

Incidental and intentional vocabulary learning: a case study of Meaning-Given, Meaning-Inferred with MC, and Pure Meaning-Inferred methods on the retention of L2 word meanings in a Chinese University

Qin LI

College of Foreign Languages, China Three Gorges University, China

Email address:

jennylih@sina.com

To cite this article:

Qin LI. Incidental and Intentional Vocabulary Learning: A Case Study of Meaning-Given, Meaning-Inferred with MC, and Pure Meaning-Inferred Methods on the Retention of L2 Word Meanings in a Chinese University. *Education Journal*. Vol. 2, No. 4, 2013, pp. 138-148. doi: 10.11648/j.edu.20130204.16

Abstract: This paper reports a case study investigating and comparing three word-learning methods [i.e. Meaning-Given, Meaning-Inferred with Multiple Choices (Meaning-Inferred with MC as below), and Pure Meaning-Inferred] in two modes of learning (i.e. incidental and intentional learning, ICL & ITL respectively as below) in terms of the retention of L2 word meanings in a Chinese University. Findings suggest that the mode of ITL led to significantly higher retention than the mode of ICL did. However, in terms of different word-learning methods, different results appeared. It has been suggested that it is crucial for teachers to balance the use of the two learning modes, input more modifications directing students to process the lexical information more elaborately, and put more emphasis on the functions of rehearsal and reactivation of new lexical information.

Keywords: ICL, ITL, Word-Learning Methods, Retention of L2 Word Meanings

1. Introduction

Many learners of a second (L2) or foreign language worry about the heavy burden of tackling thousands of words in vocabulary learning. For instance, in most of Chinese universities, non-English majors are required to pass Band 4 in the China National General English Examination (CNGEE), which usually make great demands on students' vocabulary size (nearly a threshold of 5000 English words); otherwise, they may fail to get a Bachelor's degree. However, many college English teachers appear to take a negative attitude towards vocabulary teaching, often regarding vocabulary learning as the work of students themselves', since they assume that university students have acquired sufficient ability to learn vocabulary through extensive reading (e.g. ICL) outside of class and intensive reading in class (since very few formal investigations about the real situation of vocabulary teaching in Chinese universities were found in the literature review, this claim is based primarily on the teaching experience of the author and her colleagues teaching at a

Chinese university). However, the comparatively lower retention found in ICL has led some other researchers to further evaluate the mode of ICL and the devaluated role of intentional learning (ITL) which has been argued to contribute significantly to the beginners' and intermediate learners' vocabulary learning in the later stages of language acquisition [1] [2]. Thus, many studies have been conducted in terms of the two modes of learning, but with no conclusive result regarding which is the superior mode. Furthermore, it has been argued that input modifications may help promote the efficiency of vocabulary learning by engaging students more fully in processing the new lexical information and leading them to focus more clearly on the target words while completing other language activities, such as listening and reading [3] [4] [5]. Thus, the question of which word-learning method is more effective and efficient for L2 vocabulary acquisition has become one of the most debated issues in the two learning modes respectively in L2 pedagogy. In particular, the meaning-given and meaning-inferred methods have been the focus of most of the concern in previous studies [3] [4] [5] [6] [7] [8] [9].

The inconclusive answer to the above question has motivated the author to do the present experiment with Chinese College English learners because the previous studies have rarely employed them as subjects. Second, the current vocabulary teaching situation at Chinese universities, as stated earlier, has promoted me to carry out this investigation to make an empirical contribution to the long-contested issue of the meaning-given and meaning-inferred (with or without modifications) methods with respect to the two modes of learning (ICL & ITL). Hence, three research questions are addressed in this study:

- (1) Does intentional vocabulary learning (ITL) lead to better retention of word meanings than incidental vocabulary learning (ICL)?
- (2) Which of the three methods leads to the best retention of word meanings in the mode of incidental learning (ICL)?
- (3) Which of the three methods leads to the best retention of word meanings in the mode of intentional learning (ITL)?

2. Literature Review

2.1. Incidental and Intentional Vocabulary Learning

According to Craik [10] and Eysenck [11], incidental and intentional learning can be distinguished 'in terms of the use of pre-learning instructions that do, or do not, forewarn subjects about the existence of a subsequent retention test'. As Hulstijn [12] notes, 'incidental vocabulary learning refers to the learning of vocabulary as the by-product of any activity not explicitly geared to vocabulary learning, with intentional vocabulary learning referring to any activity aiming at committing lexical information to memory'. Psychological and pedagogical studies about these two different modes of learning have been conducted by many researchers since the beginning of the twentieth century [10] [11] [12]. Among numerous studies conducted on first and second language (L1 & L2) learning, different aspects of language have been studied in terms of these two learning modes: for example, vocabulary, grammar, etc.. However, in line with the purpose of the current study, incidental and intentional vocabulary learning in second language are the main focus of this article and the relevant definitions above will be adopted in this study.

2.2. Word-learning Methods

According to many studies conducted before [6] [7] [8] [11] [12], ICL through reading accounts for a substantial proportion of vocabulary acquisition. However it is also suggested that the probability of acquiring an unfamiliar word from a single exposure in context without attention is disappointingly low, merely between .10 to .15 [13] or even less, according to Herman et al. [14] who set the score at about .05. Since the processing activities may enhance L2 vocabulary learning, it has been widely advocated that

input modifications tend to guide learners to pay more attention to target words and process the lexical information more elaborately, which would subsequently lead to a better retention of target words. To foster more effective ICL and ITL, researchers have called for attention to other provisos: for example, input modifications with different kinds of word-learning methods [3] [4] [5]. Hence, the question arises about which method leads to better retention of new words with respect to two main word learning methods---the meaning-given method (the meaning of unfamiliar words is given in the form of a L1 translation or L2 synonym) and the meaning-inferred method (the meaning of target words is inferred by readers from the context). This has been a hotly-debated issue with regard to vocabulary acquisition in foreign language teaching and learning [5] [9] [15]. However, the pure meaning-inferred method appears to be in an inferior position [5], different modifications have been made for the meaning-inferred method to avoid those so-called disadvantages, such as multiple-choice procedure (i.e. MC procedure, in which a L2 synonym or L1 translation of a target word accompanied with several distractors are presented after the target word in the text for subjects to infer the meaning), inferring with the aid of a dictionary, and concise context, etc. [3] [5] [16].

It appears to be extremely difficult to draw a conclusion concerning the exact effects of the meaning-given and meaning-inferred methods in both ICL and ITL, even if different modifications have been made to help learners with more accurate inferring. Therefore, many issues about these two word-learning methods with or without modifications in ICL and ITL remain inconclusive. Moreover, rarely did researchers employ Chinese subjects of English language learning in their experiments, which stimulated the author to carry out the present study in a Chinese university to explore the different effects of the three word learning methods [i.e. Meaning-Given, Meaning-Inferred with Multiple Choices (Meaning-Inferred with MC as below), and Pure Meaning-Inferred] in two modes of learning (i.e. incidental and intentional learning, ICL & ITL respectively as below), and hopefully the results can have some pedagogical implications for Chinese College English teachers and students to facilitate their vocabulary teaching and learning.

3. Methodology

3.1. Subjects

In general, this study adopted 'an approach to cluster random sampling' [17]. The subjects in this study were seven intact English listening and speaking classes of sophomores in a second-level university---TG University in China (with a total of 262 students). One class (37 students) of the seven attended the pilot study and the other six classes (225 students in all) took part in the main study. All of the subjects were native speakers of Mandarin

Chinese and also learners of English as a foreign language. In this university, for freshmen and sophomores who are non-English majors, there are four levels of general English proficiency to reach, namely, Band 1, Band 2, Band 3, and Band 4. [These levels have been classified according to the test program of the China National General English Examinations (CNGEE).] The subjects in this study were students of the approximate English proficiency, who had already passed Band 3 and were aiming at passing Band 4. In addition, all of them were studying towards various majors in different departments in the University, who chose English teachers by themselves through the internal campus network. Therefore, to a great extent, they were randomly arranged in each class in terms of gender and majors. Hence, basically they shared an equal footing in learning the target words of which they had no pre-knowledge, a proviso that was later checked with the results of the immediate post-tests.

3.2. Instruments

3.2.1. The Target Words

Twenty unfamiliar words were selected from the reading text, based on three criteria: (1) assumed unfamiliarity to the subjects; (2) the feasibility of being inferred through context; (3) words in the glossary of Band 4 designed by the committee of the CNGEE. Each criterion was mainly judged by the author, who has eleven years of experience in teaching English to non-English majors at a Chinese university. As final confirmation, the ten target words were selected from the result of a pilot study with respect to these three criteria. All of these words were the ones which the students knew the least among those twenty unfamiliar words which were initially chosen. The ten target words in the study, in order of appearance in the text, are: orientation, tilt, canopy, fungi, robust, ascend, epiphyte, nutrient, symbiotic, and reptile (six nouns, two verbs, and two adjectives).

3.2.2. The Reading Text

The reading passage was adopted from the reading-comprehension section of a sample examination paper of Band 4. The passage is about the tropical rainforests and contains 903 words. It was selected on several grounds: (1) Subjects are supposed to have some general idea of the topic but little knowledge of the words relevant to this issue; (2) Most of words in the passage are roughly within the range of their level for facilitating comprehension [18]. Five multiple-choice reading comprehension questions were also adapted from the same material. In addition, the passage was revised in such a way that the understanding of the ten target words was generally relevant to the completion of the comprehension questions (see Appendix B). However, the reading text was presented in three different ways (see Appendices C, D, & E). For Groups 1 and 2, ten Chinese equivalents were given to the ten target words in the reading text; for Groups 3 and 4, three choices of Chinese translation (one correct choice

with two choices of distractors) were given to each of the ten target words in the reading text for subjects to choose while inferring meanings through context; for Groups 5 and 6, the text was presented without meanings to the target words.

3.2.3. The Immediate Post-tests

Two vocabulary post-tests (words-in-isolation test and words-in-context test) were administered in this study to measure the subjects' retention of the meanings of the target words immediately after the reading comprehension, following the sequence of the words-in-isolation test first (Post-test 1) and then the words-in-context test (Post-test 2) (see Appendices F & G). Each contained the ten target words and the subjects were asked to provide the Chinese equivalents for these words. The only difference between these two tests was that in Post-test 2 subjects were given the original reading text again without any cue, which might result in a different level of retention. However, the adoption of immediate post-tests in this study without delayed post-test was mainly because the author's research questions focused on the differential effect of various types of information processing during, yet not after, the initial exposure to new lexical information [12].

3.3. Procedures

The selection of the reading material and the ten target words drew heavily on the results of a pilot study, which will be detailed below, followed by an account of the procedures of the main study.

3.3.1. The Pilot Study

Prior to the main study, a pilot study was conducted to serve the following functions: (1) to assess the likelihood of the target-word familiarity among learners of similar language proficiency with the subjects in the main study; (2) to examine the appropriateness of the difficulty level of the reading materials to be used in the main study; (3) to investigate generally how long it takes subjects to finish the reading comprehension. One intact class (37 students) was selected randomly from the seven classes (as stated in Section 3.2.1) who did not attend the main study for the sake of initial exposure to the target words. All the text-reading worksheets (see Appendix A) and a piece of clearly-stated instruction were printed out and given to the English teacher of this class. Moreover, she had been well-informed of the purpose and requirements of the pilot study. Among the 20 assumed-unfamiliar words, 10 target words of the lowest rate of accuracy were selected to be used in the main study. As the students in the main study were of similar language proficiency, it was assumed that most of the target words would be unfamiliar to them. Additionally, all the students in the pilot study finished the task within 15 minutes. Therefore, the time limit of 15 minutes was adopted in the main study.

3.3.2. The Main Study

Three English teachers (one randomly responsible for two groups) of the six groups (as stated in Section 3.2.1) had been well-trained and fully-informed of the purpose and requirements of the main study by the author. Three sets of carefully-written instructions for the three teachers and all the materials needed in the main study were printed out and distributed to each teacher one day before the main study. After all the teachers indicated that they understood what they needed to do, the main study was carried out within one week (but not simultaneously due to the time difference of the arrangement of the English classes at TG University). Generally, all the subjects were not informed of the purpose of the study for the sake of the nature of the experiment, but were told that it was a small in-class task only. The subjects in Groups 1, 3, & 5 were not forewarned of the following vocabulary post-tests (the mode of incidental learning) whereas Groups 2, 4, & 6 were told in advance that they would later be tested on the target words (the mode of intentional learning). Then the corresponding text-reading sheets (see Appendices C, D, & E) were distributed to each group and after 15 minutes all were collected by the teacher. Immediately afterwards, subjects were given Post-test 1. Five minutes later, all the worksheets were collected by the teacher and Post-test 2 was given. Finally, all the worksheets were collected by the teacher after five minutes.

3.4. Data Analysis

The vocabulary post-tests were scored by the three English teachers (mentioned above) with the scoring instructions and the standard keys clearly instructed. A word that was not given with Chinese translation or wrongly given was assigned a score of 0. A correct response received 1 point (the total score for each test was 10). If an answer was controversial in terms of the degree of the semantic approximation, discussion with the author and the other two teachers was sought for this item.

Several steps were taken in the analysis of the retention scores in the two learning modes (ICL & ITL). First, an independent t-test was performed to compare the mean retention scores of the two modes on the whole. In order to observe more details, six t-tests were performed to further compare the mean retention scores of each method in the two modes. Second, since the main unit of analysis was subjects' retention scores of three methods in two post-tests in the two modes respectively, retention scores of the two post-tests in two modes were then submitted to a 3×2 analysis of variance (two-way mixed ANOVA), with method as the between-subjects factor (meaning-given, meaning-inferred with MC, pure meaning-inferred) and test type as the within-subjects factor (post-tests 1 & 2). For the ANOVA of each mode, post-hoc analyses (Scheffé tests) were then administered to pinpoint the location of the statistically significant differences among groups. Finally, in order to give an indication of the strength of the findings, the effect size *r* of effects in each mode was then calculated

(a threshold of .5 was used here). An alpha level of .05 was adopted for all statistical tests.

4. Results and Discussion

4.1. Results and Discussion of Research Question 1

Does intentional vocabulary learning (ITL) lead to better retention of word meanings than incidental vocabulary learning (ICL) ?

4.1.1. Results

To obtain an overview of the two modes of learning (ICL & ITL), the mean scores and standard deviations of all the retentions scores of both post-tests (1 & 2) in two modes were tabulated respectively. They are displayed in Table 1 on the following page. An independent t-test was then administered to compare the mean retention scores in the two modes. In general, the mean retention score in the mode of ITL was significantly higher than that in the mode of ICL, $t(448) = 2.40, p < .05$ (see Table 1 for details).

Table 1. Comparison of mean retention scores of two modes of incidental and intentional learning

Modes	<i>n</i>	M	SD	<i>t</i> -test for Equality of Means		
Incidental	228	4.4605	2.4266	<i>t</i>	<i>df</i>	<i>Sig.</i> (2-tailed)
Intentional	222	5.0586	2.8478	2.400	448	.017

Note. The possible maximum score was 10. *n* refers to the number of retention scores in two post-tests.

To obtain a more specific picture of how different methods led to different retention in the two post-tests in each mode, six t-tests were performed to further compare the mean retention scores of each method in each post-test for the two modes (see Table 3). The mean scores and standard deviations are displayed in Table 2.

Table 2. Mean retention scores of post-tests 1 & 2 in the modes of incidental and intentional learning (ICL & ITL)

Methods	Incidental Learning		Intentional Learning	
	Post-test 1	Post-test 2	Post-test 1	Post-test 2
Meaning-given	2.6053	4.6053	6.3611	7.4444
Meaning-inferred with MC	5.9750	6.8500	5.5135	6.9730
Pure meaning-inferred	2.8333	3.5556	2.2105	2.1053

Table 3. Comparisons of mean retention scores of post-tests 1 & 2 between the modes of incidental and intentional learning (ICL & ITL)

Method	Post-test	<i>t</i> -test for Equality of Means		
		<i>t</i>	<i>df</i>	<i>Sig.</i> (2-tailed)
Meaning-given	1	7.841	72	.000*
	2	6.500	72	.000*
Meaning-inferred with MC	1	.979	75	.331
	2	.268	75	.790
	Pure	1.812	72	.074#
meaning-inferred	2	4.151	72	.000*

Note. * significant at the .05 level. # marginally significant

The t-test results indicated that in terms of the meaning-given method, the mean scores of the two post-tests in the mode of ITL were significantly higher than those in the mode of ICL, which was consistent with the result of the general comparison between the two modes as stated earlier. However, in terms of the pure meaning-inferred method, the mean score of Post-test 2 in the mode of ITL was significantly lower than that in the mode of ICL, and the difference between the mean score of Post-test 1 in the former and that in the latter was marginally significant, which ran counter to the result of the general comparison between the two modes. Lastly, in terms of the meaning-inferred with MC method, there were no significant differences between the mean scores of either post-test in the mode of ITL or those in the mode of ICL, although the mean score in each post-test in the former was slightly lower than that in each post-test in the latter.

To summarize, on the whole, the mean score of all the retention scores in both post-tests in the mode of ITL was significantly higher than that in the mode of ICL. However, in terms of different methods, the results of comparisons were somewhat different.

4.1.2. Discussion

The present study reports a generally significant superiority of ITL to ICL. Krashen [19] claimed in his Input Hypothesis that acquisition only occurs when learners' attention is focused on meaning. In addition, it is argued that at least some degree of conscious attention is necessary for ICL, which is suggested to be peripheral attention (compared to focal attention in the ITL) [20] [21]. It is evident from the lower mean score ($M=4.4605$, $SD=2.4266$) of the two post-tests in the mode of ICL, a single exposure without any attention or with peripheral attention to the word itself is rarely enough to leave a sufficiently deep imprint in the memory to keep the word available for retrieval in the word retention (see Table 1 for details).

Furthermore, as has been claimed by Hulstijn [5], the mode of ITL may invoke all kinds of rehearsal and memorization techniques, which may in turn promote a deeper imprint in the memory, even with only one single exposure, but with enough focal attention to the target word itself. Hence, the finding that the mode of ITL led to a significantly higher mean score ($M=5.0586$, $SD=2.8478$) in the two post-tests could suggest that learners' focal attention towards unfamiliar words may enhance L2 vocabulary learning substantially.

As has been stated above, the focal attention on the meaning of the target words in ITL may lead to better retention in terms of the meaning-given method, since subjects may have used more rehearsal and memorization techniques to memorize the meanings of the target words given to them. However, in the mode of ICL, subjects may have focused on reading comprehension itself, paying no attention or some

peripheral attention to the meanings of the target words. Therefore, with respect to the meaning-given method, the mean scores in both post-tests in ITL were significantly higher than those in ICL.

The comparison result in terms of the meaning-inferred with MC method in the two modes conflicts with that in the Hulstijn's study [5] in which it was found that the mean retention scores of both post-tests in ITL were much higher than those in ICL. However, in the present study, no significant difference was found between them. The plausible explanation is that although in the mode of ICL less attention may have been paid to the meanings of the words, the subjects may still have made some mental effort to infer the meanings with the help of multiple choices, which may lead to some retention. Hence, the difference between the two modes was not significant.

With regard to the pure meaning-inferred method, the comparison result between the two modes ran completely counter to the general comparison as stated earlier. Interestingly, this is also in conflict with some studies: for instance, those conducted by Hulstijn [5], and Mondria [9]. A plausible explanation for the different findings in the present study is that subjects in the mode of ICL may have tended to infer the meanings of the target words to help their reading comprehension whereas subjects in ITL may have failed to infer the meanings of the target words correctly. This could have happened even though the subjects had been forewarned of the subsequent vocabulary tests. Actually, in the pure meaning-inferred method, no cue had been given to the subjects as to what the target words could be, the subjects could have failed to pay attention to them even if in the mode of ITL.

To sum up, despite the differences of word-learning methods, the mode of ITL may generally lead to better retention than the mode of ICL. However, the adoption of different methods may affect the retention results in the two modes, which would indicate that some amount of input modifications appear to help with the retention of word meanings in spite of different learning modes (as has been stated in Section 2.2). In particular, although the mode of ICL seems to be inferior to the mode of ITL, it could be superior to ITL on some occasions with some amount of input modifications (e.g. word-learning methods) and elaboration (e.g. mental effort and elaborateness of processing). It seems to be consistent with the argument that ICL and ITL can be treated as complementary activities, since the characteristics of rehearsal and memorization of ITL could compensate for the inherent disadvantages of ICL to some extent. Moreover, to some degree the exclusively wide coverage of vocabulary acquisition through ICL may compensate for the limitations of ITL.

4.2. Results and Discussion of Research Question 2

Which of the three methods leads to the best retention of word meanings in the mode of incidental learning (ICL)?

4.2.1. Results

Mean retention scores of the two post-tests (1 & 2) in the mode of ICL are displayed in Table 4. To examine whether there was a significant effect for each factor, retention scores were further submitted to a 3x2 analysis of variance (two-way mixed ANOVA), with method as the between-subjects factor (meaning-given, meaning-inferred with MC, pure meaning-inferred) and test type as the within-subjects factor (Post-tests 1 & 2) (see Table 5). In order to evaluate the strength of the findings here, the effect size *r* of each effect (namely the effects of method, test type, and method x test type) was then calculated (a threshold of .5 was adopted here, see Table 6 for details).

Table 4. Number of subjects, mean retention scores, standard deviations of post-tests 1 & 2 in the mode of incidental learning (ICL)

Methods	n.	Post-test 1		Post-test 2	
		M	SD	M	SD
Meaning-given	38	2.6053 ^a	1.5341	4.6053 ^a	1.6691
Meaning-inferred with MC	40	5.9750 ^b	2.1896	6.8500 ^b	2.2366
Pure meaning-inferred	36	2.8333 ^a	1.6125	3.556 ^a	1.6638

Note. The possible maximum score was 10. n refers to the number of subjects in each group. Means in the same column that do not share the same letter differ significantly at $p < .05$ by Scheffé test. In Post-test 1, for example, there was a significant difference between the meaning-given and meaning-inferred with MC methods, whereas there was no significant difference between the meaning-given and pure meaning-inferred methods.

Table 5. ANOVA on the retention scores of post-tests 1 & 2 in the mode of incidental learning (ICL)

Source	Sum of Squares	df	Mean Square	F	p
Between-subjects					
Method	475.822	2	237.911	43.990	.000
S within-group error	600.323	111	5.408		
Within-subjects					
Test type	81.802	1	81.802	56.822	.000
Test type x Method	18.381	2	9.191	6.384	.002
Test type x S Within-group error	159.799	111	1.440		

Note. S=subjects.

Table 6. The effect size *r* for ANOVA on the retention scores of post-tests 1 & 2 in the mode of incidental learning (ICL)

ΩMethod	ΩTesttype	ΩTesttype*Method
.5233	.5733	.2123

Note. A threshold of .5 of effect size *r* was adopted here. For instance, the effect of method had a large effect size whereas the effect of the interaction of test type and method had a small effect size.

A substantively significant method effect was revealed in the ANOVA results, $F(2, 111) = 43.99, p < .001, r = .52$, as well as a substantively significant test type effect, $F(1, 111) = 56.82, p < .001, r = .57$. Another significant method x test type interaction was indicated here, $F(2, 111) = 6.38, p < .05$, but with a small effect size, $r = .21$. As a significant method effect was obtained here, Scheffé post-hoc multiple-range tests were performed to pinpoint the location of the differences among groups (see Table 4 for details).

The Scheffé test on the retention scores of Post-test 1 (words-in-isolation post-test) in the mode of ICL indicated that there was a significant difference between the mean scores of the meaning-given and meaning-inferred with MC methods, as well as between the mean scores of the meaning-inferred with MC and pure meaning-inferred methods. No significant difference was revealed between the mean scores of the meaning-given and pure meaning-inferred methods, although the mean retention score of the latter was slightly higher than that of the former.

A similar pattern was observed in the retention scores of Post-test 2 (words-in-context post-test) in the mode of ICL with a second Scheffé test. The mean retention score of the meaning-inferred with MC method was significantly higher than those of the meaning-given and pure meaning-inferred methods. However, no significant difference was shown between the retention scores of the meaning-given and pure meaning-inferred methods, although the retention score of the former was a little higher than that of the latter.

As depicted earlier, a substantively significant test type effect and a significant method x test type interaction with a small effect size were found apart from a substantively significant method effect (see Table 5 for details). The mean retention score of each method in Post-test 2 (words-in-context) was generally higher than that of the corresponding method in Post-test 1 (words-in-isolation). Thus, the pattern of the mean retention scores in these three methods in Post-test 2 remained similar to that in Post-test 1, suggesting that the effect of method persisted in different types of post-test. Finally, the meaning-inferred with MC method led to the best retention in both post-tests in the mode of ICL.

4.2.2. Discussion

The findings in the mode of ICL in this study provide strong support for the argument that elaborateness of processing plays a critical role in facilitating retention [5]

[6] [11]. The most plausible explanation for these may be that in the mode of ICL, subjects tended to pay no attention or little peripheral attention to the target words, thus the elaborateness of processing appeared to help a lot with the retention of word meanings. In terms of the meaning-given method, the meanings of target words were given to them directly, which might have made the subjects ignore these words in order to accomplish the reading comprehension task. The inferiority of the pure meaning-inferred method here could be attributed to the inherent disadvantages of the method itself, which may have led to wrong inferences and in turn led to wrong retentions. The finding of the substantively significant superiority of the meaning-inferred with MC in this study may further support the argument that the MC procedure tends to reduce the chance of inferring wrong meanings (through pure meaning-inferred method) and thus compensates for the limited information provided by the context itself.

However, this runs counter to Watanabe's [16] findings that there was no significant difference between the meaning-inferred with MC and meaning-given methods in ICL. This difference could possibly be ascribed to the selection of the distractors used in the MC procedure in which the distractors 'should not be too plausible nor too implausible' [5]. The assessment of the appeal of distractors is not an easy job, which may in turn affect the retention results of MC procedure. The veracity of the foregoing reasons remains to be verified in future research.

The finding of no significant difference between the meaning-given and pure meaning-inferred methods in ICL corresponds to many other studies reviewed in Section 2: for example, Mondria's and Watanabe's studies [9] [16]. However, according to the theory of elaborateness of processing, it has always been expected that the meaning-inferred method would outperform the meaning-given method in terms of the retention of word meanings, since it seems to be assumed that deeper elaborative processing may create all kinds of links among the word meanings, the context, and the knowledge of the learner [9]. The explanation for the unexpected result of no significant difference between these two methods may be sought in the pure meaning-inferred method itself. The first explanation is that in the mode of ICL, since no attention or little peripheral attention may have been paid to the target words, the subjects may have ignored those words and mainly concentrated on the reading comprehension task, or that the subjects could have inferred the meanings of the target words but with wrong inferences, which finally led to worse retention. The second explanation may be that the quantity of the memorizing activity in the meaning-inferred method is less than that in the meaning-given activity, or the quality of the memorizing activity in the former is worse than that in the latter [9]. Although in the mode of ICL, it is rare for the subjects to memorize the target words on purpose, it cannot be ignored that there exists the possibility for them to memorize words information unconsciously. Therefore, the inferring process tends to be

more effort-consuming and it may affect the memorization of the target information, which may affect the effect of elaborative processing.

As depicted earlier, a substantively significant test type effect was revealed as well as a significant method \times test type interaction but with small effect size in the present study. The mean retention score of each method in the word-in-context post-test was generally higher than that of the corresponding method in the word-in-isolation post-test. It can be suggested that the rich information provided by the context may lead to better retention. The repetition of presenting the same context to the subjects may help retention of word meanings to some extent. Although with only one more chance of reactivation of the target words in the design of the present study, better results were indicated. Hence, the reactivation and rehearsal of new information may be helpful in terms of retention of word meanings.

To summarize, the great effect of the rich and elaborative processing has been shown in the mode of ICL with the result that the best retention was led by the meaning-inferred with MC method. However, due to the inherent disadvantages -- possibly less quantity and inferior quality of memorization of the pure meaning-inferred method -- it failed to be superior to the meaning-given method in ICL. This suggests that the elaboration may not automatically lead to better retention. Lastly, a significant test type effect and a significant method \times test type interaction may suggest that reactivation and rehearsal of the new information could lead to better retention in the mode of ICL.

4.3. Results and Discussion of Research Question 3

Which of the three methods leads to the best retention of word meanings in the mode of intentional learning (ITL)?

4.3.1. Results

Mean retention scores of the two post-tests (1 & 2) in the mode of ITL are displayed in Table 7. To determine the effect of each factor, retention scores were then submitted to another 3×2 analysis of variance (two-way mixed ANOVA), with method as the between-subjects factor (meaning-given, meaning-inferred with MC, pure meaning-inferred) and test type as the within-subjects factor (post-tests 1 & 2) (see Table 8). In order to evaluate the strength of the findings here, the effect size r of each effect (namely the effects of method, test type, and method \times test type) was then tabulated (a threshold of .5 was used here, see Table 9 for details).

The ANOVA results indicated a substantively significant method effect, $F(2, 108) = 84.50, p < .001, r = .66$. A significant test type effect, $F(1, 108) = 35.39, p < .001$ was found with a comparatively smaller effect size, $r = .49$ (compared with the threshold of .5). Another significant method \times test type interaction was revealed here, $F(2, 108) = 12.11, p < .001$, but with only a small effect size, $r = .30$. As a substantively significant method effect was obtained here, Scheffé post-hoc multiple-range tests were then

administered to pinpoint the location of the differences among groups (see Table 7 for details).

Table 7. Number of subjects, mean retention scores, standard deviations of post-tests 1 & 2 in the mode of intentional learning (ITL)

Methods	n.	Post-test 1		Post-test 2	
		M	SD	M	SD
Meaning-given	36	6.3611 ^a	2.4975	7.4444 ^a	2.0763
Meaning-inferred with MC	37	5.5135 ^a	1.9239	6.9730 ^a	1.7398
Pure meaning-inferred	38	2.2105 ^b	1.3388	2.1053 ^b	1.3313

Note. The possible maximum score was 10. n refers to the number of subjects in each group. Means in the same column that do not share the same letter differ significantly at $p < .05$ by Scheffé test. In Post-test 1, for example, there was a significant difference between the meaning-given and pure meaning-inferred methods, whereas there was no significant difference between the meaning-given and meaning-inferred with MC methods.

Table 8. ANOVA on the retention scores of post-tests 1 & 2 in the mode of intentional learning (ITL)

Source	Sum of Squares	df	Mean Square	F	p
Between-subjects					
Method	988.192	2	494.096	84.495	.000
S within-group error	631.546	108	5.848		
Within-subjects					
Test type	36.622	1	36.622	35.390	.000
Test type × Method	25.061	2	12.530	12.109	.000
Test type × S Within-group error	111.759	108	1.035		

Note. S=subjects.

Table 9. The effect size r for ANOVA on the retention scores of post-tests 1 & 2 in the mode of intentional learning (ITL)

ω_{Method}	ω_{Testtype}	$\omega_{\text{Testtype*Method}}$
.6552	.4863	.3016

Note. A threshold of .5 of effect size r was adopted here. For instance, the effect of method had a large effect whereas the effect of the interaction of test type and method had a small effect.

The Scheffé test on the retention scores of Post-test 1 (words-in-isolation post-test) in the mode of ITL indicated that the mean score of the pure meaning-inferred method differed significantly from those of the meaning-given and meaning-inferred with MC methods, but no significant difference was found between the mean scores of the latter two methods, even though the mean score of the

meaning-given method was slightly higher than that of the meaning-inferred with MC method.

A similar pattern was observed in the retention scores of Post-test 2 (words-in-context post-test) in the mode of ITL with another Scheffé test. There was a significant difference between the mean scores of the meaning-given and pure meaning-inferred methods, as well as between those of the meaning-inferred with MC and pure meaning-inferred methods. However, no significant difference was revealed between the mean scores of the meaning-given and meaning-inferred with MC methods, although the mean score of the meaning-given method was slightly higher than that of the meaning-inferred with MC method.

As delineated earlier, a significant test type effect with slightly smaller effect size ($r = .49$ compared with the threshold of .5) and a significant method × test type interaction with a small effect size were indicated in addition to a substantively significant method effect (see Table 8 for details). The mean retention score of each method in Post-test 2 (words-in-context post-test) was generally higher than that of the corresponding method in Post-test 1 (words-in-isolation), except that the mean score of the pure meaning-inferred method in Post-test 2 was slightly lower than that in Post-test 1. Furthermore, the pattern of the mean retention scores in these three methods in Post-test 2 remained similar to that in Post-test 1, suggesting that the effect of method persisted in different types of post-tests. Finally, the meaning-given method led to the best retention in both post-tests in the mode of ITL, despite no significant difference from that of the meaning-inferred with MC method.

4.3.2. Discussion

A substantively significant method effect was found in the mode of ITL, which suggests that different word-learning methods may lead to different levels of retention. First, the significant difference between the meaning-inferred with MC and the meaning-given method in the mode of ICL disappeared in the mode of ITL, with similar findings to be found in Hulstijn's study [5]. The plausible explanation is that the difference between the meaning-given and meaning-inferred with MC methods might be washed out by the effect of rehearsal and memorizing techniques of the mode of ITL [5]. In the mode of ITL, the subjects' focal attention might have been paid to the target words and more mental effort might have been made to memorize the word meanings. However, due to the typical characteristics of the meaning-inferred with MC and pure meaning-inferred methods, more mental effort might have been made to infer the word meanings while memorizing them. Hence, the meaning-given method appeared to be superior to the other two methods. However, it may be the elaborateness of the processing and the advantages of MC procedure of the meaning-inferred with MC method that made no significant difference between the meaning-given and meaning-inferred with MC methods,

even though the mean retention scores in both post-tests of the former were slightly higher than those in the latter.

Although the superiority of the meaning-inferred with MC method in the mode of ICL disappeared in the mode of ITL, it still remains a significant difference from the pure meaning-inferred method, the explanation for which might be the inherent disadvantages of the latter and the help of multiple choices of the former besides context itself, as has been discussed in Section 4.2.2, which further supports the argument that MC procedure may help reduce the inherent disadvantages of the pure meaning-inferred method.

Different from the findings in the mode of ICL, the mode of ITL may lead subjects to pay focal attention to the target words for the following vocabulary post-tests. Hence, in the meaning-given method, subjects might have focused most of their attention on memorizing the word meanings which were given to them directly. However, in terms of pure meaning-inferred method, no cue for the target words was given; therefore, subjects might have ignored some of the target words while doing the reading comprehension task. Even if they had paid attention to the target words, more mental effort might have been made to infer the word meanings, which may have produced some wrong inferences which may have been retained afterwards. Therefore, some amount of input modifications may help to some extent with the retention of word meanings in the mode of ITL.

As delineated earlier, it is indicated in the mode of ITL in the present study that there was a significant test type effect with marginally medium effect size as well as a significant method \times test type interaction but with small effect size. The mean retention scores of the meaning-given and meaning-inferred with MC methods in the word-in-context post-test were generally higher than those of the corresponding methods in the word-in-isolation post-test, which is similar to the findings of both significant effects in the mode of ICL in the present study. It can be argued that the information richness provided by the context may lead to better retention. However, an exception appeared in the present study that, with respect to the pure meaning-inferred method, showed that the mean retention score in Post-test 2 was a bit lower than that in Post-test 1 (see Table 7). One plausible explanation is that guesswork played a role in the pure meaning-inferred method, which led to wrong inferences. The subjects who presented the right answers in Post-test 1 might have given wrong answers in Post-test 2 due to guesswork and not based on the memorizing of meanings inferred successfully. The veracity of the reasons above remains to be verified in future investigations.

To sum up, the effect of rehearsal and memorizing techniques of the ITL may negate the difference between the meaning-given and meaning-inferred with MC methods, since focal attention to the target words (meaning-given method in ITL) and the elaboration of processing (meaning-inferred with MC method) may both lead to good retention of word meanings. In addition, the disadvantages

of the pure meaning-inferred method may be reduced by a MC procedure. However, the meaning-given method led to significantly higher retention than the pure meaning-inferred method did, which may be mainly due to the inherent disadvantages, the possible less quantity of memorization, the possible inferior quality of memorization of the pure meaning-inferred method (as discussed in Section 4.2.2), and the typical characteristics of ITL mode itself (e.g. rehearsal and memorizing techniques). Lastly, a significant test type effect and a significant method \times test type interaction may suggest that reactivation and rehearsal of the new information could lead to better retention in the mode of ITL, as well as in the mode of ICL.

5. Conclusion

5.1. Summary of the Findings of the Present Study

This study was designed to address three main research questions: (1) does intentional vocabulary learning lead to better retention on word meanings than incidental vocabulary learning? (2) Which of the three methods leads to the best retention of word meanings in the mode of ICL? (3) Which of the three methods leads to the best retention of word meanings in the mode of ITL? The findings of the present study can be summarized as follows:

First, with regard to the first research question comparing the two modes of ICL and ITL, the major findings can be summarized by two points: (1) On the whole, the mode of ITL led to significantly higher retention than the mode of ICL did. (2) However, in terms of different word-learning methods, different results appeared. With respect to the meaning-given method, the mean scores of both post-tests in the mode of ITL were significantly higher than those in the mode of ICL, which was consistent with the general finding; however, there was no significant difference between the two modes in terms of the meaning-inferred with MC method. Conversely, in terms of pure meaning-inferred method, the mean score of Post-test 1 in the mode of ITL was lower than that in the mode of ICL, which was marginally significant; in addition, the mean score of Post-test 2 in the former was significantly lower than that in the latter. This indicates that with regard to word-learning methods, the comparative results of the two modes may be different, which may further suggest that attention to the target words (i.e. the mode of learning) may not be the only factor which affects the retention of word meanings, but that the word-learning method may also play some role.

Second, with respect to the second research question comparing the three word-learning methods in the mode of ICL, the major findings are: (1) A substantively significant method effect was revealed and the meaning-inferred with MC method led to the best retention of word meanings, but no significant difference was found between the meaning-given and pure meaning-inferred methods. (2) A substantively significant test type method was indicated as

well as a significant method \times test type interaction but with a small effect size. The patterns in both post-tests were similar. Additionally on the whole the retention scores were higher in the word-in-isolation post-test (Post-test 1) than those in the word-in-context post-test (Post-test 2). These findings may suggest that without or with some peripheral attention to the target words in the mode of ICL, the elaborateness of process and reactivation of new information may help with the retention of word meanings.

Third, in terms of the last research question comparing the three word-learning methods in the mode of ITL, the major findings are: (1) A substantively significant method effect was revealed. There was a significant difference between the meaning-given and pure meaning-inferred methods, as well as between the meaning-inferred with MC and pure meaning-inferred methods. But the significant difference between the meaning-given and the meaning-inferred with MC methods in the mode of ICL disappeared in the mode of ITL. On the whole, the meaning-given method led to the best retention of word meanings, but with no significant difference from the meaning-inferred with MC, although the mean scores in the two post-tests of the latter were slightly lower than those of the former. (2) A significant test type effect with slightly smaller effect size than the threshold r of .5 and a significant method \times test type interaction with small effect size were indicated in the mode of ITL. The patterns in both post-tests were similar. In addition, the mean scores in Post-test 2 in terms of the meaning-given and meaning-inferred with MC methods were generally higher than those in Post-test 1; however, the opposite result appeared in terms of the pure meaning-inferred method. The findings may suggest that with focal attention, the significant difference between the meaning-given and the meaning-inferred with MC methods may be washed out by the rehearsal and memorizing techniques (i.e. the mode of ITL).

5.2. Pedagogical Implications

Based on the findings, the pedagogical implications of the present investigation are apparent. To begin with, the general superiority of ITL may suggest that ITL must have a place in the second language curriculum, 'complementary to (not instead of) the well-established principles of incidental and contextual learning' [12], the rehearsal and memorizing techniques of which do help the retention of word meanings to some extent. The more attention paid to the different features (e.g. semantic and orthographic, etc.) of target words, the more likely it is that the new lexical information will be retained. Therefore, it may be helpful that teachers teach vocabulary explicitly in some ways by designing different tasks to lead students to focus on the lexical information while doing other language activities. The various comparison results in terms of word-learning methods between the two modes could indicate that attention may not be the only element leading to good

retention, but also different input modifications which guide learners to process the lexical information more elaborately appear to have an effect on the retention. Thus, the two learning modes of ICL and ITL may coexist, alternate with each other, and compensate for each other's disadvantages.

Second, in terms of ICL, the findings of the current investigation could indicate that the elaborateness of processing with the help of cues (input modifications, here, refers to the MC procedure) could compensate for the inherent disadvantages of ICL. Since it is generally agreed that ICL makes major contributions to the acquisition of a large proportion of vocabulary [19] [21], the design of input modifications aiming at guiding learners to engage more deeper mental processing in lexical information may promote the efficacy of this mode, which is generally claimed to lead to very low retention. Therefore, if teachers assign student lots of extensive reading to complete outside of class for the sake of vocabulary acquisition, some appropriate input modifications of reading materials guiding students to invest more mental effort to target words might help them with better retention. Furthermore, teachers may help students reactivate the new lexical information using various kinds of language activities to consolidate the retention. However, one more thing that requires attention is the appeal of distractors used in MC procedure, as Hulsitjn [5] points out that teachers play a very important role in deciding which distractors are too plausible or too implausible for the students. If students work on their own out of class without the teachers' direction, other input modifications (e.g. meaning-given) may be given to them, apart from MC procedure which needs teachers' special design.

Third, with regard to ITL, the role of attention and rehearsal has been shown clearly in the findings. It may be suggested that 'when the efficiency of the vocabulary acquisition is the main objective, the meaning-given method is preferable to the meaning-inferred method' [9]. In addition, the meaning-given method has less crucial requirements for the context and no requirements for the inferring skills of students [1].

To summarize, it has been suggested that it is crucial for teachers to balance the use of the two learning modes, input more modifications directing students to process the lexical information more elaborately, and put more emphasis on the functions of rehearsal and reactivation of new lexical information.

5.3. Limitations of the Present Study and Suggestions for Future Research

5.3.1. Limitations of the Present Study

In terms of limitations, this study is liable to the following restrictions. First, no pre-tests were given in the present study and this may have affected the validity of the equality of variances. Due to time limits, this study failed to give pretests to all the subjects in the main study to test

their language proficiency beforehand. Therefore, all the variances were assumed to be equal based on the superficial and objective conditions of subjects without pretests, which might have affected the validity of the results to some extent. Second, students might have made wild guesses on post-tests for fear of poor grades, due to the limited exposure to the target words that probably only measured students' receptive knowledge, which in turn might have affected the experimental results. Third, also due to time limits, the research failed to provide subjects with delayed post-tests, which might have only measured the retention based on short-term memory. Delayed post-tests could have measured students' long-term memory, which may have produced a different picture from that drawn by the present investigation.

5.3.2. *Suggestions for Future Research*

Future research could investigate the retention on other word features, apart from the semantic feature in this study, such as morphonological, prosodic, and pragmatic features. Additionally, the validity of the conclusions in the present study could be investigated for students at lower levels, who might have more difficulty with inferring and memorization as well.

References

- [1] Huckin, T. and Coady, J., Incidental vocabulary acquisition in a second language: A review, *Studies in Second Language Acquisition*, 1999, 21:181-193.
- [2] Lewis, M., *The Lexical Approach*. Hove: Language Teaching Publications, 1994.
- [3] Fraser, C. A., Lexical processing strategy use and vocabulary learning through reading, *Studies in Second Language Acquisition*, 1999, 21:225- 241.
- [4] Fukkink, R. G., Blok, H., & de Glopper, K., Deriving word meaning from written context: A multicomponential skill, *Language Learning*, 2001, 51:477- 496.
- [5] Hulstijn, J. H., Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In *Vocabulary and Applied Linguistics*, P. J. L. Arnaud & H. Bejoint, Eds. London: Macmillan, 1992, 113-125.
- [6] Berlyne, D. E., Carey, S. T., Lazare, S. A., Parlow, J., & Tiberius, R., Effects of prior guessing on intentional and incidental paired-associate learning, *Journal of Verbal Learning and Verbal Behavior*, 1968, 7:15-32.
- [7] Forlano, G., Hoffman, M. N. H., Guessing and telling methods in learning words of a foreign language, *Journal of Educational Psychology*, 1937, 28:632-636.
- [8] Luppescu, S., & Day, R. R., Reading, dictionaries, and vocabulary learning, *Language Learning*, 1993, 43:263-287.
- [9] Mondria, J. A., The effects of inferring, verifying, and memorizing on the retention of L2 word meanings, *Studies in Second Language Acquisition*, 2003, 25:473-499.
- [10] Craik, F. I. M. & Lockhart, R. S., Levels of processing: A framework for memory research, *Journal of Verbal Learning and Verbal Behavior*, 1972, 11:671-684.
- [11] Eysenck, M.W., Incidental learning and orienting tasks. In *Handbook of Research Methods in Human Memory and Cognition*, C. R. Puff, Eds. New York: Academic Press, 1982, 197-228.
- [12] Hulstijn, J. H., Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In *Cognition and Second Language Instruction*, P. Robinson, Eds. Cambridge: Cambridge University Press, 2001, 258-286.
- [13] Nagy, W. On the role of context in first- and second-language vocabulary learning. In *Vocabulary: Description, Acquisition and Pedagogy*, N. Schmitt & M. McCarthy, Eds. Cambridge: Cambridge University Press, 1997, 64-83.
- [14] Herman, P. A., Anderson, R. C., Pearson, P. D., & Nagy, W. E., Incidental acquisition of word meaning from expositions with varied text features, *Reading Research Quarterly*, 1987, 22:263-284.
- [15] Bialystok, E., Inferencing: Testing the "hypothesis-testing" hypothesis. In *Classroom Oriented Research in Second Language Acquisition*, H. W. Seliger & M. H. Long, Eds. Rowley, MA: Newbury House, 1983, 104-123.
- [16] Watanabe, Y., Input, intake, and retention: Effects of increased processing on incidental learning of foreign language vocabulary, *Studies in Second Language Acquisition*, 1997, 19:287-307.
- [17] Mackey, A. & Gass, S. M., *Second Language Research: Methodology and Design*. Mahwah, NJ: London: Lawrence Erlbaum Associates, 2005.
- [18] Laufer, B., How much lexis is necessary for reading comprehension? In *Vocabulary and Applied Linguistics*, P. J. L. Arnaud & H. Bejoint, Eds. London: Macmillan, 1992, 126-132.
- [19] Krashen, S. D., We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis, *The Modern Language Journal*, 1989, 73:440-464.
- [20] Ellis, R., Factors in the incidental acquisition of second language vocabulary from oral input: A review essay, *Applied Language Learning*, 1994, 5:1-32.
- [21] Schmidt, R., Attention. In *Cognition and Second Language Instruction*, P. Robinson, Eds. Cambridge: Cambridge University Press, 2001, 3-32.