International Journal of Linguistics, Literature and Translation
ISSN: 2617-0299 (Online); ISSN: 2708-0099 (Print)
DOI: 10.32996/ijllt
Journal Homepage: www.al-kindipublisher.com/index.php/ijllt

IJLLT
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

How do Adjunct Input Strengthen L2 Intentional Vocabulary Learning?

Long Fei
Associate Professor, School of Foreign Languages, Harbin University of Commerce, Harbin, China
Corresponding Author: Long Fei **E-mail**: fayeloong@126.com

| ABSTRACT
This study used a pretest-posttest-delayed posttest design at one-week intervals to determine the extent to which gloss (G), gloss with one example sentence (GE) and dictionary (D) adjunct input contributed to intentional vocabulary learning. Fifty senior high school students learning English were assigned to three experimental groups according to their English diagnostic test. Knowledge breadth (form-meaning connection) and depth (synonym discrimination, derivation production, collocation production) were both absorbed in our measurement. The results showed that they all enhance more gains on both immediate and delayed posttests. Furthermore, synonym discrimination stayed especially constant in memory than collocation production. Although no significant differences were found between the three groups in knowledge depth overall, GE showed better retention in collocation production while D contributed more to derivation production and synonym discrimination at length. Intriguingly, a negative relationship was found between familiarity with vocabulary knowledge and vocabulary learning of synonym discrimination. The study provides further support for the use of various degrees of adjunct input for intentional language learning. The results are discussed along with students' feedback from the questionnaire.

| KEYWORDS
L2 vocabulary learning; intentional learning; gloss; dictionary

| ARTICLE INFORMATION
ACCEPTED: 01 January 2025 **PUBLISHED:** 11 January 2025 **DOI:** 10.32996/ijllt.2025.8.1.8

1. Introduction
L2 (second language) vocabulary knowledge can be acquired both intentionally and incidentally. It is a common belief that the vast majority of L1 words are learned incidentally (Nation, 2001; Webb & Nation, 2017). Thus, abundant studies have given insights into incidental learning in L2 as well. However, the limited amount of L2 input suggests that intentional learning might be responsible for most vocabulary growth (Cobb, 2007; Laufer, 2003, 2005; Webb, 2008a). Due to the scarcity of intentional learning research, the argument continues for differences in their learning outcomes and the effect of intentional learning needs to be further explored (Cobb, 2016; McQuillan, 2016; Zhang, 2020).
Dictionary is an imperative intentional learning strategy as an adjunct input (Hulstijn, 1996; Nation 2001). Webb & Nation (2017) proposed that it is probably best used with words that are already partially known. Scholars have argued that providing fewer senses in the dictionary and always giving the core meaning would result in better retention of the meaning of a word. It would be of great value to investigate whether different degrees of adjunct input result in distinct vocabulary learning gains of partially known words accordingly. Gloss, as another adjunct input, is easier to access than dictionary entries with a brief explanation in a text (Hulstijn, Hollander, & Greidanus, 1996).
Many studies emphasized consequences of different forms, types or modes of them separately (Zandieh, 2012; Knight, 1994; Jung, 2016; Zhang, 2021; Huang, 2013; Ko, 2012; Zhang, 2020; Yanagisawa, 2020). Only a few studies highlight the differences in effect between them (Hulstijn, 1996; Levine, 2004). Additionally, they were all considered under incidental learning conditions. Therefore, to fill this gap, this study takes gloss as a modest degree of input and gloss with one example sentence to a word (hereinafter referred to as GE) as a moderate one to investigate distinctions brought from intentional learning.
When it comes to measuring learning gains as thoroughly as possible, both the breadth and depth of word knowledge are required. Most researchers are concerned about vocabulary breadth with form recognition and form-meaning connection

measures (Hill, 2003; Jung, 2016; Malone, 2018; Zhang, 2021). There is a scarcity of focus on vocabulary learning in terms of its depth, as Zhang (2020) suggested that future studies may adopt more dimensions of vocabulary knowledge to examine the effects (of dictionary use) beyond forming-meaning mappings.

The current study explores intentional vocabulary improvement of partially known words with various adjunct inputs to determine their effects and differences. Four aspects of vocabulary knowledge (form-meaning connection, synonym discrimination, derivation production and collocation production) were measured to provide a profound assessment of retention, especially the depth of it from the last three measures. Earlier research found that the frequencies of word exposure can increase vocabulary learning (Knight, 1994; Hulstijn, 1996; Webb, 2015; Malone, 2018). Given that more frequencies imply more familiarity. Finally, the relationship between the degree of the grasp of the known words and vocabulary gains was examined.

2. Background

2.1 Adjunct input of gloss and dictionary

Textual input of a second language (L2) has long been recognized as essential in vocabulary acquisition (Zhang, 2021). In vocabulary learning activities, gloss is a meaning-focused input from reading and dictionary is widely acknowledged as an intentional learning strategy. (Nation 2001; Webb & Nation 2017; Zhang, 2020).

A vocabulary gloss is a short definition or an explanation of the word (Jacobs et al.,1994; Nation, 2001), in different forms, languages, locations and media (Zhang, 2021; Jung, 2016; Ko, 2012; Yanagisawa, 2020).

Dictionary intends to provide more sufficient and complete contextual cues of one word with instances of its word families, collocations, phrases, and illustrative sentences. Two major kinds of receptive use of it are gaining meanings as an incidental aid and learning words as an intentional input (Webb & Nation, 2017). Some scholars reckoned that it is usually more effective (Hulstijn etal. 1996; Hill, 2003) while the other considered it inefficient (Zhang, 2021). Most studies laid emphasis on gloss or dictionary use under incidental learning separately (Zandieh, 2012; Knight, 1994; Jung, 2016; Zhang, 2021; Huang, 2013; Ko, 2012; Zhang, 2020; Yanagisawa, 2020). Only a few empirical studies provided evidence of distinctions between them (Hulstijn, 1996; Levine, 2004). This is to ignore the fact that they all took dictionary use as optional.

Dictionary, an important intentional learning strategy, probably can be best used with words that are already partially known (Webb & Nation, 2017). Furthermore, it was partially proved that better retention of the meaning of a word may result from fewer senses in the dictionary and always the core meaning (Lowie, 2003). In addition, although the degree of explicitness in the provided glosses also differs across the studies, no comparisons focus on the degree difference itself either (Jung, 2016). Thus, it would be novel and innovative to combine them together in terms of the degree of input to fill the gap regarding their comparisons with intentional learning of partially known words.

2.2 The measure of vocabulary language learning

Since most studies investigate gloss and dictionary effects on incidental learning conditions separately, learning outcomes with them are solely as well. For instance, Jung (2016) studied gloss and Knight (1994) focused on dictionary use. Yanagisawa (2020) and Zhang (2021) made a meta-analysis to explore the overall effects of different glosses while Zhang (2020) made a meta-analysis to find out that of different dictionaries. In these cases, measures of effect are all taken on incidental learning of new words. There are relatively few studies of L2 vocabulary growth of priorly known words (Webb & Nation, 2017). Current studies concerning prior knowledge highlight the different level between groups instead of the knowledge growth itself within group. Feng (2020) and Webb (2015) are both the cases investigating the former question.

With regards to the dimensions of measurement, there are mainly two different divisions. First, receptive knowledge (words that are recognized when heard or read) and productive knowledge (words that can be called to mind and used in speech or writing) with the former being greater. Second, breadth of word knowledge and depth of word knowledge. Vocabulary breadth primarily involves form-meaning mapping knowledge. Vocabulary depth denotes how well a word is known as it might involve knowledge of paradigmatic relations (antonymy, synonymy, hyponymy, gradation) and syntagmatic relations (collocational restrictions) (Milton, 2009; Haastrup, 2000; Zhang, 2020). Most researchers are concerned about vocabulary breadth with form recognition and form-meaning connection measures (Hill, 2003; Jung, 2016; Malone, 2018; Zhang, 2021).

Fruitful studies provide evidence that they both boost language acquisition and reading comprehension (Knight, 1994; Hulstijn, 1996; Zandieh, 2012; Jung, 2016; Webb, 2015). The overall outcome of vocabulary learning would be strengthened by frequency of occurrence (Feng, 2020; Webb, 2015; Hulstijn, 1996). Only one study (Dai et al.2019) investigated whether dictionary use would help learners acquire vocabulary depth knowledge (collocational knowledge in their case). Zhang (2020) suggested that future studies may adopt a dimensions approach to examine the effects in improving other aspects of vocabulary knowledge beyond form-meaning mappings.

3. Research questions

1. To what extent do different degrees of adjunct inputs increase vocabulary learning gains (form-meaning connection, synonym discrimination, derivation production and collocation production) under incidental condition?
2. Do different degrees of adjunct inputs increase learning gains in breadth and depth differently and to what extent?

3. What is the relationship between the degree of familiarity with words and vocabulary learning gains?

4. Method

4.1 Participants

A total of 50 students of grade 3 from a class of a senior high school in Chengdu, Sichuan Province took part in this study with similar ages of 17 or 18 (M = 17.46, SD = .50) and nearly averaged gender (24 females and 26 males). The participants were assigned to three groups for this study according to a recent municipal English diagnostic test (maximum score: 150) and further statistics indicated that there was no significant difference between the three groups F (2, 47) = .004, p = .996. The participants achieved moderate scores on the test (M = 101.62, SD = 14.59), indicating that they were all at an immediate level of English proficiency given that the threshold score for first-tier universities is 97. Kolmogorov-Smirnov test of normality indicated that the participants were normally distributed among groups (p >.05) based on their text scores.

An English teacher provided feedback on the target words to ensure they are just learned recently.

Prior to starting the study, verbal consent was obtained stating that involvement in the study was voluntary.

4.2 Research instruments

Target words

Due to the setting and purpose of test, target words were selected first, for they have to be in line with the following conditions. First, at least two words which are of higher extent of familiarity in participants' prior knowledge should be in paradigmatic relation with them (synonyms). This serves as a standard for synonym discrimination tests and a comparison in replying RQ3. Therefore, we compared words in the textbook with synonyms in Modern Collegiate English Synonym Discrimination and TEM-4 tests to select primary words.

Second, their derivation and collocation patterns are in accordance with participants' cognitive stage, which espouses rational learning and feedback. Third, it is without doubt that they should be newly learned and partially grasped by participants, which increased the internal validity of the study (Nation & Webb, 2011).

Ultimately, ten target words were selected including 5 verbs, 4 nouns and one adjective (attain, utilize, prolong, abolish, retort, upbringing, personnel, aftermath, illusion, artificial).

As Grains (1986) put it that retention in short term memory is not effective if the number of chunks of information exceeds seven and productive knowledge is even harder (Milton, 2009), 6 words and only 3 words out of them were adopted in derivation and collocation production test respectively.

Materials

As for materials of distinct degrees of adjunct input, which are also the independent variables in this study, dictionary materials (high degree) were printed papers provided with primitive information of the target word, its derivational words, and collocations if there exists excluding other information in Oxford Intermediate Learner's English Chinese Dictionary (8th Edition). Then, gloss materials (low degree) ended up removing the following parts based on dictionary ones. 1. All further explanation such as example sentences of target words. 2. Additional meanings (i.e., other than the meaning in the synonym test). 3. All further explanation of derivational words and collocations. Finally, based on the gloss, an example sentence concerning the meaning is attached to the word. On top of that, they are all provided in the same way in terms of language (bilingual), pattern (parts in bold), mode (textual) and manner of provision (separate printed papers).

To clarify the findings, a questionnaire was also designed utilizing five-point Likert Scale for participants to choose the extent of material's role exerted in synonym discrimination (words with partial familiarity and higher familiarity respectively), derivation and collocation learning gains. Point settings were 0-4 expressing ideas from no use to a great use. Before these four choices, there was a primary question asking if they had viewed the materials given wholly. Data with a negative answer would be excluded.

Dependent measures

1. Form-meaning connection

Section A of the test examined form-meaning connection of the single-word items with 10 multiple-choice items. The set of the options was drawn on Pavia (2019). Each items contained four meaning options, including a. the correct answer in L1; b. a distracter in word families with the target word; c. a distracter of the other target word and d. I don't know. All multiple-choice items contained such a pattern in this study (one correct answer, two distracters and I don't know). An example question is showed in (A):

2. Synonym discrimination

Section B explored synonym discrimination between newly learned words and more familiar ones with 30 multiple-choice items in the form of a sentence context. Three words (one target word and two more familiar words) were tested in every three

different items with the same three words options. Sentence questions were subtly selected or refined from TEM-4 and Modern Collegiate English Synonym Discrimination to prevent coincidence in all given materials. An example question is provided in (B):

3. Derivation production

Section C required participants to give a derivational word (change of part of speech) that occurred in the assisted materials, which served as the depth of knowledge of the target word as it indicated how well the word was known by participants. Such derivation is also a key point of study and texts in high school. Six original words were given, and they wrote behind each of them. An example question is projected in (C):

4. Collocation production

Section D, another key point in such stage, contributed to the retention of collocations in materials. Three words in primary form were given and they wrote behind each of them. An example question is illustrated in (D):

4.2 Procedure

The whole study was conducted over two weeks. First, the participants were divided in three groups with the same level according to their diagnostic test scores. Groups consisted of three experimental groups: GE (with gloss and an example sentence) (n = 18), D (with dictionary) (n = 16) and G (with gloss) (n = 16). Given that this study barely focuses on the growth within groups and differences of input degree, there is no control group. In the first week, all participants completed the pretest of four sections first. Then, they read the materials and accomplished a posttest immediately. Items in section B with target words were put at the rear in the first two tests to avoid guessing scores. A week later, participants completed the delayed posttest with a different order in each section and a questionnaire attached to it.

It rates a mention that the test sections, with section A referring to language breadth and section B-D referring to language depth, were completed separately in order to avoid reciprocal factors.

4.3 Analysis

SPSS (Version 26) for MacOS was used to analyze the data, removing all the missing data. In four test sections, all data were scored dichotomously with 1 for a correct response and 0 for an incorrect response. Scores in section B were given only in 10 items with target words. Scores were given in section C and D as long as their answers met with testing aim. To answer the first research question, repeated measures ANOVA were used to compare scores within each group at different times of testing (pretest, immediate posttest, and delayed posttest). To answer the second question, One-Way ANOVA was used to determine significant between-group differences in their vocabulary learning gains across the four sections. Before using that, analysis using One-Way ANOVA indicated that for pretests there is no significant difference between three groups in section A ($F (2, 47) = 1.32$; $p = .28$; partial $\eta2 = .053$), section B ($F (2, 47) = .47$; $p = .63$; partial $\eta2 = .02$), section C ($F (2, 47) = .77$; $p = .47$.; partial $\eta2 = .032$) and in section D ($F (2, 47) = 1.05$; $p = .36$; partial $\eta2 = .04$). Weak effect sizes may partly be due to small samples here. No violation of the assumptions of normality, and homoscedasticity had occurred in our data. Statistic in their overall pretest scores also indicated that we had successfully selected target words which they had been partly known ($M = 11.76$, $SD = 4.01$, account for 61.9% of the full score).

To answer the third research question, 10 items of target words in section B are considered as low familiarity while the others are higher ones. Another 10 out of section B were chose as our subjects in our analysis excluding much familiar and easier ones. Each item on the test was scored individually as not learned (incorrect on the pretest and incorrect on the posttests), guessed (correct on the pretest and incorrect on the posttests), known (correct on the pretest and the posttests), and learned (incorrect on the pretest and correct on the posttests). Next, the data from the not learned and learned scores were used in Pearson's r to assess the relationship between degree of familiarity and learning gains.

5. Results

The descriptive statistics of vocabulary test scores on each section of the test are presented first in Table 1. To answer the first research question, repeated measures ANOVA were used to compare scores within each group at different times of tests. To answer the second research question, One-Way ANOVA was adopted between groups to determine whether they differed significantly in their vocabulary learning gains. To answer the third research question, Pearson's r correlation coefficients between degree of familiarity and learning gains for each group at each grade level were examined for the immediate posttest and delayed posttest in section B of the test. Ultimately, the descriptive statistics of the feedback from the questionnaire are illustrated.

The pairwise comparisons between pretest and immediate posttest scores and between pretest and delayed posttest scores are shown in Table 2. Post-hoc comparisons using the Scheffe test are illustrated in Table 4.

[there should be table 1 which is in the submission file]

TABLE 2. Pairwise comparison for different sections for the test

Time of testing (*i*)	Time of testing (*j*)	Word knowledge dimension	Difference between means (*i*-*j*)	Std error	*p*	Cohen's *d*
GE		Form-meaning				
1	2	connection	-2.889***	1.367	.000	2.077
	3		-2.278***	1.994	.000	1.471
GE		Synonym				
1	2	discrimination	-2.278***	1.987	.000	1.212
	3		-3.444***	2.502	.000	1.556
GE		Derivation				
1	2	production	-1.444**	1.580	.001	1.066
	3		-1.000*	1.495	.011	0.826
GE		Collocation				
1	2	production	-1.000**	1.084	.001	1.142
	3		-.667*	1.328	.048	0.880
D		Form-meaning				
1	2	connection	-2.938***	2.144	.000	1.722
	3		-1.938**	2.294	.004	1.057
D		Synonym				
1	2	discrimination	-2.750**	2.671	.001	1.155
	3		-3.375***	2.964	.000	1.533
D		Derivation				
1	2	production	-2.188***	.910	.000	2.269
	3		-.750*	1.291	.035	0.602
D		Collocation				
1	2	production	-.188	.403	.083	1.076
	3		-.125	.342	.164	1.348
G		Form-meaning				
1	2	connection	-2.188***	1.559	.000	1.296
	3		-1.312*	1.815	.011	0.807
G		Synonym				
1	2	discrimination	-2.188**	2.257	.001	0.850
	3		-2.375**	2.579	.002	1.065
G		Derivation				
1	2	production	-1.875***	1.500	.000	1.537
	3		1.500**	1.633	.002	1.181
G		Collocation				
1	2	production	-.813*	1.223	.018	1.328
	3		-.563	1.094	.057	1.027

Note. Test time 1 = Pretest, 2 = Immediate posttest, 3 = Delayed posttest.
*p < .05, **p < .01, ***p < .001

5.1 Comparisions within groups

On the form-meaning connection section of the test, all experimental groups embodied significant differences between scores from the pretest to the immediate posttest and from pretest to the delayed posttest with strong effect size (GE: F (2, 34) = 32.82; p < .001; partial η2 = .659, D: F (2, 30) = 19.36; p < .001; partial η2 = .563, G: F (2, 30) = 15.92; p < .001; partial η2 = .515).

On the synonym discrimination section, all experimental groups, again, held significant difference between scores from the pretest to the immediate posttest and from pretest to the delayed posttest with big effect size (GE: F (2, 34) = 24.75; p < .001; partial η2 = .593, D: F (2, 30) = 16.09; p < .001; partial η2 = .518, G: F (2, 30) = 10.52; p < .001 partial η2 = .412).

On the derivation production section, experimental groups continued to vary significantly from the pretest to the immediate posttest and from pretest to the delayed posttest (GE: F (2, 34) = 8.75; p < .001; partial η2 = .34, D: F (2, 30) = 25.65; p < .001; partial η2 = .631, G: F (2, 30) = 16.58; p < .001; partial η2 = .525). GE declined in large effect size obtained previously.

On the collocation production section, experimental groups all embodied relatively slighter difference between the pretest to the immediate posttest and between the pretest to the delayed posttest with smaller effective size as well. Furthermore, D even

showed no significant difference with a weak effect size (GE: F (2, 34) = 7.93; p = .001; partial η2 = .318, D: F (2, 30) = 2.56; p = .094; partial η2 = .146, G: F (2, 30) = 4.85; p = .015; partial η2 = .244).
The pairwise comparisons between pretest and immediate posttest scores as well as pretest and delayed posttest scores indicated that for section A, B and C, there was a significant increase for all experimental groups with Cohen's d over 0.5 or 1 with strong effect size. In addition, as for section D, there was significant differences between pretest and delayed posttest for GE and G with strong effect sizes (GE: p = .001, Cohen's d = 1.14; G: p = .018, Cohen's d = 1.32). And they both showed moderate differences between pretest and delayed posttest scores (GE: p = .048, Cohen's d = 0.88; G: p = .057, Cohen's d = 1.02). D presented no significant difference between pretest and immediate posttest scores (p = .083, 95% CI [-.40, .02]) as well as pretest and delayed posttest scores (p = .164, 95% CI [-.31, .06]).

5.2 Comparisions between groups

Between groups of GE, D and G, the results showed that there was no statistically significant difference with an extremely small effect size (F (2, 47) = 0.68; p = .943; partial η2 = .003) on the overall immediate posttest. When the results for different sections of the test were considered separately, there was especially no significant difference found for synonym discrimination (F (2, 47) = .091; p = .913; partial η2 = .004) and form-meaning connection (F (2, 47) = .13; p = .879; partial η2 = .005) followed by derivation production (F (2, 47) = 1.92; p = .157; partial η2 = .076). Most differences existed in collocation production (F (2, 47) = 3.25; p =.048; partial η2 = .121) with a modest effect size.
Further results were elaborated with three fine-grained comparisons (GE-D, GE-G, D-G) on the immediate posttest here. Post-hoc comparisons using the Scheffe test for the form-meaning connection section, indicated no significant difference in three of them with slight effect sizes (GE-D: p = .972, Cohen's d = .10; GE-G: p = .879, Cohen's d = .17; D-G: p = .966, Cohen's d = .10). By the same token, there were no significant differences in synonym discrimination between them with 95% confidence interval including zero value (GE-D: p = .916, Cohen's d = .02; GE-G: p = .965, Cohen's d = .09; D-G: p = .989, Cohen's d = .04). In terms of derivation production, there was a significant difference between GE-D (Cohen's d = .76) and D-G (Cohen's d = .61) with moderate effect size. For collocation production, there was a significant difference between GE-D with a large effect size (p = .049; 95% CI [.004, 1.83]; Cohen's d = 1.08).
The results presented no statistically significant difference (F (2, 47) = .122; p = .886; partial η2 = .005) between three groups on the delayed posttest in general. Then the results for different sections of the test were illustrated individually, there was no significant difference found for form-meaning connection (F (2, 47) = .166; p = .847; partial η2 = .007), for synonym discrimination (F (2, 47) = .318; p = .729; partial η2 = .013), for derivation production (F (2, 47) = .193; p = .825; partial η2 = .008) and for collocation production (F (2, 47) = 1,77; p = .18; partial η2 = .070).
Further results were elaborated with three fine-grained comparisons (GE-D, GE-G, D-G) on the delayed posttest, post-hoc comparisons indicated no significant difference in form-meaning connection with weak effect sizes (GE-D: p = .868, Cohen's d = .18; GE-G: p = .999, Cohen's d = .01; D-G: p = .898, Cohen's d = .17), in synonym discrimination with small effect sizes as well (GE-D: p = .906, Cohen's d = .16; GE-G: p = .932, Cohen's d = .12; D-G: p = .730, Cohen's d = .28)
Results were in the same principle in derivation production. As for collocation production, there appeared to be significant differences between GE-D (Cohen's d = .79) and G-D (Cohen's d = .51) with moderate effect sizes.

5.3 Relationship between the degree of familarity and gains

Pearson's r correlation coefficients between familiarity with words and learning gains for three experimental groups were examined for the immediate posttest and delayed posttest in terms of synonym discrimination. On the immediate posttests overall, there was a moderate and negative correlation between the familiarity of words and test scores (r = -.548, n = 50, p < .05). There was also a moderate and negative correlation between them on the delayed posttest overall (r = -.524, n = 50, p <.05).
Regarding each group, they showed no significant difference on the delayed posttest. On the immediate posttest, G embodies significantly negative correlation between the familiarity of words and test scores (r = -.611, n = 16, p < .05) followed by GE (r = -.548, n = 18, p < .05). D indicated the slightest negative correlation relatively (r = -.416, n = 16, p < .05).
Descriptive statistics of the immediate posttest scores are displayed in table 3.

TABLE 3. Descriptive statistics for the synonym discrimination of learned words

	G (*n* = 16)		GE (*n* = 18)		D (*n* = 16)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Prior learned words	3.11	1.01	4.18	1.15	5.26	1.24
Newly learned words	7.50	2.85	7.28	2.08	7.63	2.28

Note. Newly learned words just refer to the target words and the full scores of the two kinds are both 10.

TABLE 4. Post-hoc comparison for different groups

Time of testing	Comparison group	Word knowledge dimension	*p*	95% confidence interval for the difference	
				Upper bound	Lower bound
2	GE-D	Section A	.972	-1.282	1.060
	GE-D	Section B	.916	-2.442	1.748
	GE-D	Section C	.185	-2.188	.316
	GE-D	Section D	.049*	.004	1.829
	GE-D	Overall depth	.941	-2.974	2.251
	GE-G	Section A	.879	-1.408	-.936
	GE-G	Section B	.965	-1.438	-1.076
	GE-G	Section C	.936	-.356	1.482
	GE-G	Section D	.621	-.559	1.267
	GE-G	Overall depth	.999	-2.661	2.564
	G-D	Section A	.966	-1.081	1.331
	G-D	Section B	.989	-2.280	2.030
	G-D	Section C	.350	-2.044	.543
	G-D	Section D	.327	-.377	1.502
	G-D	Overall depth	.958	-3.000	2.375
3	GE-D	Section A	.868	-1.042	1.598
	GE-D	Section B	.906	-2.039	1.428
	GE-D	Section C	.993	-1.251	-.1.376
	GE-D	Section D	.181	-.222	1.513
	GE-D	Overall depth	.938	-2.439	3.245
	GE-G	Section A	.999	-1.293	1.348
	GE-G	Section B	.932	-1.477	1.991
	GE-G	Section C	.891	-1.563	1.063
	GE-G	Section D	.734	-.597	1.138
	GE-G	Overall depth	.970	-2.565	3.120
	G-D	Section A	.898	-1.109	1.608
	G-D	Section B	.730	-2.347	1.222
	G-D	Section C	.843	-1.038	1.664
	G-D	Section D	.573	-.518	1.268
	G-D	Overall depth	.994	-2.799	3.049

6. Discussion

This study first expanded on earlier research by providing empirical evidence if adjunct inputs can strengthen vocabulary knowledge of form-meaning connection, synonym discrimination, derivation production and collocation recognition. On the other hand, whether different degrees of inputs boost those learning gains differently was also explored. Finally, the relationship between the degree of word familiarity and vocabulary learning gains was also examined. Additionally, the results were discussion along with the feedback from the questionnaire.

6.1 How do different degrees of adjunct inputs increase learning gains

In answer to the first question, the results indicated an overall gain in vocabulary knowledge of 27.93% in mean scores compared with the full score from pretest to immediate posttest for GE, of 21.38% for D and of 19.82% for G. The results are consistent with earlier research that gloss and dictionary have a positive effect on L2 vocabulary leaning (Knight, 1994; Hulstijn, 1996; Ko, 2012; Yanagisawa, 2020). It is without doubt that incidental leaning as a by-product activity and intentional leaning are both imperative for L2 vocabulary learning growth (Pavia, 2019).

The results from pretest to delayed posttest of the experimental groups revealed a gain of 25.48% for GE, of 21.38% for D and 19.82% for G.

In separate vocabulary knowledge sections compared with their maximum scores, GE acquired more gains in collocation production and form-meaning connection, followed by synonym discrimination and derivation production from pretest to immediate posttest.

But it acquired the most gains in synonym discrimination from pretest to delayed posttest. D obtained most gains for derivation production and form-meaning connection, followed by synonym discrimination from pretest to immediate posttest. However, it obtained the most gains in synonym discrimination from pretest to the delayed posttest. G acquired higher gains for derivation production and collocation production from pretest to immediate posttest and higher gains for derivation production and synonym production from pretest to delayed posttest.

It is apparent that overall form-meaning connection is easier to acquired which is in line with previous studies that form-oriented tasks resulted in higher gains (Hill, 2003；Zhang, 2020). Such tasks have lied in the measurement of abundant studies in terms of language breadth.

This study contributes to furnishing data adopting ampler dimensions of tests to investigate gains in more aspects. As Zhang (2020) pointed that only one study investigated whether dictionary use would help learners acquire vocabulary depth knowledge (collocational knowledge in their case) and appealed to studies in more aspects considering the inconclusive results now.

In terms of the depth of knowledge, three groups all obtained various gains in different sections. From the feedback of questionnaire, they all gave higher scores in synonym discrimination. The main reason is that section C and D are productive knowledge, which is inherently more difficult (Milton, 2009; Zhang, 2021). However, their feedback was not in line with their actual results to some extent. It is intriguing that GE intended to bring better effects in collocation production on the immediate posttest but in synonym discrimination on the delayed posttest which held water from the questionnaire (section B: M = 2.24, SD = .75; section D: M = 1.82; SD = .88). G obtained high scores in derivation production in two tests. Since the characteristic of gloss is brief and efficient, learner may incline to choose the brief information as their input as well. As for D, it acquired higher gains in derivation production and synonym discrimination. It goes along with Huang's (2013) study that students paid more attention to derivations when using dictionary. In addition, more input for a word brings about better cognition in discrimination.

In summary, we can conclude that modest degree of input (G) can strengthen knowledge depth more in both tests. Moderate input (GE) strengthens only a bit more gains in breadth on the immediate test and more gains in-depth on the delayed posttest. High degree of input (D) shows more gains in depth as well. They all showed better long-term retention in depth. Overall, they all inclined to have better retention of synonym discrimination in both short and long terms and worse in collocation production which agreed with the speculation that learners might not learn many things about the words through quick study. For example, some frequent collocates put forward by Webb & Nation (2017).

6.2 Do adjunct inputs increase learning gains differently

In answer to the second research question, the results indicated that different degrees of input exerted a similar role in form-meaning connection but brought different effects in the depth of knowledge for collocation production on both test and for derivation production on the immediate posttest. According to their questionnaire then, there were mixed results of D in terms of collocation production. They gave high scores on the questionnaire but obtain the least score and decreased the most during two tests. The empirical evidence proved Huang's (2013) finding that learners seldom used a dictionary to learn the appropriate usage of a word.

Since there are scarce studies comparing different inputs such as gloss and dictionary and barely few in terms of depth knowledge resulting from them, this study contributes to some novel empirical findings in this case.

Dai et al. (2019) found that dictionary use would help learners acquire collocational knowledge. On the contrary, GE (gloss with one example) performed the best during both tests in our case. Moreover, D (dictionary) tends to cultivate better knowledge in derivation production and synonym discrimination. It is consistent with prior study of students' preference for consulting a dictionary. Huang (2013) found that students learned derivations and additional meanings more than the usage of a word. Synonym discrimination, individually, not only kept but also increased its effects in the delayed posttest. It lends credibility to Haastrup (2000), who put forward that learners build up semantic fields by adding terms to them as they elaborate their vocabularies and creating links between words they already know and new L2 words. Moreover, central to depth of knowledge is the process of network building. G showed the least increase in this section due to its lack of information.

In summary, we can conclude that Moderate input (GE) strengthens more gains in collocation production which was contrary to D. High degree of input (D) shows more gains derivation production and synonym discrimination. G did not distinguish itself in any section.

6.3 Relationship between the degree of familarity and gains

Previous study has laid emphasis on the relationships between learners' prior vocabulary knowledge and incidental vocabulary learning. However, their prior knowledge is distinctions between groups with different language levels (Feng, 2020). In that case, our study provides a new perspective on prior knowledge. It is interesting that three groups showed a negative relationship between familiarity with the words and learning gains of synonym discrimination. Additionally, the degree of input illustrates a

similar tendency. The higher degree of input, the slighter negative relation there is. From our perspective, the following reasons may contribute to these results. First, according to Haastrup (2000), lexical progression involves a process of network building where the organizational links in the mental lexicon are strengthened, e.g., through a process of gradual differentiation within a certain lexical field. And network building is an extremely slow process with some subfields much more difficult than others. Therefore, maybe the negative result is an illustration of the stage of construction of linking. There is not enough input of new words to let them assure the discrimination from prior learned words. That lends credibility to the similar tendency of the degree of input. Second, our question setting may cause confusion during their restructure of semantic network when accommodating new words. Speculations also indicate that more research results are imperative.

6.4 Limitations and furture research

Several limitations of this study should be noted.

First, in this study the participants were in limited sample size. To produce more tenable results with more validity and reliability, in future studies, it would be of great value to increase the sample volume as much as possible.

Second, in this study, we measured gains in terms of breadth and depth. However, those measurement of depth especially the synonym section may exert a subtle influence on form-meaning connection in breadth as additional input. Finally, as for synonym discrimination section, maybe more scientific ways of testing were needed to measure different acquisition of more familiar words and newly learned words exactly. It is not out of question that results indicating a negative relationship between familiarity of learned words and gains may be influenced by the set of questions in synonym discrimination.

7. Conclusion

Overall, this study showed that three degrees of input could ensure better retention of knowledge depth in both the short and long terms. Furthermore, GE facilitated collocation production apparently, D performed better in terms of derivation production and synonym discrimination and G showed weak strength in fine sorted sections but showed obvious strength in knowledge depth overall. The results suggest that there is value in learning with all three degrees of input within the classroom. Furthermore, given that different inputs can cater to different learning aims, teachers and students can select one under corresponding circumstances thus enhancing learning effect and efficiency. Additionally, during the restructure of semantic network of newly learning words based on prior words, rich input of new words would be needed to facilitate the network budling process.

Funding: The research is supported by the Ministry of Education Humanities and Social Sciences Youth Foundation for Research Funds: A Study on Discourse Semantic Representation System and Resource Construction for Language (No. 23YJCZH150) and 2024 Heilongjiang Province Postdoctoral Research Startup Fund Project "Research on the Construction of Discourse Semantic Networks for Generative Artificial Intelligence".

Conflicts of Interest: Declare conflicts of interest or state “The authors declare no conflict of interest.”

References

[1] Cobb, T. (2016). Numbers or numerology? A response to Nation (2014) and McQuillan (2016). Reading in a Foreign Language 28, 299–304.

[2] Dai, Y., Wu, Z., & Xu, H. (2019). The effect of types of dictionary presentation on. the retention of metaphorical collocations: Involvement load hypothesis vs. cognitive load theory. International Journal of Lexicography, 32(4), 411-431.

[3] Feng, Y., & Webb, S. (2020). Learning vocabulary through reading, listening, and viewing: Which mode of input is most effective? Studies in Second Language Acquisition, 42(3), 499-523.

[4] Haastrup, K., & Henriksen, B. (2000). Vocabulary acquisition: Acquiring depth of. knowledge through network building. International Journal of Applied Linguistics, 10(2), 221-240.

[5] Hill, M., & Laufer, B. (2003). Type of task, time-on-task and electronic dictionaries in incidental vocabulary acquisition.

[6] Huang, S., & Eslami, Z. (2013). The use of dictionary and contextual guessing. strategies for vocabulary learning by advanced English-language learners. English Language and Literature Studies, 3(3), 1.

[7] Hulstijn, J., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. The Modern Language Journal, 80, 327–339.

[8] Jung, J. (2016). Effects of glosses on learning of L2 grammar and vocabulary. Language Teaching Research, 20(1), 92-112.

[9] Knight, S. (1994). Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities. The modern language journal, 78(3), 285-299.

[10] Ko, M. H. (2012). Glossing and second language vocabulary learning. Tesol Quarterly, 46(1), 56-79.

[11] Laufer, B. (2005). Lexical frequency profiles: from Monte Carlo to the real world. A response to Meara (2005). Applied Linguistics, 26(4), 582 - 588.

[12] Laufer, B. (2003). Vocabulary acquisition in a second language: Do learners really. acquire most vocabulary by reading? Some empirical evidence. Canadian Modern Language Review/Revue Canadienne des Langues Vivantes, 59, 567–587.

[13] Levine, A., Bejarano, Y. A. E. L., Carrell, P., & Vered, L. I. L. Y. (2004). Comparing. dictionary definitions and glosses in hypertext foreign language reading: Facilitating foreign language reading comprehension and vocabulary acquisition. CATESOL Journal, 16(1), 59-68.
[14] McQuillan, J. (2016). Time, texts, and teaching in vocabulary acquisition: A rebuttal to Cobb (2016). Reading in a Foreign Language, 28, 307–318.
[15] Malone, J. (2018). Incidental vocabulary learning in SLA: Effects of frequency, aural. enhancement, and working memory. Studies in Second Language Acquisition, 40(3), 651-675.
[16] Milton, J. (2009). Measuring second language vocabulary acquisition (Vol. 45). Multilingual Matters.
[17] Nation, I.S.P. (2001). Learning vocabulary in another language. Cambridge: Cambridge.
[18] Pavia, N., Webb, S., & Faez, F. (2019). Incidental vocabulary learning through listening to songs. Studies in Second Language Acquisition, 41(4), 745-768.
[19] Webb, S., & Chang, A. C. S. (2015). How does prior word knowledge affect vocabulary learning progress in an extensive reading program? Studies in Second Language Acquisition, 37(4), 651-675.
[20] Webb, S., & Nation, P. (2017). How vocabulary is learned. Oxford, UK: Oxford University Press.
[21] Webb, S. (2008a). Receptive and productive vocabulary size. Studies in Second Language Acquisition, 30(1), 79 - 95.
[22] Yanagisawa, A., Webb, S., & Uchihara, T. (2020). How do different forms of glossing contribute to L2 vocabulary learning from reading? A meta-regression analysis. Studies in Second Language Acquisition, 42(2), 411-438.
[23] Zandieh, Z., & Jafarigohar, M. (2012). The effects of hypertext gloss on. comprehension and vocabulary retention under incidental and intentional learning conditions. English Language Teaching, 5(6), 60-71.
[24] Zhang, C., & Ma, R. (2021). The effect of textual glosses on L2 vocabulary acquisition: A meta-analysis. Language Teaching Research.
[25] Zhang, S., Xu, H., & Zhang, X. (2020). The effects of dictionary use on second language vocabulary acquisition: A meta-analysis. International Journal of Lexicography, 34(1), 1-38.