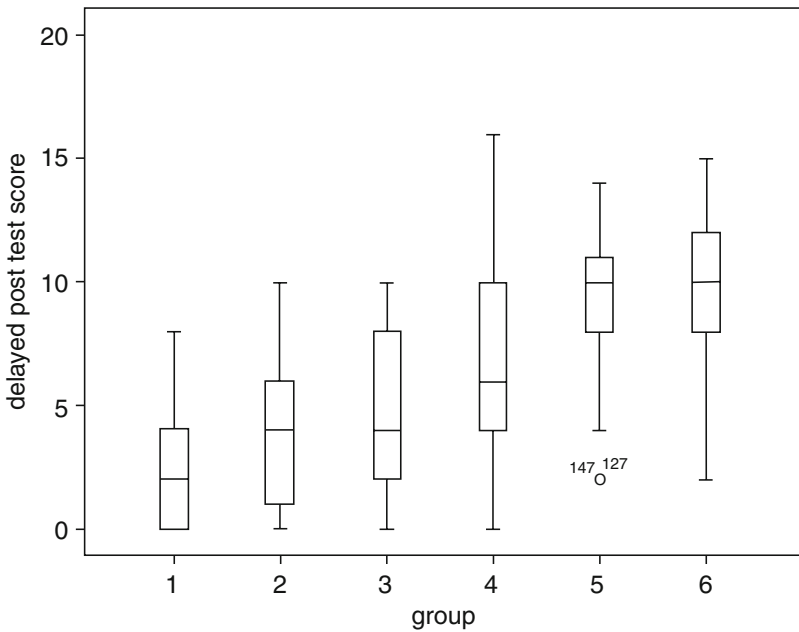


Figure 7.2 Delayed post test scores across groups

Table 7.2 Median of scores among the three groups based on the involvement load index

Groups based on the Involvement Load Index	Tasks	Involvement index	N	Median	
				Immediate	Delayed
Group A	Task 1 & Task 2	0	54	7.00	4.00
Group B	Task 3 & Task 4	2	67	10.00	6.00
Group C	Task 5 & Task 6	3	64	13.50	10.00

Table 7.3 Post hoc comparison of intergroup differences based on involvement load index

Group	Immediate post-test		Delayed post-test	
	z	p (adjusted)	Z	p (adjusted)
A-B	-3.05	.007	-3.01	.008
A-C	-6.43	.000	-7.44	.000
B-C	-3.60	.001	-4.71	.000

Table 7.4 Intergroup differences between group 1 and the other five groups (post hoc comparisons following Kruskal Wallis test)

	Z	p (adjusted for multiple comparisons)	r (effect size)
Groups 1–2	-2.442	ns	–
Groups 1–3	-2.060	ns	–
Groups 1–4	-4.693	<.001	.63
Groups 1–5	-5.273	<.001	.80
Groups 1–6	-6.214	<.001	.83

comparisons with the Bonferroni correction, the post hoc comparisons carried out on the immediate post-test show that Task 1 is not significantly different from Task 2 or Task 3; however, it is significantly different from Tasks 4, 5 and 6 (see Table 7.4). Effect sizes were computed manually for the Kruskal Wallis test, following the procedure outlined in Field (2013, p. 227 and p. 248).

The results for the delayed post-test were very similar in that group 1 was significantly different from groups 4, 5 and 6 but not from the other groups. For reasons of space these results are not reported in detail here.

We investigated the relative importance of the three components (*need*, *search* and *evaluation*) on vocabulary acquisition by regrouping participants into two groups according to the presence or absence of each component: \pm *need*, \pm *search* and \pm *evaluation* (see Table 7.5 for details). We named each group based on the name and the status of the components. For example, Group NA referred to the group which performed the task where the *need* factor was absent; Group NP meant that the group which performed the task where *need* was present. The groupings based on the factors *search* and *evaluation* were created in similar ways. As we can see in Table 7.5, the median values in the factor-present group tended to be higher than in the factor-absent group in both the immediate and the delayed post-tests.

For *need*, the differences between the two groups were significant in the immediate post-test (Mann Whitney U test, $U = 1773.5$, $p < .001$) and the delayed post-test ($U = 1599.50$, $p < .001$), but for *search* the differences were not significant in either post-test. The grouping based on *evaluation* did result in significant differences in the immediate post-test ($U = 1794.50$, $p < .001$) as well as the delayed post-test ($U = 1596.00$, $p < .001$).

Table 7.5 Median of scores based on classifications of involvement components

Components	Groups	N	Tasks	Median	
				Immediate Post-test	Delayed Post-test
Need	Group NA	54	Task 1, 2	7	4
	Group NP	131	Task 3, 4, 5, 6	12	8
Search	Group SA	87	Task 1, 2, 4, 6	12	6
	Group SP	98	Task 3, 5	10	8
Evaluation	Group EA	84	Task 1, 2, 3	8	4
	Group EP	101	Task 4, 5, 6	13	8

Table 7.6 Effect sizes of group differences based on the classification according to *need*, *search* and *evaluation*

Group	Immediate post-test	Delayed post-test
	r	r
Group Need	.394	.434
Group Search	ns	ns
Group Evaluation	.499	.541

Thus, there are significant differences between the groups in their scores on the immediate post-test and the delayed post-test, if the grouping variable is *need* or *evaluation*, but not when *search* is chosen as the group factor. The effect sizes displayed in Table 7.6 for the three groupings indicate that *need* and *evaluation* are relevant factors in initial vocabulary acquisition and retention whilst *search* is not. A stronger effect size was found for *evaluation* than for *need* in the immediate post-test, and the same was true for the delayed post-test. This means that *evaluation* has a stronger impact on scores than *need*. The relative weight of each of the components of the involvement load model is therefore as follows: *evaluation* > *need* > *search*. Effect size differences appear to be slightly higher for the delayed post-tests than for the immediate post-tests for *need* as well as *evaluation*, but these differences are not significant.

Finally we focused on the impact of differences in the involvement load of tasks on the degree of vocabulary loss between the immediate

Table 7.7 Vocabulary loss between the immediate and the delayed post tasks

Vocabulary Loss	Mean	SD
Task 1	-2.91	2.234
Task 2	-3.90	3.833
Task 3	-4.10	2.857
Task 4	-4.51	3.280
Task 5	-3.13	2.617
Task 6	-3.85	2.949

post-test and the delayed post-test. Contrary to our expectations, all groups lost a roughly equal number of words between the immediate post-test and the delayed post-test (see Table 7.7). A Mann Whitney U test was used to investigate whether the differences between the six conditions with respect to the number of items lost after one week were significant, but these analyses did not reveal any significant differences.

7.5 Discussion

First of all it is important to note that all six treatments led to some acquisition of target words, which confirms that incidental vocabulary acquisition through reading is possible. However, different tasks had different effects on vocabulary learning and retention: the results revealed that scores increased from task 1 to task 6, which lends support to the ILH in that tasks with a higher involvement load are more effective for vocabulary acquisition than those with a lower involvement load, and confirms earlier findings of Paribakht and Wesche (1997) that practising words in post-reading activities supports vocabulary learning. Although Laufer and Hulstijn (2001) have argued that the depth of processing tasks would necessarily require a longer amount of time to complete, this study has shown that even when the time-on-tasks is controlled across tasks, the tasks with higher involvement load led to better vocabulary retention both in the immediate post-test and in the delayed post-test. This confirms earlier findings of Kim (2008; 2011), who also controlled for time-on-task, but contrasts with Yaqubi et al. (2010), who found students who carried out a task to which they had allocated a high Involvement Index obtained lower scores than students who performed tasks with a lower Involvement Index. We think the poor results of Yaqubi et al.'s task with a high Involvement Index might

be due to the fact that students had to look up the meanings of words in a dictionary. This might have been too difficult and time-consuming, and might have distracted from reading and understanding the text. In addition, it is not clear whether students brought the same dictionaries or different dictionaries, which is a confounding factor. Finally, it is not clear whether the authors controlled for time-on-task, which is essential to be able to evaluate the differences in type-of-task.

The fact that we found no significant differences between Task 1 and Task 2 shows that the nature of post-reading activities does matter. If questions about the text are unrelated to the target items (as in Task 2), students do not feel the need to learn the words, and the two tasks are equally (in)effective. However, the absence of significant differences between Task 1 and Task 3 indicates that answering questions that are relevant for the target items does not necessarily increase students' engagement with the words to a sufficient degree. It is only when they need to *evaluate* the words in the text against other words (as in Tasks 4, 5 and 6) that they need to process the words more deeply and this increases their chances of remembering the words in a post-test.

No differences were found in initial vocabulary learning or retention among students who carried out Task 5 or Task 6, which were equivalent with respect to the Involvement Index but differed from each other because Task 5 was more input-oriented and Task 6 more output-oriented. This confirms Laufer and Hulstijn's (2001) prediction that input- and output-oriented tasks will be equally beneficial for vocabulary acquisition if the involvement load is kept constant across tasks. Our results therefore provide little support for the findings of Yaqubi et al.'s study (2010), who concluded that tasks that are equivalent from the perspective of involvement but differ from each other because of their orientation towards input or output do not necessarily lead to the same results.

Our study provides clear evidence that the three components (*need*, *search* and *evaluation*) differ significantly from each other with respect to their impact on incidental vocabulary acquisition. The effect sizes revealed that the largest proportion of the variance was explained by *evaluation*, followed by *need*, with *search* in third position. Other researchers also concluded that *evaluation* is the most important component of involvement. Kim (2011) claims that this is particularly the case for initial vocabulary learning, but in our study this was also found to be the case for vocabulary retention. Although *need* is the second most important factor among the three, we found it is difficult to control or manipulate. For instance, Task 1 and Task 2 were designed *not* to trigger

need, but we could not ascertain that they did not trigger any need to learn the words among the learners. Maybe some participants in Task 2 felt a strong need to learn the words just because they were curious about the target words, although the post-reading tasks did not require any need to comprehend the target words. Hence, *need* remains hard to measure as its role depends to some extent on learners' motivation and attitudes towards the task.

Our study also revealed that there was a decrease in the mean scores among all groups from the immediate post-test to the delayed post-test, which is to be expected as learners often forget some newly learned words after a few days. Therefore, the results indicate that reinforcement of newly learned words is still needed if students are to remember them in the longer term, regardless of the amount of involvement load of the vocabulary learning task. The fact that effect sizes were slightly higher for the delayed post-test is interesting in the light of Kim's (2011) comments about the importance of investigating the long term effects of tasks with different involvement loads on the acquisition of L2 vocabulary.

Our experiment confirms Laufer and Hulstijn's (2001) suggestion that the impact of *search* on incidental vocabulary acquisition might be lower than that of the other two components, because *search* was found to have no significant effect on incidental vocabulary acquisition in our study. A possible reason for the lack of a significant effect of *search* may be the manipulation of the construct itself in the current study. In Laufer and Hulstijn's (2001) study, *search* was triggered by consulting a dictionary or teachers. In the current study, however, *search* was operationalized by referring students to a glossary at the end of the text with only L1 equivalents. We operationalized the presence or absence of *search* by the location of the glossary, that is *search* was absent when a marginal glossary was provided, but present when the glossary was presented at the end with words in alphabetical order. This kind of *search* was relatively limited by comparison with the approach suggested by Laufer and Hulstijn, and it may explain why in our study *search* was found to have little impact on vocabulary learning and retention.

7.6 Pedagogical implications

In the present study, we have seen how certain reading tasks can contribute to vocabulary acquisition, which may have important implications

for L2 teaching and learning in general, but in the Chinese context in particular. The study demonstrates, first of all, that learning vocabulary through reading is possible and feasible, but reading with enhancement activities tends to be more effective. It is particularly important that we have shown that incidental vocabulary learning works in the Chinese context, because there is less awareness in China of the potential of incidental vocabulary learning and learners tend to rely on rote learning to enhance their vocabularies (Li, 2004). As we have shown, task 1 (where translation equivalents were given in the text) and task 2 (with post-reading questions which are irrelevant to the target words) have a low involvement load because learners do not need to engage with the new words at all. Post-reading activities which are relevant to the target words, such as those in tasks 3, 4, 5 and 6 which require learners to engage with the meaning of the new words in a variety of ways, are much more beneficial for vocabulary acquisition. Teachers should therefore be aware of the importance of the involvement load of tasks they develop. Aiming at designing tasks with a high involvement load will not necessarily limit teachers' choice of task types. As the result from the comparison between Task 5 and Task 6 suggests, there is no significant difference between input-oriented tasks and output-oriented tasks both in initial vocabulary learning and in retention. As long as a high involvement load can be induced, teachers have many options in designing reading tasks.

Because learners are likely to forget some of the vocabulary items they have learned after a certain amount of time has elapsed, it is necessary for teachers to provide repeated exposures and additional tasks to maintain the initial vocabulary gains. For instance, teachers can create opportunities for students to encounter the same words in different contexts and to process the words several times in doing various post-reading tasks. Finally it may be beneficial for teachers as well as learners to investigate to what extent practice with incidental vocabulary learning changes students' own vocabulary learning strategies (Schmitt, 1997) and their perceptions of their own learning.

7.7 Conclusion

The present study set out to investigate the effects of different tasks on incidental vocabulary acquisition. In an attempt to test the Involvement Load Hypothesis (ILH), an experiment was conducted among Chinese students in a secondary vocational school whose

proficiency was estimated to be at A2 level on the CEFR. To the best of our knowledge, this was the first test of the ILH among students with a relatively low level of proficiency. The results showed that the students learned more words in reading tasks with a higher involvement load than in tasks with a lower involvement load both in the immediate post-test and the delayed post-test, which is in accordance with Hulstijn and Laufer's (2001) finding that tasks with higher involvement load lead to better vocabulary learning and retention. While through mere reading students can learn a certain number of words, this method is far from effective. We also found that the three components of involvement construct did not carry the same weight. *Evaluation* turned out to be the most important of the three and *search* was the least important. Students carrying out output-oriented tasks did not outperform those doing input-oriented tasks with the same involvement load. Thus, our study does not support the findings of Yaqubi et al. (2010) on this point.

As with Hulstijn and Laufer (2001), the current study focused on learning and retention of word meanings only. It remains to be seen whether tasks with a higher involvement load will lead to better learning of other aspects of word knowledge, as described in Nation (2001). For example, we do not know whether learning of derivational suffixes, formulaic sequences or collocations improves in tasks with a higher involvement load. Hence, it is recommended that future research should address the effects of involvement load on the learning of other aspects of vocabulary knowledge. Further research should also look into the long-term effects of tasks with different involvement loads, as the current study suggests the different effects of tasks persisted after one week, but we do not know if these effects would be measurable later on and which types of activities help support vocabulary retention in the longer term.

Appendix 7.1 Tasks 1–6

Reading passage:

When the computer is down

The most frightening words in the English language are “Our computer is down”. You hear these words more and more when you are on business. The other day I was at the airport, where I was waiting for a ticket to Washington. But the girl in the ticket office said, “I’m sorry, our computer is down. That’s the reason why we can’t sell tickets.”

I looked down at the computer and every passenger was just standing there drinking coffee and staring at the black screen. Then I asked her, “What do all you people do?”

“We give the computer the information about your trip, and then it tells us whether you can fly or not.”

After the girl told me they had no backup computer, I said, “Let’s forget the computer. What about your planes? They are still flying, aren’t they?”

“I couldn’t tell without asking the computer.”

“Are there any other airlines that are flying to Washington within the next hours?”

“I wouldn’t know,” she said, pointing at the dark screen, “Only ‘IT’ knows. ‘IT’ can’t tell me.”

By this time there were quite a few people standing in lines. Word soon spread to other travellers that the computer was down. Some people went white, some people started to cry and still others kicked their luggage...

Target words: frightening staring at screen backup airline spread kick luggage

Task 1

Gender: Birthday:

Directions: read the following passage for fun and get the general meaning in 20 minutes.

When the Computer Is Down

The most **frightening** (可怕的) words in the English language are “Our computer is **down** (停止运行的)”. You hear these words more and more when you are on business. The other day I was at the airport, where I was waiting for a ticket to Washington. But the girl in the ticket office said, “I’m sorry, our computer is down. That’s the reason why we can’t sell tickets.”

I looked down at the computer and every passenger was just standing there drinking coffee and **staring at** (盯着) the black **screen** (屏幕). Then I asked her, “What do all you people do?”

“We give the computer the information about your trip, and then it tells us whether you can fly or not.”

After the girl told me they had no **backup** (备用的) computer, I said, “Let’s forget the computer. What about your planes? They are still flying, aren’t they?”

“I couldn’t tell without asking the computer.”

“Are there any other **airlines** (航线) that are flying to Washington within the next hours?”

“I wouldn’t know,” she said, **pointing at** (指向) the dark screen, “Only ‘IT’ knows. ‘IT’ can’t tell me.”

By this time there were quite a few people standing in lines. Word soon **spread** (传播, 散布) to other travellers that the computer was

down. Some people went white, some people started to cry and still others **kicked** (踢) their **luggage** (行李)...

Task 2

Gender: Birthday:

Directions: read the following passage and complete the comprehension questions in 20 minutes.

See the reading passage in **Task 1**

Reading comprehension

- Where was the writer the other day?
A. at home; B. at an airport; C. in a hotel; D. in a computer store
- Which city was the writer taking the plane to?
A. Washington B. Paris C. Tokyo D. London
- Why the writer was going to that place? According to the passage, the most possible answer should be
A. Visiting his/her friends.
B. Visiting his/her family.
C. Travelling.
D. On business.
- Where did the girl work?
A. She was working in the ticket office.
B. She was working in the school.
C. She was working in the bank.
D. She was working in the restaurant.
- According to the passage, which of the following words were heard more and more in the English language?
A. The tickets were sold out.
B. Contact us during working hours.
C. The airline has been cancelled.
D. Our computer is down.
- Why the girl told the writer that they couldn't sell tickets?
A. Because the computer was down.
B. Because the tickets were sold out.
C. Because the writer had no money.
D. Because the airline had been cancelled.
- What were the other passengers drinking while standing in the line?
A. They were drinking cola.
B. They were drinking coffee.
C. They were drinking fruit juice.
D. They were drinking water.

8. According to the girl, were the planes still flying?
- A. Yes.
 - B. No.
 - C. She couldn't tell without asking the computer.
 - D. She was unwilling to tell.

Task 3

Gender: Birthday:

Directions: read the following passage and complete the comprehension questions in 20 minutes.

When the Computer Is Down

The most **frightening** words in the English language are "Our computer is **down**". You hear these words more and more when you are on business. The other day I was at the airport, where I was waiting for a ticket to Washington. But the girl in the ticket office said, "I'm sorry, our computer is down. That's the reason why we can't sell tickets."

I looked down at the computer and every passenger was just standing there drinking coffee and **staring at** the black **screen**. Then I asked her, "What do all you people do?"

"We give the computer the information about your trip, and then it tells us whether you can fly or not."

After the girl told me they had no **backup** computer, I said, "Let's forget the computer. What about your planes? They are still flying, aren't they?"

"I couldn't tell without asking the computer."

"Are there any other **airlines** that are flying to Washington within the next hours?"

"I wouldn't know," she said, **pointing at** the dark screen, "Only 'IT' knows. 'IT' can't tell me."

By this time there were quite a few people standing in lines. Word soon **spread** to other travellers that the computer was down. Some people went white, some people started to cry and still others **kicked** their **luggage**...

Reading comprehension

1. According to the passage, "Our computer is down" are the most ___ words in the English language.
- A. exciting
 - B. terrible
 - C. helpful
 - D. cheerful

Vocabulary glosses:

airline n. 航班

backup adj. 备用的

down adj. 停止运行的

frightening adj. 可怕的

kick v. 踢

luggage n. 行李

2. What was the airline that the writer was taking?
- A. The airline flying to Washington.
 - B. The airline flying to Paris.
 - C. The airline flying to Tokyo.
 - D. The airline flying to London.
3. What was the matter with the computer?
- A. It was working actively.
 - B. It was breaking into pieces.
 - C. It was fine.
 - D. It stopped working.
4. When the computer was down, the screen turned to be
- A. black B. green C. red D. yellow
5. What could the girl in the ticket office do for the passengers without asking the computer?
- A. She could sell a ticket.
 - B. She could write out a ticket.
 - C. She could answer the passenger's questions.
 - D. She could do nothing.
6. If there had been a backup computer, which of the following situation would NOT happen?
- A. The girl could do nothing.
 - B. The girl could sell a ticket.
 - C. The girl could answer the passenger's questions.
 - D. Everything would continue working.
7. Which of the following statement is NOT mentioned?
- A. Some people went white.
 - B. Some people quarrelled with the girl.
 - C. Some people started to cry.
 - D. Some people kicked their luggage.
8. The last paragraph suggests that
- A. A modern computer won't be down
 - B. Computers can take the place of humans
 - C. Sometimes a computer may bring suffering to people
 - D. There will be great changes in computers.

point at 指向
screen n. 屏幕
spread v. 传播, 散布
stare at 盯着

Task 4

Gender: Birthday:

Directions: read the following passage and complete the comprehension questions in 20 minutes.

When the Computer Is Down

The most **frightening** words in the English language are “Our computer is **down**”. You hear these words more and more when you are on business. The other day I was at the airport, where I was waiting for a ticket to Washington. But the girl in the ticket office said, “I’m sorry, our computer is down. That’s the reason why we can’t sell tickets.”

frightening adj. 可怕的
down adj. 情绪低落
adj. 停止运行的
prep. 向下

I looked down at the computer and every passenger was just standing there drinking coffee and **staring at** the black **screen**. Then I asked her, “What do all you people do?”

stare at 盯着
screen n. 屏幕
n. 纱窗
v. 掩藏, 遮蔽
v. 放映, 播放

“We give the computer the information about your trip, and then it tells us whether you can fly or not.”

After the girl told me they had no **backup** computer, I said, “Let’s forget the computer. What about your planes? They are still flying, aren’t they?”

backup n. 增援, 援助
adj. 备用的

“I couldn’t tell without asking the computer.”

“Are there any other **airlines** that are flying to Washington within the next hours?”

airline n. 航班

“I wouldn’t know,” she said, **pointing at** the dark screen,
“Only ‘IT’ knows. ‘IT’ can’t tell me.”

point at 指向

By this time there were quite a few people standing in lines. Word soon **spread** to other travellers that the computer was down. Some people went white, some people started to cry and still others kicked their **luggage**...

Reading comprehension (see the reading comprehension in Task 3)

spread v. 展开, 铺开
v. 传播, 散布
v. 扩散, 蔓延
kick v. 踢
v. 踢球得分;
射门得分
luggage n. 行李

Task 5

Gender: Birthday:

Directions: read the following passage and complete the comprehension questions in 20 minutes.

When the Computer Is Down

The most **frightening** words in the English language are “Our computer is **down**”. You hear these words more and more when you are on business. The other day I was at the airport, where I was waiting for a ticket to Washington. But the girl in the ticket office said, “I’m sorry, our computer is down. That’s the reason why we can’t sell tickets.”

I looked down at the computer and every passenger was just standing there drinking coffee and **staring at** the black **screen**. Then I asked her, “What do all you people do?”

“We give the computer the information about your trip, and then it tells us whether you can fly or not.”

After the girl told me they had no **backup** computer, I said, “Let’s forget the computer. What about your planes? They are still flying, aren’t they?”

“I couldn’t tell without asking the computer.”

“Are there any other **airlines** that are flying to Washington within the next hours?”

“I wouldn’t know,” she said, **pointing at** the dark screen, “Only ‘IT’ knows. ‘IT’ can’t tell me.”

By this time there were quite a few people standing in lines. Word soon **spread** to other travellers that the computer was down. Some people went white, some people started to cry and still others **kicked** their **luggage**...

Vocabulary glosses: airline n. 航班 backup n. 增援, 援助 adj. 备用的 down adj. 情绪低落 adj. 停止运行的 prep. 向下 frightening adj. 可怕的 kick v. 踢 v. 踢球得分; 射门得分	luggage n. 行李 point at 指向 screen n. 屏幕 n. 纱窗 v. 掩藏, 遮蔽 v. 放映, 播放 spread v. 展开, 铺开 v. 传播, 散布 v. 扩散, 蔓延 stare at 盯着
--	--

Reading comprehension (see the reading comprehension in Task 3)

Task 6

Gender: Birthday:

Directions: read the following passage and complete the comprehension questions in 20 minutes.

See the reading passage in Task 4

Making sentences with the following words.

1. frightening
2. stare at
3. screen
4. backup
5. airline
6. spread
7. kick
8. luggage

References

- Adolphs, S., & Schmitt, N. (2003). Lexical coverage of spoken discourse. *Applied Linguistics*, 24(4), 425–438.
- Anderson, J. R. (1995). *Cognitive Psychology and Its Implications* (4th ed.). New York: Freeman.
- Cheng, Ch. M., & Yang, M.-Y. (1989). Lateralization in the visual perception of Chinese characters and words. *Brain and Language*, 36(4), 669–689.
- Chen, X., Ramirez, G., Luo, Y., Geva, E., & Ku, Y.-M. (2012). Comparing vocabulary development in Spanish- and Chinese speaking ELLs: The effects of metalinguistic and sociocultural factors. *Reading and Writing*, 25(8), 1991–2020.
- Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press. Retrieved from http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf.
- Craik, F. I. M., & Lockhart, R. S. (1972). Level of processing: A framework for memory research. *Journal of verbal learning and verbal behaviour*, 11(6), 671–684.
- Ellis, N. C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning*, 43(4), 559–617.
- Field, A. (2013). *Discovering statistics: Using IBM SPSS statistics* (4th ed.). London: Sage.
- Gass, S. (1999). Discussion: Incidental vocabulary learning. *Studies in Second Language Acquisition*, 21(2), 319–333.
- Hill, M., & Laufer, B. (2003). Type of task, time-on-task and electronic dictionaries in incidental vocabulary acquisition. *International Review of Applied Linguistics*, 41(2), 87–106.
- Hulstijn, J. H. (2003). Incidental and intentional learning. In C. J. Doughty & M. H. Long (eds). *The Handbook of Second Language Acquisition* (pp. 349–381). Oxford: Blackwell.
- Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal

- glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal*, 80(3), 327–339.
- Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, 51(3), 539–558.
- Hunt, A., & Beglar, D. (1998). Current research and practice in teaching vocabulary. *The Language Teacher*, 22(1), 7–25.
- Joe, A. (1995). Text-based tasks and incidental vocabulary learning. *Second Language Research*, 11(2), 149–158.
- Joe, A. (1998). What effects do text-based tasks promoting generation have on incidental vocabulary acquisition. *Applied Linguistics*, 19(3), 357–377.
- Jones, F. R. (1995). Learning an alien lexicon: A teach-yourself case study. *Second Language Research*, 11(2), 95–111.
- Kim, Y. J. (2008). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning*, 58(2), 285–325.
- Kim, Y. J. (2011). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning*, 61(Suppl. 1), 100–140.
- Ko, Myong Hee (2005). Glosses, comprehension, and strategy use. *Reading in a Foreign Language*, 17(2), 125–143.
- Larrañaga, P., Treffers-Daller, J., Tidball, F., & Gil Ortega, M. (2012). L1-transfer in the acquisition of manner and path in Spanish by native speakers of English. *International Journal of Bilingualism*, 16(1), 117–138.
- Laufer, B., & Hulstijn, J. H. (2001). Incidental vocabulary acquisition in a second language: the construct of task induced involvement. *Applied Linguistics*, 22(1), 1–26.
- Lawson, M. J., & Hogden, D. (1996). The vocabulary-learning strategies of foreign-language students. *Language Learning*, 46(1), 101–135.
- Li, X. (2004). *An analysis of Chinese learners' beliefs about the role of rote learning in vocabulary learning strategies*. Unpublished doctoral dissertation. University of Sunderland, Sunderland.
- Nagy, W., & Herman, P. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. Mckeown & M. Curtis (eds.). *The nature of vocabulary acquisition* (pp. 19–35). Hillsdale, NJ: Lawrence Erlbaum.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. London: Newbury House.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, 63(1), 59–81.
- Nation, P. (2007). The four strands. *Innovation in Learning and Teaching*, 1(1), 1–12.
- Nation, I. S. P., & Wang, K. (1999). Graded readers and vocabulary. *Reading in a Foreign Language*, 12(2), 355–380.
- Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary development. In J. Coady & T. Huckin (eds.). *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 174–200). Cambridge: Cambridge University Press.
- Pellicer-Sanchez, A. (2012). Automaticity and speed of lexical access: Acquisition and assessment. PhD thesis, University of Nottingham.

- Schmidt, R. (1994). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. C. Ellis (ed.). *Implicit and explicit learning of languages* (pp. 165–210). London: Academic Press.
- Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (eds.). *Vocabulary: Description, acquisition and pedagogy* (pp. 199–227). Cambridge: Cambridge University Press.
- Schmitt, N. (2007). Current perspectives on vocabulary teaching and learning. In J. Cummins & C. Davison (eds.). *International handbook of English language teaching: Part II* (pp. 827–841). New York: Springer.
- Shao, X. (2014). A study of Chinese college students' English reading anxiety. *American Journal of Educational Research* 2(5). Retrieved from: pubs.sciepub.com/education/2/5/10/
- Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition*, 15(2), 165–179.
- Smith, S., Kilgarriff, A., & Sommers, S. (2008). Making better wordlists for ELT: Harvesting vocabulary lists from the web using WebBootCat. Retrieved from <http://www.kilgarriff.co.uk/Publications/2008-SmithKilgSommers-DAE-BetterWordlists.doc>
- Yaqubi, B., Rayati, R. A., & Allemzade Gorgi, N. (2010). The involvement load hypothesis and vocabulary learning: The effect of task types and involvement index on L2 vocabulary acquisition. *The Journal of Teaching Language Skills*, 2(1), 145–163.