

The Effect of Group Work on English Vocabulary Learning

Su-Fei Lin¹

¹ Language Centre, Chia-Yi University, Taiwan

Correspondence: Su-Fei Lin, Language Centre, Chia-Yi University, Taiwan. Tel: 886-5-271-7977.

Received: March 26, 2018 Accepted: May 8, 2018 Online Published: May 15, 2018

doi:10.5539/jel.v7n4p163 URL: <https://doi.org/10.5539/jel.v7n4p163>

Abstract

This study investigated the effectiveness of group work (GW) in EFL vocabulary learning by second year, non-English major, university students in Taiwan, in comparison with working individually (IW). The students (N=44) worked in mixed ability groups of 3-4 or in IW to complete vocabulary exercises following reading activities. The classroom intervention followed a repeated measures design with alternating sessions (one week IW, one week GW) for 12 weeks. In order to measure students' word knowledge gains, the modified vocabulary knowledge scale was used in pre-, post- and delayed-post tests, and the scores from the tests were analyzed with paired t tests. Qualitative information about vocabulary discovery and retention was further obtained from interviews with 24 students conducted after the classroom intervention. Results showed that students' overall improvement in vocabulary knowledge with group work was significantly higher than that with individual work on immediate post-tests, though both treatments had a beneficial effect. Later retention of word knowledge after GW was only 2% higher than that with IW. Interpretations and implications of these findings are discussed.

Keywords: vocabulary learning, group work, individual work, zone of proximal development (ZDP), involvement load

1. Introduction

Vocabulary development closely relates to learners' reading ability (Krashen, 2011; Nation, 2013). First, in order to understand and learn efficiently from unsimplified text, learners need to already possess a vocabulary of at least 5,000 word families, which provides coverage of 98% of authentic text (Laufer, 1992; Hirsh and Nation, 1992). This requisite coverage is for reading at a reasonable level (Laufer and Ravenhorst-Kalovski, 2010; Schmitt, Jiang and Grabe, 2011). Ways of improving students' word knowledge, to this level include glosses (Watanabe, 1997, Myong, 2012), dictionary use (Scholfield, 1997; Jee and Hyeok, 2016), direct vocabulary instruction (Schmitt, 1997; Maki and Thomas, 2015) etc. Secondly, however, reading itself, especially extensive reading (Perfetti, 2007) is seen by some as an important method to learn vocabulary. Studies have however found limited vocabulary learning through meaning-focused reading (Horst, Cobb and Meara, 1998; Waring and Takaki, 2003; Rosszell, 2007). By contrast, positive effects on students' vocabulary learning have been found from explicit activities conducted after intensive reading (e.g. matching the target word with a definition or synonym, fill-in-the-blanks exercises with target words, inferring the meaning of new words from the sentence context) (Knight, 1994; Paribakht & Wesche, 1997; Laufer, 2000; Min, 2008).

Such vocabulary activities may be performed by students either doing individual work (IW) or group work (GW). Surprisingly, however, despite the effectiveness widely claimed for group work in language learning (Ohta, 2001; Swain, 2002; Storch, 2005; Jones, 2006, Lasito & Storch, 2013; Dobao, 2014) no studies have investigated the conduct of such vocabulary activities through group work, particularly considering the impact on the discovery and retention/consolidation of lexical information separately. This study therefore attempts to investigate the effect of incorporating collaborative group work into the above mentioned activities to find out its precise influence on students' vocabulary learning, in particular considering vocabulary discovery and retention separately.

2. Literature Review

Nation (2013) recognizes three general processes by which lexical information about a word is discovered and remembered, i.e. learned. These processes consist of "noticing", "retrieval" and "creative use" (or generative use) (p.102). We now therefore review how each of these might be enhanced by either GW or IW.

2.1 Nation's Perspective on Vocabulary Learning

2.1.1 Noticing

Giving attention to at least some aspect of a new word is the first step in the process of vocabulary learning. In order to gain knowledge of a word, learners need to be aware of the word and recognize it as a useful language item (Ellis, 1991; Schmidt, 1990) removed from its sentence context rather than just “as a part of a message” (Nation, 2013, p.103). That is to say, the word need to be mentally separated from its sentence context and becomes the focus of the learners' attention as a language item about which some information need to be ‘discovered’ (Schmitt, 1997). If one does not notice that one does not know something (e.g. about a word), then one will not take steps to ‘discover’ the information that one does not know (Schmitt, 1997).

It has been recognized that reading individually is not entirely conducive to noticing new words. Often the learner's focus may be on understanding the meaning of the text, not on noticing and learning new words (Nation, 2013). Learners may pay some passing attention to new words, at a low level of consciousness, but they may skip them (as may be encouraged by the teacher in a reading class, e.g. for skimming), or give them some minimal possible meaning to support the continuation of figuring out the wider text (Grabe, 2009). Studies have therefore found low rates of vocabulary learning through individual meaning-focused reading (Horst, Cobb and Meara, 1998, Waring and Takaki, 2003; Rosszell, 2007).

According to Vygotsky's (1978, 1986) constructivist theory of learning, by contrast, an individual's learning occurs through communication with others in the social group. This would imply that more vocabulary might be noticed and learnt from reading if GW was in some way involved. Vygotsky further states that learning occurs when learners interact collaboratively in the zone of proximal development (ZPD), which is the space between what a learner can do without help and what a learner can do with help from a more competent member of the group (Vygotsky, 1978). Noticing and discovering new lexical information from reading could be seen as an activity within the ZPD, and so capable of being facilitated by a more able individual who prompts noticing and provides assistance to the novice through conversation in the group (Aljaafreh and Lantolf, 1994). Notably GW allows for more knowledgeable ‘others’ additional to the teacher to be involved in this process. However, while the general GW literature seems to suggest benefits of GW for learning, the vocabulary literature reveals little specially about GW benefits for vocabulary noticing (Ohta, 2001; Swain, 2002; Storch, 2005).

2.1.2 Retrieval

The second step in enabling a word to be learned, according to Nation, is retrieval. This primarily assists retention of what has been noticed, or what Schmitt (1997) calls ‘consolidation’. When a word is retrieved, the mental connection between word form and meaning is strengthened (Baddeley, 1990; Nation, 2013): the learner subconsciously evaluates and compares the word with other words which he is able to recognize and then chooses the one most suitable for the present context thus strengthening memory for the lexical information (Beheydt, 1987; Pavlenko, 2009).

Moreover, learning a word fully involves discovering and retaining a number of different types of lexical information (e.g. its precise spelling, sound, and part of speech). It is unlikely, however, that students will notice all of these facets of a word fully after only one exposure. Thus repeated retrieval has a second benefit that further such inherent aspects of the word may be discovered and start to be memorized. Indeed, in order to equip learners with enough information about a word to use the word accurately in production, it is suggested that a minimum of twelve exposures is required (Beck, Mckeown & Omanson, 1987).

Clearly, therefore, after the initial meeting, opportunities need to arise for learners to repeatedly meet or use a new word, so that retrieval occurs (Schmitt and McCarthy, 1997). While this may in part be determined by the teacher or the teaching materials, clearly performing vocabulary tasks via GW is likely to engender more retrieval than the same tasks done as IW.

Exposing learners to a large amount of words repeatedly purely in reading texts has been seen by some as a key way for them to learn vocabulary, since it requires repeated retrieval, but such reading is essentially an IW activity. Studies of the relationship between L2 reading and vocabulary learning have in fact shown that increasing the amount of extensive reading (where words recur, and so need to be repeatedly retrieved) does lead to noticeable vocabulary learning, but that it is slow in comparison with the amount that has to be read (Elley and Mangubhai, 1983; Ferris, 1988; Pitts, White and Krashen, 1989). Furthermore, many EFL learners, including ours who are non English majors, do not have the time or motivation to read extensively in the FL.

Thus reading supplemented with another activity enhancing vocabulary learning has been considered in several studies (Knight, 1994, Paribakht and Wesche, 1997, Laufer, 2000, 2003). Min (2008) examined the effectiveness

of reading plus vocabulary-enhancement activities (RV) versus narrow reading (NR) (repeated reading of thematically-related articles) for vocabulary acquisition and retention among EFL secondary-school students in Taiwan. The results showed that the RV group gained significantly more knowledge of target words than the NR group both on the post-test and on the delayed post-test. In these studies, however, the possible role of IW versus GW in performing the task was not considered.

In GW, learners are expected to correct others' errors and to give explanations to other group members. This should not only help the group member who lacks knowledge to discover more lexical information, but also prompt the explainer to retrieve word knowledge and reconsider the meaning of the target word in relation to the rest of words, which should enhance their learning of further aspects of word as well. GW also must help all learners retrieve through repeated speaking and listening, not just reading, and attention may be given not only to words targeted in an RV task, but also to non-targeted words. Once again, while we can deduce from the general literature the putative benefits of GW, the vocabulary literature says little about this in relation to retrieval.

2.1.3 Generative Use

The final step in learning a new word according to Nation is generative use. Generative use refers to meeting or using the previously-met word in a context that is different from the context in which the word was met previously. For example, if a student meets the word 'inherited' used as a verb in the sentence: 'Some scientists believe that a person's personality is mainly inherited', and then meets the word again in another sentence: 'He inherited a fortune from his grandmother', the learner will need to rethink the meaning and use of the word 'inherited' and this will help the student remember the word (Nation, 2013). Generative use then again enables further, more contextual, aspects of lexical information to be discovered about words (e.g. multiple meanings, collocation, associated grammatical structures), as well as at the same time promoting retention/consolidation. This is consistent with 'levels of processing hypothesis' in cognitive psychology which implies that word retention relates to the amount of attention given to, and the variety of types of manipulation applied to a new word (Craik and Lockhart, 1972; Leow and Mercer, 2015; Baddeley and Hitch, 2017).

Once again, in performing vocabulary tasks through GW, there is expected to be much less predictability about what contexts words will be used and heard in, since each member of the group has their own history of prior exposure to and use of the word, and this will be reflected in their GW interactions. By contrast in IW the learner is limited to the context provided in the learning materials and his/her own prior experience of the word. While in general it is recognized that the interaction in collaborative group work provides students with opportunities to build and practice their newly-constructed knowledge (Panitz, 1999), to our knowledge this has not been directly investigated in post-reading vocabulary tasks. The vocabulary literature again says very little about GW benefits for vocabulary generation, although the general GW literature seems to suggest such benefits of GW for learning (Ohta, 2001; Swain, 2002; Storch, 2005, Dobao, 2014).

2.2 Involvement Load

A somewhat distinct approach from the above is the 'involvement load hypothesis' about vocabulary learning. Laufer and Hulstijn (2001), who have suggested that the involvement in processing target words, and hence the benefit for learning, is affected by three task related features: need, search and evaluation. Need is the requirement for the target word in order to complete the task, such as needing a particular word to fill in the gap in a sentence correctly. "Need does not exist if the target vocabulary is not needed to complete the task. Need is moderate if the task requires the target vocabulary, and it is strong if the learner feels the need for the vocabulary, e.g. a genuine communicative task" (Nation, 2013, p.98). Clearly, when reading a text, the need for a new word that is encountered may vary depending on how far understanding it is crucial to adequately understanding the meaning of the text as a whole. In dedicated vocabulary tasks following reading, however, the need for the target words in the exercises will be much enhanced regardless of whether the task is performed as IW or GW. For the present study, the need for the target words in completing the tasks is moderate because the tasks require it, and need does not come from the learner. The need for words other than the target ones is not promoted by the task, however, and it is possible that such a need may be created where a fellow group member perceives the need to find out about a non-target word in the task which a particular learner working in IW would have overlooked.

Search is the attempt to find the target word information required by the tasks, for example, by using a dictionary to look up the meanings or forms of the target word. "Search does not exist if the word forms or their meanings are supplied as a part of the task. Search is moderate if learners have to search for the meaning of the item, and strong if learners have to search for the form to express a meaning" (ibid). In the tasks we envisage, the requisite information is not all supplied with the task. The search for the forms and meanings of target words is rated low because they are provided by the teacher in work prior to the task. However, for other aspects of target words

and for non-target words search is moderate as the learner has to discover the information. Here the search resources will differ between GW and IW since, in the former, fellow group members can provide information additionally to a dictionary.

Evaluation involves the comparison of target word information with the context required in the tasks and then deciding if the word selected fits the context appropriately. "Evaluation is moderate if the context is provided and is strong if the learner has to create a context" (ibid). In our study, evaluation is moderate since the context is provided for both parts of the tasks. However, we would argue that richer evaluation, so more involvement, may occur in GW than IW since members of the group may collectively come up with more reasons for why a word does or does not fit a context than a learner working alone.

Although some research has been done which supports the involvement load hypothesis (e.g. Laufer and Hulstijn, 2001; Hulstijn and Laufer, 2008; Kim, 2011), as far as we are aware the potential role in it of the mode of working on a task, by IW or GW, has not been considered.

2.3 Studies of IW or GW in Relation to Vocabulary

In studies of interaction in general classroom activities it is reported that learners tend to focus particularly on vocabulary and other linguistic forms (Williams, 1999, 2001; Leaser, 2004; Kim, 2008). More specifically, Kim (2008) examined the effectiveness of collaborative and individual tasks on the acquisition of L2 vocabulary by Korean-as-a-second-language learners (KSL). Thirty-two, intermediate-level, KSL speakers in a preparatory Korean language program completed a pre-test, a dictogloss task, and two post-tests over a three-week period. The dictogloss was a procedure in which a text was read to a group of learners, and while the text was being read the second and third times, learners were asked to take notes, including noting words and phrases from the text. After listening to the text three times, learners in the individual work group were required to reconstruct the text individually while thinking aloud, whereas collaborative group learners reconstructed the text with a partner. The results showed that learners who worked in the collaborative groups performed significantly better on the vocabulary tests than the learners who worked individually, even though learners in both groups exhibited similar numbers of LREs. LREs are language-related episodes, defined by Swain and Lapkin (1998) as "any part of a dialogue where the students talk about the language they are producing, question their language use, or correct themselves or others" (p. 326).

In addition, in GW activity in general, it has been shown that peers can be experts and novices at the same time: more proficient and less knowledgeable learners both may contribute knowledge to each other in order to increase the level of their performance (Donato, 1994; Anton and DiCamilla, 1998; Swain and Lapkin, 1998; Ohta, 2000, 2001). A study by Watanabe (2008) explored the interaction of adult ESL (English as a Second Language) learners with either a higher- or a lower-proficiency peer during pair problem solving, and investigated their perceptions of the interactions with their partners. The study showed that the three participants preferred working with group members who contributed their ideas, regardless of their proficiency level. The analysis of the interactive aspects of the LREs further indicated that the more proficient learners and the less proficient learner became resources for each other by repeating each other's utterances, which helped them to elaborate and to build up each other's understanding (DiCamilla and Anton, 1997). For this reason we will involve learners of unequal proficiency level in our groups for GW.

2.4 Purpose of This Study

In order to test general claims made about the benefits of GW in a specific context, this study therefore aimed to investigate the following questions: 1. Will students learn the target words better through group work than through individual work? Will that change over the period of the study? 2. What aspects of word knowledge do the students claim to discover through group work vs individual work? 3. What aspects of word knowledge do the students claim to retain better through group work vs individual work?

3. Method

3.1 Participants

The participants were 44 second year, L1 Chinese, non-English major students from various departments attending the required English reading course to improve their basic reading ability and vocabulary size for further ESP courses. There were 28 male and 16 female students, aged from 18 to 20, with pre-intermediate English proficiency as measured with the national Taiwanese GEPT test.

3.2 Overall Procedure

All 44 participants took pre-test of target word knowledge (see Appendix 3), carried out tasks in IW and GW,

took an immediate post-test, a delayed post-test, and 12 participated in interviews after the classroom intervention. For GW activities, the 44 students were arranged into groups of 3, including members with high, mid- and low prior knowledge of the target words in each group (Donato, 1994; Ohta, 2000). For IW activities in other weeks, the same 44 participants worked individually in class (see Appendix 1).

In each weekly lesson, first a reading text was introduced from the textbook Day and Yamanaka (2007), *Cover to Cover 1*, Oxford University Press. After the schema activating activities, the teacher explained the general meaning of reading text, vocabulary, and intended reading comprehension strategy to the students. Students then read the text and completed the reading comprehension exercise individually. This was followed by two vocabulary exercises provided in the textbook, which were completed either through IW or GW, and lasted for 15-20 minutes. 12 texts were used over the 12 weeks of the intervention: seven pairs of vocabulary exercises were completed with IW, five with GW in alternating weeks (See Appendix 2).

3.3 Training Sessions

While IW was very familiar to students, so not deemed to require training, this could not be assume for the collaborative GW of our study. Students were introduced to the idea of ‘collaborative’ GW as GW where students are grouped intentionally rather than self-selected, so as to have varied levels of prior knowledge, but with each member individually responsible for interacting and contributing to the tasks performed, with no specific organization or procedure imposed on the group from outside. During the GW training session in week 3, handouts on disputational talk, cumulative talk, and exploratory talk were given and explained to students (Mercer, 1995, 1996; Wegerif and Mercer, 2000). Students observed and understood how to co-construct knowledge in groups through exploratory talk, while comparing this with disputational talk, and with cumulative talk. Later the students were given a vocabulary task and asked to discuss the target words by using exploratory talk. When students were practising, the teacher went round the classroom to check on them and to encourage them to use exploratory talk during discussions.

3.4 Test Instrument

A modified version of the Vocabulary Knowledge Scale (VKS) test (Paribakht & Wesche, 1993) was used for all tests of the targeted items (Appendix 4). In this kind of test, the testee answers a series of graded questions (categories) about each vocabulary item tested, revealing how many different kinds of information about the word they know. The modification took account of the fact that, besides word form, meaning, and synonym (tested by the existing VKS), the vocabulary tasks also practised antonyms, collocations, inflections, and parts of speech. In the modified VKS self-report categories, students were therefore given the additional option (category IV) of providing the antonym of the target word. It was also stipulated that when using the target words to make sentence (category V), they should use sentences which they had not seen before and not read in the exercise, because the researcher wanted to preclude participants using the strategy of memorizing one of the example sentences from the tasks in order to answer this question.

The two pre-tests between them covered all the words to be targeted in the exercises. Immediate post-tests were given directly after each vocabulary task was finished (10-15 minutes) to check each student’s short-term retention of vocabulary information for each of the six words targeted. The delayed post-test, two weeks after finishing the intervention, was to check each student’s change in knowledge retention between pre- and post-test and covered 40 vocabulary items studied earlier in either IW or GW modes.

The derivation of scores from student test responses was again a slightly modified version of the original VKS system. Responses to the set of questions about each word are evaluated and mapped onto scores as shown in Appendix 4, yielding a score of 1-5 for each word tested. For example, if the testee chooses category II ‘I have seen this word before, but I don’t know what it means’ they are awarded a score of two for at least being familiar with the written form of the word. However, they may also receive a score of two if they stated they had knowledge of the word in a sentence: _____’, but showed that they did not know its meaning. A score of 4 was given if the word was used with the appropriate meaning and correct grammar, even if another part of the sentence contained errors (e.g. In many society, sleep deprivation is becoming part of the culture).

3.5 Interviews

Semi-structured interviews with 12 participants were conducted in weeks 14-17 (out of class time). The participants were interviewed in pairs for 30-80 minutes, and the information gathered was recorded on a digital voice recorder. They were asked their experiences and opinions concerning vocabulary discovery and retention using GW (N=24), and IW (N=24), (See Appendix 5).

3.6 Data Analysis

The test scores were converted to percent, for ease of comprehension, and analyzed with SPSS (Version 16) to compare any gains in word knowledge which the students had made through IW and GW. Interview data was fully transcribed in the original Mandarin Chinese and then translated into English. Each audio file was checked several times to ensure there was no possibility of missing any data. The analysis procedure for coding, categorization, description and interpretation suggested by Bogdan and Biklen (2007) and Patton (2002) was used to analyze the student interview data. In order to ensure the reliability of the categorization and coding (Mackey and Gass, 2011), the researcher involved three of her colleagues as second judges. The codes assigned differently from the researcher by second judges were discussed and changed if needed.

4. Results

4.1 Changes Between Pre-Test and Immediate Post-Test

Table 1. Vocabulary scores from pre- and immediate post-tests for IW

	IW 1	IW 2	IW 3	IW 4	IW 5	IW 6	IW 7
Pre-test	36.82	19.48	17.54	29.55	19.57	17.64	12.50
Immediate post-test	48.53	40.81	49.51	68.63	65.81	65.56	64.46
Improvement	11.71	21.33	31.97	39.08	46.24	47.92	51.96

Table 2. Mean vocabulary scores from pre-and immediate post-tests for GW

	GW 1	GW 2	GW 3	GW4	GW 5
Pre-test	19.19	20.74	14.05	18.99	21.41
Post-test	48.96	64.32	68.75	70.44	71.48
Improvement	29.77	44.58	54.70	51.45	50.07

As seen in Tables 1, 2, with both IW and GW immediate post-test mean scores were always higher than pretest scores, as would be expected given the amount of instruction and learning in between in both treatments. The vocabulary scores in the pre-tests for IW ranged from 12.50% to 36.82%, while the vocabulary scores in the immediate post-tests ranged from 40.81% to 68.63%. The vocabulary scores in the pre-tests for GW ranged from 14.05% to 21.41%, while the vocabulary scores ranged from 48.96% to 71.48% in the immediate post-tests.

In both treatments prior knowledge of the words as reflected in the pretests varied somewhat from week to week, as might be expected and this variation is reflected to some extent in the post-test scores for IW but not for GW. The GW post-test scores steadily rose from session to session, reaching 71.5% in the final session. The rise in posttest scores did show signs of leveling off, however, suggesting that scores would not reach 100% for a long time, if at all. This could be a sign that the richer amount of input in GW improves learners' knowledge up to a certain level regardless of their initial level of knowledge while IW simply adds a certain amount of knowledge to what was already known.

Post-test scores rose over the period of the study for both IW and GW, though the pattern is more consistently progressive over time for the latter, where we see a curve possibly tending to eventually level off at a post-test score of around 75%. More importantly improvement scores (=post-test minus pretest) rose for both treatments over sessions. This shows that both treatments had a progressive effect of increasing the pre-post improvement of knowledge over time, not just GW, despite IW being the familiar mode of working and GW the innovation. This will be discussed further below.

Crucially for our study, however, the trajectory of change in scores over time shows greater benefit from GW than from IW. The immediate post-test scores end up in the 70s for the former but not for IW. The significance of the difference was confirmed by a paired sample t test comparing overall improvement scores in GW with those in IW ($t=7.279$, $p<.001$). The result was similar to Kim's dictogloss study (2008), described earlier, which showed that learners who worked in the collaborative groups performed significantly better on the vocabulary tests than the learners who worked individually.

4.2 Delayed Post-Test Results



Figure 1. Vocabulary scores from delayed post-test for GW and IW

The mean vocabulary scores from the delayed post-tests for GW and IW showed that the students retained a little more vocabulary knowledge through GW (38%) than through IW (36%). The paired sample test showed that the mean difference (1.85) was statistically significant ($p=0.027$). This, therefore, again supported the superiority of GW, albeit by a score difference that was very small in terms of pedagogical importance. The overall most noticeable thing about the delayed post-test scores, however, was that they were considerably lower than the average immediate posttest scores for both activities (GW=64.8%, IW=57.6%). This might indicate that rapid forgetting/attrition had set in regardless of activity.

Attrition was to be expected (Jones, 2006; Kim, 2008), however, and the high rate of apparent forgetting might be because the vocabulary tested in the delayed post-test was not included in the final examination (taken by the students as part of the course requirements, but not part of the research). The final examination took place just after the delayed post-test and tested vocabulary mainly from the texts introduced in weeks 15-17, which were not included in our study. Students would have been much more motivated to review all the target words learned before if they had been included in the final examination, and this might have affected the result of the delayed post-test.

4.3 Aspects of Word Knowledge Reported as Discovered in IW vs GW

While the interviews cannot reliably tell us how often various kinds of lexical information were discovered or retained, they can give us an idea of the range of types that were considered, and the sources of information and how the information was handled.

Table 3. Examples of aspects of word knowledge discovered in IW

Aspects of word knowledge discovered in IW	Example	Students
Synonym of target words	<i>disable/crippled</i>	10A-S, 10B-S
Antonym of target words	<i>disable/strong</i>	11A-S, 9A-S
Other meaning of target words (polysemy: <i>disorder</i>)	<i>disorder/physical illness or a disorganized situation</i>	11B-S
Prefix (e.g. <i>ir-</i>)	<i>rational/irrational</i>	10A-S
Suffix (e.g. <i>-less</i>)	<i>valuable/valueless</i>	10B-S
Spelling of target words (homonymy)	<i>employee/ employer</i>	12B-S

Table 4. Example of aspects of word knowledge discovered in GW

Aspects of vocabulary discovered in GW	Example	Students
Synonym of target words	<i>plump/puffy</i>	8A-S, 8B-S, 9A-S
Antonym of target words	<i>break down/ working</i>	8B-S
Synonym and antonym of target words	<i>interrupt/disturb, stop/disconnected</i>	7A-S, 7B-S, 9A-S
Other meaning of target words (polysemy)	<i>Image/picture</i>	5B-S, 6A-S, 9A-S
Prefix (e.g. <i>un- or in-</i>)	<i>unload/load</i>	4B-S
Suffix (e.g. <i>-ed</i>)	<i>relax/relaxed</i>	8A-S
Lexical grammatical features	<i>confident/ confident of</i>	9A-S, 9B-S
Phrase	<i>unload/take down</i>	9A-S

As seen in Table 3 and 4, students reported that they discovered more types of word knowledge when working in GW than in IW. Both in IW and in GW, students reported that they found out synonyms, antonyms and meanings of target words but since the tasks were explicitly about meanings, synonyms and antonyms of target words, this is not surprising. No other lexical information about words was required by the tasks, although some was by the tests (e.g. grammatical properties of words, collocation). Aspects not strictly required by either were affixes, compounds, phrases, and pronunciation. Spelling, though required to be written in both, was supplied in both. It is therefore interesting that quite a lot of these other types of information were reported as discovered.

Knowledge of prefixes, and suffixes, which students discovered in both groups, was not required in the tasks except insofar as it might relate indirectly to antonyms, where the affix showed oppositeness (e.g. un-happy, dis-connect, care-less) and is known to help with remembering words and meanings, even though the students did not say that (the morphological decomposition strategy). The precise spelling of target words (only mentioned as being discovered in IW but inevitably discovered in fact in both modes) was also knowledge needed in relation to the tests (e.g. making sentences with target words or writing down the synonym or antonym of the target words in the tests). The other types of word knowledge reported as discovered in GW included lexical grammatical features and phrases which were not specifically required in the tasks but were relevant to the tests (Category V, making sentences with target words).

While there were not great differences in the kinds of lexical information involved, clearly the sources, and variety and quality, of information differed between IW and GW. In IW, the synonym, antonym and additional meaning of target words which students discovered were generally from the dictionary, since students depended on their own efforts to complete the tasks and the dictionary was the main source for word information (10A-S, 11A-S, 11B-S). In group work, word knowledge was obtained from the contribution of other group members (8B-S) or from the dictionary (8A-S, 8B-S, 9A-S). [8B-S] showed more active engagement with the information when he said: 'The antonym of break down is operating according to my dictionary. Group members found out that working can also be the antonym of break down. The word, working is even easier than operating. I never thought that working can be the antonym of break down, but the group members do. So I use the word working as the antonym of break down.' 7A-S showed more sources for word information when he said: 'During group work, the group members use different types of dictionary; for example, one groupmate uses cell phone as dictionary, another groupmate uses paper dictionary, the other uses electronic dictionary. The information about the synonym or antonym of words the groupmates get from their dictionaries could be different. For example, the synonym of interrupt from groupmate A is disturb, stop from groupmate B and disconnected from groupmate C. So I learn disturb, stop, and disconnected are all the synonym of interrupt from the group members. 9A-S importantly mentioned greater processing of word information through group discussion. His basic dictionary only provided meaning and pronunciation of words, and the meaning provided in the dictionary sometimes could not be applied to the sentence in the task: an example was the task of circling the word that did not belong in 'unload (par.5) unpack take down pack'. He then discussed the possible meanings of the word he got from the dictionary with group members and decided on a meaning that properly fitted the task.

In IW, students discovered the use of prefixes and suffixes from the teacher and the dictionary (10A-S, 10B-S). In group work, students again discovered prefixes and suffixes both from group discussion and from the dictionary (4B-S), but used as a shared resource. 8A-S that discovered the synonym of casually could be the word, relaxed, simply by adding -ed to relax from the dictionary before telling group members this.

Knowledge of spelling of words (different from any homonyms), lexical grammatical features and phrases were targeted in the tests: students were required to make sentences with target words and to write down words correctly. It became apparent however that students were using knowledge of the test to guide what they did in the tasks, not just what the task required. Some lexical-grammatical features and the use of phrases were reported only in GW. 9A-S illuminated something of the thinking that could be promoted in GW when he reported that he discovered a lexical grammatical feature, using v-ing after -confident-, when explaining the meaning of the word, -confident-, to group members. In order to answer a group member's question, he re-read the question and checked the dictionary again; finally he made a sentence with the target word by using the lexical information he found. By contrast, 8B-S showed that thinking in individual work was less rich than in group work. He said: 'I will not know so much about the knowledge of words when I work on the task alone. What I do is to memorize as much vocabulary as I can without knowing the lexical grammatical use of the words clearly'.

Overall, students seemed to discover richer lexical information in GW than in IW. This seemed to be because, in GW, students had more stimulation to notice new aspects of words, and more opportunities to retrieve words and add to word knowledge from other group members during group discussions; they also had the chance to learn more pieces of word knowledge by giving help (e.g. explaining things) to their fellow group members.

4.4 The Aspects of Word Knowledge Reported as Retained in IW vs GW

Table 5. Aspects of word knowledge retained in IW vs GW (N=24)

Aspects of word knowledge remembered in IW	Example	Students
Synonym and antonym of target words	target words in general	10B-S
Meaning of target words	target words in general	10B-S,
Spelling of target words	target words in general	10A-S
Grammatical use of words	frustrate	10A-S, 10B-S
Aspects of word knowledge remembered in GW		Students
Synonym of target word	prevent/stop	8A-S
Meaning of target word	unique/extraordinary/ one and only	7B-S,
Compound noun	remote/remote control	8A-S
Phrase	pay off	8B-S
Lexical grammatical use of words	target words in general	9A-S

As can be seen from Table 5, the aspects of word knowledge students in both groups said they remembered, not surprisingly, spanned the same range of types as we saw above for discovery. Hence we will not go over each separately again here.

The factors which students reported to help them memorize synonyms, antonyms and meanings of target words in IW were target words repeatedly (10A-S), including having no interruption from group discussion, having to depend on oneself, and asking the teacher for help (10B-S). An example of teacher impact on retention in IW was remembering the grammatical use of words (10A-Ss, 10B-S), because of help from teacher correction rather than because they used a dictionary. The teacher's corrections seemed to make certain word knowledge more memorable. 10a-S for example said 'I make a sentence with the new word, frustrate: I am frustrating. Then, the teacher corrects the mistake in the sentence; the sentence becomes I am frustrated. In this situation, I remember easily about the grammatical use of the word frustrate'.

In GW, the factors that helped students to memorize many aspects of target words were claimed to be learning this information from group discussion (8A-S), and contributing word knowledge to group members (7B-S). One way in which this occurred was given by 8A-S who reported memorizing a synonym of the target word, because group members mentioned a synonym word that was easy to remember (e.g. synonym of prevent could be stop) during group discussion. Again, 8B-S reported that he understood and remembered the phrases discussed in group discussions, since group members used better examples to help him understand: 'I can remember better about the phrases or words that are discussed in the group discussion. For example, the phrases pay off in the sentence: Sandy took out many loans for medical school, so it will take them a long time to pay off his debts. Besides, group members often tell me better examples to help me understand about it'. The implication here is that he sees better understanding as leading to better memory.

Furthermore, the very act of receiving information in discussion was claimed to help. As 8A-S said: 'A group member finds out that remote can also be remote control. After discussion, all of group members understand that remote can be distant or remote control. In order to answer this question in the task, the former is better. So I understand and remember a new meaning of the word now'.

Not only receiving but also giving information in GW seemed to help. 9A-S pointed out: 'I specially remember about certain words because groupmates asked me about the meaning or lexical grammatical use of those words. So I explain the meaning or grammatical use of words to group members which help me to have deep impression about them'.

Overall, once again the processes of GW were reported as providing richer support for retention than IW, both in terms of multiple retrievals and generative use, in Nation's terms, and through their interactive nature. Nevertheless, as we have seen, this benefit was only weakly reflected in the delayed post-test score evidence ($t=1.85$, $p=.027$). (We used paired t tests again to compare posttest scores, but note that it was not comparing groups but the same people on two sets of vocabulary that learnt in GW and that learnt in IW).

5. Discussion and Conclusion

5.1 Vocabulary Learning Through GW and IW

The test results showed that students in the very short term attained higher scores through GW than through IW. GW fulfilled the expectations from the general literature (e.g. Donato, 1994; Anton and DiCamilla, 1988; Lasito & Storch, 2013; Dobao, 2014), allowing more knowledgeable others, other than the teacher, to be involved. This increased the opportunities for the desirable processes identified by Nation (2013) to occur, such as lexical information being noticed, retrieved and generatively used. It enhanced the learning of different aspects of word knowledge through error correction, explanation, suggestion, and sharing of resources, as compared with learning in IW. Moreover, it enhanced task involvement since doing any task with others in GW required more attention, and richer searching than just talking to oneself in IW (Laufer and Hulstijn, 2001). Furthermore, in GW, students not only had more opportunities to receive lexical information from group members but also to benefit through giving up (e.g. explanation) to group members during group discussion.

An unexpected finding was the increase in effectiveness of IW over the period. One expects a new treatment (GW) to show increasing pre-post effect over time as learners get used to it, but IW was familiar and would be expected to have a pre-post effect unchanging over the period of the study. However, its result in some way presents it as another 'new' treatment. This may be due to the fact that, unlike normally in such classes, there were regular vocabulary post-tests at the end of each class, administered for the researcher to gather data but actually also having an unanticipated pedagogical effect both on IW and GW learning (as claimed by Karpicke and Roediger, 2007). This is supported by the fact that the pre-post gain in the very first learning session of the study, which involved IW, was exceptionally low at only 11%, but increased considerably over ensuing sessions. This maybe because that was the very first session and students did not yet really anticipate the immediate post-test, even though they had been told there would be one.

The major fall off for both GW and IW at the delayed post-test was not entirely unexpected as it is common in vocabulary learning studies (Jones, 2006; Kim, 2008). In our case this might have been enhanced because the vocabulary tested in the delayed post-test was not to be included in the final course examination (taken by the students as part of course requirements, but not part of the research). Nevertheless, this result is disappointing particularly as in interview the GW students mentioned a range of benefits of GW for retention, yet remembered only a little more lexical information than IW students in the delayed post-test. Of course our tests only included the targeted vocabulary items from the textbook exercises. It is possible that if non-targeted vocabulary had been tested, GW would have shown a more marked long term benefit than IW, given the nature of the reported GW interactions. Nevertheless, this prompts the need for consideration of how to combat the rapidity of attrition of learnt vocabulary knowledge. Perhaps the message is that even the most effective method used at the initial vocabulary learning stage cannot overcome the widely recognized additional need for long term, recycling of vocabulary at regular intervals after the session where new information arises (Karpicke and Roediger, 2007; Nation, 2013).

5.2 Future Research

While the value of explicit vocabulary tasks following classroom reading is no longer really in question, there remain many facets of how such tasks are best implemented still to be investigated. We used collaborative GW to investigate vocabulary learning and retention in the present study. It would be interesting to test the effectiveness of other types of GW such as cooperative group work. Moreover, the independent part played by tested in combination with IW or GW need to be further investigated, including the kinds of informal tests which are commonly given by teachers or pedagogical vocabulary recycling purposes. Finally, longer periods of delayed retention should be measured, with and without intervening revision/recycling in some form a week or a month later before a much delayed post-test, so as to simulate real learning conditions more closely.

References

- Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal*, 78(4), 465-483. <https://doi.org/10.1111/j.1540-4781.1994.tb02064.x>
- Anton, M., & DiCamilla, F. J. (1998). Socio-cognitive functions of L1 collaborative interaction in the L2 classroom. *The Canadian Modern Language Review*, 54, 314-342. <https://doi.org/10.3138/cmlr.54.3.314>
- Baddeley, A. D. (1990). *Human memory*. London: Lawrence Erlbaum Associates.
- Baddeley, A. D., & Hitch, G. J. (2017). Is the levels of processing effect language-limited? *Journal of Memory and Language*, 1-13. <https://doi.org/10.1016/j.jml.2016.05.001>

- Beck, I., Mckeown, M., & Omanson, R. (1987). The effects and uses of diverse vocabulary instructional techniques. In M. McKeown & M. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 147-163). Hillsdale, NJ: Erlbaum. [https://doi.org/10.1016/0346-251X\(87\)90048-0](https://doi.org/10.1016/0346-251X(87)90048-0)
- Behydt, L. (1987). The semantization of vocabulary in foreign language learning. *System*, 15(1), 55-67. [https://doi.org/10.1016/0346-251X\(87\)90048-0](https://doi.org/10.1016/0346-251X(87)90048-0)
- Bogdan, R. C., & Biklen, S. K. (2007). *Qualitative research for education: An introduction to theory and methods* (5th ed.). Boston: Allyn & Bacon.
- Craik, F., & Lockhart, R. (1972). Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-84.
- Day, R. R., & Yamanaka, J. (2007). *1 Cover to Cover*. Oxford University Press.
- DiCamilla, F. J., & Anton, M. (1997). Repetition in the collaborative discourse of L2 learners: A Vygotsky perspective. *The Canadian Modern Language Review*, 53, 609-33. <https://doi.org/10.3138/cmlr.53.4.609>
- Dobao, A. F. (2014). Vocabulary learning in collaborative tasks: A comparison of pair and small group work. *Language Teaching Research*, 18(4), 497-520. <https://doi.org/10.1177/1362168813519730>
- Donato, R. (1994). Collective scaffolding in second language learning. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 33-56). Norwood, NJ: Ablex.
- Elley, W., & Mangubhai, F. (1983). The impact of reading on second language learning. *Reading Research Quarterly*, 19, 53-67. <https://doi.org/10.2307/747337>
- Ellis, R. (1991). The interaction hypothesis: A critical evaluation. In E. Sadtono (Ed.), *Language acquisition and the second/foreign language classroom* (pp. 179-211). Singapore: RELC Anthology Series, 28.
- Ferris, D. (1988). *Reading and second language vocabulary acquisition* (Unpublished paper). Department of Linguistics, University of Southern California.
- Grabe, W. (2009). *Reading in a second language: Moving from theory from theory to practice*. Cambridge: Cambridge University Press.
- Hirsh, D., & Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*, 8, 689-696.
- Horst, M., Cobb, T., & Meara, P. (1998). Beyond a clockwork Orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, 11, 207-223.
- Hulstijn, J. H., & Laufer, B. (2008). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, 51, 539-558. <https://doi.org/10.1111/0023-8333.00164>
- Jee, H. M., & Hyeok, J. C. (2016). An experimental study of dictionary use on vocabulary learning and reading comprehension in different task conditions. *International Journal of Lexicography*, 31(1), 29-52. <https://doi.org/10.1093/ijl/ecw037>
- Jones, L. C. (2006). Effects of collaboration and multimedia annotations on vocabulary learning and listening comprehension. *CALICO Journal*, 24(1), 33-58. <https://doi.org/10.1558/cj.v24i1.33-58>
- Karpicke, J. D., & Roediger, H. L. (2007). Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 33(4), 704-719. <https://doi.org/10.1037/0278-7393.33.4.704>
- Kim, Y. (2008). The contribution of collaborative and individual tasks to the acquisition of L2 vocabulary. *The Modern Language Journal*, 92(1), 114-130. <https://doi.org/10.1111/j.1540-4781.2008.00690.x>
- Kim, Y. (2011). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. <https://doi.org/10.1111/j.1467-9922.2011.00644.x>
- 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. <https://doi.org/10.1111/j.1540-4781.1994.tb02043.x>
- Krashen, S. D. (2011). *Free vocabulary reading*. US: Libraries Unlimited.
- Lasite & Storch, N. (2013). Comparing pair and small group interactions on oral tasks. *RELC Journal*, 44(3), 361-375. <https://doi.org/10.1177/0033688213500557>
- Laufer, B. (1992). The lexical plight in second language reading: Words you don't know, words you think you

- know, and words you can't guess. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 20-34). Cambridge: Cambridge University Press.
- Laufer, B. (2000). Electronic dictionaries and incidental vocabulary acquisition: Does technology make a difference? In U. Heid, S. Evert, E. Lehmann, & C. Rohrer (Eds.), *EURALEX* (pp. 849-854). Stuttgart: Stuttgart University.
- Laufer, B. (2003). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. *The Canadian Modern Language Review*, 59(4), 567-587. <https://doi.org/10.3138/cmlr.59.4.567>
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The effect of task-induced involvement load. *Applied Linguistics*, 22(1), 1-26. <https://doi.org/10.1093/applin/22.1.1>
- Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Lanugage*, 22(1), 15-30. Retrieved from <http://files.eric.ed.gov/fulltext/EJ887873.pdf>
- Leeser, M. J. (2004). Learner proficiency and focus on form during collaborative dialogue. *Language Teaching Research*, 8, 55-81. <https://doi.org/10.1191/1362168804lr134oa>
- Leow, R. P. & Mercer, J. D. (2015). Depth of processing in L2 learning: theory, research, and pedagogy. *Journal of Spanish Language Teaching*, 2(1), 69-82. <https://doi.org/10.1080/23247797.2015.1026644>
- Mackey, A., & Gass, S. (2011). *Research methods in SLA: A practical guide*. Chichester: Wiley-Blackwell. [https://doi.org/10.1016/S0959-4752\(96\)00021-7](https://doi.org/10.1016/S0959-4752(96)00021-7)
- Maki, N., & Thomas, C. V. F. (2015). Vocabulary acquisition through direct and indirect learning strategies. *English Language Teaching*, 8(10), 142-151. Retrieved from <https://files.eric.ed.gov/fulltext/EJ1079088.pdf>
- Mercer, N. (1995). *The guided construction of knowledge: Talk among teachers and learners*. Clevedon: Multilingual Matters. <https://doi.org/10.1111/j.1467-9922.2007.00435.x>
- Mercer, N. (1996). The quality of talk in children's collaborative activity in the classroom. *Learning and Instruction*, 6(4), 359-377. [https://doi.org/10.1016/S0959-4752\(96\)00021-7](https://doi.org/10.1016/S0959-4752(96)00021-7)
- Min, H. T. (2008). EFL vocabulary acquisition and retention: reading plus vocabulary enhancement activities and narrow reading. *Language Learning*, 58(1), 73-115.
- Myong, H. K. (2012). Glossing and second language vocabulary learning. *TESOL Quarterly*, 46(1), 56-79. <https://doi.org/10.1002/tesq.3>
- Nation, I. S. P. (2013). *Learning vocabulary in another language*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139524759>
- Ohta, A. S. (2000). Rethinking interaction in SLA: Developmentally appropriate assistance in the zone of proximal development and the acquisition of L2 grammar. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 51-78). Oxford University Press.
- Ohta, A. S. (2001). Peer interactive tasks and assisted performance in classroom language learning. In A. S. Ohta, *Second language acquisition processes in the classroom: Learning Japanese* (pp. 73-128). Mahwah, NJ: Lawrence Erlbaum.
- Panitz, T. (1999). Collaborative versus cooperative learning: A comparison of the two concepts which will help us understand the underlying nature of interactive learning. Retrieved from <http://home.capecod.net/~tpanitz/tedsarticles/coopdefinition.htm>
- Paribakht, T. S., & Wesche, M. (1993). The relationship between reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada Journal*, 11(1), 9-29. <https://doi.org/10.18806/tesl.v11i1.623>
- Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In T. Huckin, & J. Coady (Eds.), *Second Language Reading and Vocabulary Learning* (pp.174-200). Norwood, N. J.: Ablex Publishing Corporation.
- Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3rd ed.). California: SAGE Publications, Inc.
- Pavlenko, A. (2009). Conceptual representation in the bilingual lexicon and second language vocabulary learning. In A. Pavlenko (Ed.), *The bilingual mental lexicon: Interdisciplinary approaches* (pp. 125-160). Bristol: Multilingual Matters. <https://doi.org/10.21832/9781847691262-008>

- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*, 357-83. <https://doi.org/10.1080/10888430701530730>
- Pitts, M., White, H., & Krashen, S. (1989). Acquiring second language vocabulary through reading: A replication of the Clockwork Orange study using second language acquires. *Reading in a Foreign Language, 5*(2), 271-275.
- Rosszell, H. R. (2007). *Extensive reading and intensive vocabulary study in a Japanese university* (Unpublished doctoral thesis). Temple University Japan.
- Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129-158. <https://doi.org/10.1093/applin/11.2.129>
- Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition, and pedagogy* (pp. 199-227). Cambridge: Cambridge University Press.
- Schmitt, N., & McCarthy, M. (1997). *Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press.
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal, 95*(1), 26-43. <https://doi.org/10.1111/j.1540-4781.2011.01146.x>
- Scholfield, P. (1997). Vocabulary reference works in foreign language learning. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: description, acquisition and pedagogy* (pp. 279-302). Cambridge: Cambridge University Press.
- Storch. (2005). Collaborative writing: Product, process, and students' reflection. *Journal of Second Language Writing, 14*, 153-73. <https://doi.org/10.1016/j.jslw.2005.05.002>
- Swain, M. (2002). Peer-peer dialogue as a means of second language learning. *Annual Review of Applied Linguistics, 22*, 171-185. <https://doi.org/10.1017/S0267190502000090>
- Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal, 82*, 320-337. <https://doi.org/10.1111/j.1540-4781.1998.tb01209.x>
- Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Vygotsky, L. (1986). *Thought and language*. Cambridge, MA: MIT Press.
- Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language, 15*, 130-163.
- Watanabe, Y. (1997). Input, intake, and retention: Effects of increased processing on incidental learning of foreign language vocabulary. *Studies in second Language acquisition, 19*, 287-308. <https://doi.org/10.1017/S027226319700301X>
- Watanabe, Y. (2008). Peer-peer interaction between L2 learners of different proficiency levels: Their interactions and reflections. *The Canadian Modern Language Review, 64*(4), 605-635. <https://doi.org/10.3138/cmlr.64.4.605>
- Wegerif, R., & Mercer, N. (2000). Language for thinking: A study of children solving reasoning test problems together. In H. Cowie & G. van der Aslsvort (Eds.), *Social Interaction in Learning and Instruction* (pp. 179-192). Oxford: Pergamon.
- Williams, J. (1999). Learner-generated attention to form. *Language Learning, 51*, 303-346. <https://doi.org/10.1111/j.1467-1770.2001.tb00020.x>
- Williams, J. (2001). The effectiveness of spontaneous attention to form. *System, 29*, 325-340. [https://doi.org/10.1016/S0346-251X\(01\)00022-7](https://doi.org/10.1016/S0346-251X(01)00022-7)

Appendix A

Data Collection Overview

Week 1	a. VKS pre-test of target vocabulary in text 1-4 (20 mins)
Week 2	a. Individual work b. VKS immediate post-test (15 mins)
Week 3	a. Collaborative group work training + Collaborative group work b. VKS immediate post-test
Week 4	a. Individual work b. VKS immediate post-test
Week 5	a. Collaborative group work b. VKS immediate post-test
Week 6	a. Individual work b. VKS immediate post-test c. VKS pre-test for texts 5-12 (45 mins)
Week 7	a. Collaborative group work b. VKS immediate post-test
Week 8	a. Individual work b. VKS immediate post-test
Week 9	Mid-term exam (Not part of research)
Week 10	a. Collaborative group work + Individual work b. VKS immediate post-test
Week 11	a. Collaborative group work b. VKS immediate post-test (both GW & IW)
Week 12	a. Individual work b. VKS immediate post-test
Week 13	a. Collaborative group work b. VKS immediate post-test
Week 14	a. Individual work b. VKS immediate post-test c. Interviews
Week 15	a. Interviews
Week 16	a. Interviews
Week 17	a. Delayed VKS post-test (80 minutes). b. Interviews
Week 18	a. Final exam (Not part of research)

The group work in week 3 was considered as a training and practice session, in case some revision of the GW method might have been required at this point. In fact, however, the GW did not need to be modified since it integrated into the course well without causing problems. Due to the exigencies of the course timetable, and in order to conduct the delayed post-test (week 17) at least two weeks after the last classroom intervention (week 14), two reading texts had to be covered in week 10.

Appendix B

Examples of the Post-Reading Vocabulary Tasks

Vocabulary exercise for text 1

Looking at vocabulary in context

A. Find the words in bold in the text. For each line, circle the word that does not belong.

- attitudes** (par. 1) viewpoints outlooks heights
- increase** (par. 2) growth decrease rise
- continue** (par. 3) go on with keep on stop
- traditional** (par. 5) old-fashioned conservative modern
- responsibilities** (par. 5) answers duties tasks
- quit** (par. 5) keep on stop leave

B. Fill in the blanks with the words in bold from A.

- The soccer coach was so mean that three players _____.
- Airfares are getting more expensive because of the _____ in fuel prices.

3. Lisa has a very bad _____ at work. She's so lazy.
4. I am always busy in my new job, because I have so many different _____.
5. I hate using microwave ovens. I prefer cooking on the stove the _____ way.
6. You must _____ your diet if you want to lose more weight.

In part A, the form, meaning, synonyms, antonyms, and occasionally spelling of the target words were practised. In part B, the meaning of the target words and part of speech were practised through the exercise. Part B also showed examples of collocation that were different from the original text.

Vocabulary exercise for text 2

Looking at vocabulary in context

A. Find the words in bold in the text. Circle the correct definitions.

1. **Experts** (par. 2) are people with special knowledge or skills / who are not paid for their work.
2. If you sleep for an **average** (par. 3) of six hours a night, you sleep for exactly six hours every night / more on some nights and less on other nights.
3. **Deprivation** (par. 4) means having enough of something / not having enough of something.
4. **Symptoms** (par. 4) of an illness are the signs / causes of the illness.
5. A **debt** (par. 5) is something you must pay back / a special advantage.
6. **Disorders** (par. 6) are illness / accidents.

B. Fill in the blanks with the words in bold from A. Be sure to use the correct forms.

1. Sandy took out many loans for medical school, so it will take him a long time to pay off his _____.
2. I want the advice of an _____, so I hired the best lawyer in town.
3. The _____ temperature in San Francisco during the summer is 17 degrees.
4. That medical practice takes care of people with eating _____.
5. Food _____ is a big problem in many African countries.
6. Common _____ of flu include fever, sore throat and pain.

The first vocabulary exercise for text 2 practised the meaning of target words in relation to synonyms and antonyms in paraphrases. The second exercise for text 2 was similar to the one for text.

Appendix C

Words Targeted in the L2 Sessions

Targeted words for texts 1-4

Vocabulary:

1. a. attitudes b. increase c. continue d. traditional e. responsibilities f. quit
2. a. experts b. average c. deprivation d. symptoms e. debt f. disorders
3. a. psychic b. clients c. creative d. predict e. skeptical f. unique
4. a. concerns b. trend c. banned d. grounding e. sneak out f. addicted

Targeted words for texts 5-12

Vocabulary:

5. a. casually b. interrupt c. frustrated d. impressed e. interactive f. influenced
6. a. disprove b. population c. developed countries d. genes e. risky f. lifestyle
7. a. drifting b. mosque c. struck d. devastated e. drowned f. destroyed
8. a. images b. cosmetics c. particular d. plump e. depressed f. altered
9. a. outgoing b. rebellious c. reversed d. environment e. inherited f. dependable
10. a. remote b. efficient c. break down d. eagerly e. material f. unload
11. a. residence b. disable c. embarrassed d. pretended e. complained f. valuable
12. a. dramatic b. intercept c. evidence d. witness e. rational f. hoaxes

Appendix D

The VKS Instrument

Modified VKS used in the study

Self-report categories	
I	I don't remember having seen this word before.
II	I have seen this word before, but I don't know what it means.
III	I have seen this word before, and I <i>think</i> it means _____ (synonym or translation).
IV	I <i>know</i> this word. It means _____ (synonym, translation, and antonym).
V	I can use this word in a sentence: _____. (Write a sentence that you have not seen before or read in the text or exercise.) (If you do this section, please also do Section IV.)

VKS scoring categories--- meaning of scores (modified)

Self-report categories	Possible score	Meaning of scores
I	1	The word is not familiar at all.
II	2	The word is familiar but its meaning is not known.
III	3	A correct synonym, antonym or translation is given.
IV	4	The word is used with semantic (including good collocation) appropriateness in a sentence.
V	5	The word is used with semantic appropriateness and grammatical accuracy (including correct part of speech and inflection) in a sentence.

Appendix E

Interview Questions

1. Does group work help you discover more information about word than the word information you discover while working alone? If so, what vocabulary information so you discover through group work? How?
2. Does group work help you remember vocabulary knowledge? How? Why? Why not?
3. Does individual work help you discover vocabulary information? If so, what vocabulary information do you discover through individual work? How?
4. Does individual work help you remember vocabulary knowledge better?

Copyrights

Copyright for this article is retained by the author, with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).