

Potential Effects of Collaborative Output Task on L2 Novice Learners' Vocabulary Acquisition: A Pilot Study

Mutsumi Kawakami

Abstract—This paper reports on the potential impact of a collaborative output activity on the vocabulary acquisition of novice English learners in Japan. Twenty-six participants randomly assigned to pairs or groups of three performed a dictogloss task for three weeks. The dictogloss is an integrated skills activity that promotes oral interaction and written output through discussion of the text they heard. The current study aimed to examine the potential benefit of the activity on their vocabulary knowledge and their learning gain perceptions by modifying the dictogloss activity for novice learners. The results indicated that the participants significantly increased their meaning and usage aspects of vocabulary knowledge after the treatment but the retention of usage was superior to knowledge of meaning. Furthermore, the participants perceived the dictogloss activity as being effective on their productive skills for language learning. Some suggestions are proposed to enhance novice learners' attention to form and meaning.

Index Terms—Second language acquisition, focus on Form, interaction, collaborative learning.

I. INTRODUCTION

The Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) has put pressure on English teachers to seek better ways to promote English communication skills in their students. Productive vocabulary knowledge is indispensable for both oral and written communication. However, it is difficult to picture students being engaged in real communication in many English classes in Japan owing to lack of opportunities to use their English knowledge in meaningful form-focused input and output. Because of limited class time allocated to vocabulary instruction in many English classrooms in Japan, students take control of vocabulary learning. They traditionally memorize word meanings by using lists, written and oral drills, and flashcards with little opportunity to use the words in authentic contexts. Beginners specifically tend to memorize word meanings and spellings by looking at a wordlist without paying attention to other aspects of word knowledge such as syntactic features and pronunciation of words, while they feel a difficulty in using the learned words in writing and speaking [1].

A. Productive Vocabulary Knowledge and Learning

In vocabulary learning, both incidental and deliberate learning are considered important. There is some evidence that incidental learning is better for learning more about

words whose meanings have already been learned ([2], [3]), other prior research suggests that deliberate learning such as by using word cards may be an efficient way of learning some aspects of word usage [4]. Nation ([5], p.2) claims four prominent strands of vocabulary learning: meaning-focused input, meaning-focused output, language-focused learning, and fluency development. English teachers in universities should create opportunities for their students to learn through these four strands, especially meaning-focused input and output with deliberate attention to language form in classrooms, which learners have little experience of before they enter university.

Laufer [6] compared two activities: Focus on Form: reading and dictionary use followed by a discussion and comprehension questions, and Focus on FormS: memorizing words in a list and sentence completions. She also examined the effectiveness of subsequent intentional learning with a word list. Her findings showed that the Focus on FormS group gained target words on both passive and active word recall tests significantly greater than the Focus on Form group. However, the subsequent intentional learning reduced the difference between the two groups. The tests employed in this study were recall tests, which asked the learners to choose L1 or L2 translation from options (passive knowledge) and to write counterparts of L1 or L2 words (active knowledge). The results of her study imply that learning words in context should be followed by form focused learning.

In his study examining the effect on vocabulary acquisition of two task types, writing and reading, Webb [7] argues that time spent on task is one of the most important factors in acquiring vocabulary. Productive tasks (writing) may be more effective on the gains of all aspects of vocabulary acquisition over receptive tasks (reading) as learners tend to spend more time on productive tasks.

B. Interaction during Tasks and Language Learning

Research over the past two decades has demonstrated that collaborative learning during tasks can facilitate second language vocabulary acquisition, contingent upon learner interaction (e.g., [8], [9]). These studies are based on sociocultural theory whose starting point is language as a trigger of cognitive process involved in second language learning not just as a tool of communication. These studies use output-based tasks that enable learners to focus on form to complete the task, such as text reconstruction tasks in groups. As Swain and Lapkin [9] mention in their Output Hypothesis, learners become aware of problems in language they produce resulting in noticing a gap between their interlanguage and target language forms and thus test their hypotheses about the target language. Language related

Manuscript received December 24, 2017; revised February 24, 2018.

M. Kawakami is with Science and Engineering Department, Tokyo Denki University, Saitama, Japan (e-mail: mutsumik@mail.dendai.ac.jp).

episodes (LREs) are defined as “any part of a dialogue where students talk about the language they are producing, question their language use, or correct themselves or others” [10]. The efficacy of interaction on language acquisition depends on the quality and quantity of interaction in a pair or a group.

Storch [11] compared qualities of LREs performed among four tasks: multiple choice, rational definition (cloze), text reconstruction and composition. The results showed that the text reconstruction task most enhanced attention to grammatical choices; the cloze task promoted the least. She reported that more structured tasks restricted learners’ focus on grammar. Hence, text reconstruction tasks performed by dyads seem to have the most potential to encourage them to focus on form leading to L2 acquisition.

Kim [8] examined the impact of LREs on vocabulary acquisition of Korean as a second language learners through dictogloss, comparing dyads and individual LREs. Her findings showed that there were no significant differences in the number of LREs between the two conditions, but assistance given from peers may enhance vocabulary learning. Additionally, results of the delayed posttest administered two weeks after the treatment indicated that learners gave correct answers when they successfully resolved LREs.

Williams [12] investigated the relationship between LREs and L2 development for eight learners of different proficiencies during a text reconstruction task in an eight-week classroom treatment. The results of her study illustrated that the learners correctly answered the lexical and grammatical items on the tailor-made test when LREs were successfully resolved. The results also showed that the higher proficiency students were more likely to benefit from information provided by their peers than the lower proficiency students.

From a psychological approach to task-based L2 teaching, dictogloss [13] is a teaching technique that creates opportunities for learners to learn words in context through interaction. Learners are asked to listen to a passage and work together in small groups to reconstruct the original passage during a dictogloss task. This method helps induce the cognitive process of noticing specific language features through input and pushed output through which learners hypothesize how the target language works and test such hypotheses [14]. Swain and Lapkin [10] claim that productive tasks are important and effective for L2 acquisition when learners work together and focus on form, meaning, and function of language through feedback from their peers.

Most of these studies are conducted in ESL contexts and the participants were above the intermediate proficiency level. There are few studies that examine the efficacy of interactional tasks on the productive vocabulary retention for low proficiency students in EFL contexts. The current study aimed to explore the effects of a dictogloss task on novice English learners at a university in Japan. Data regarding their evaluation and application of the task was also collected. The research questions addressed were:

1. Does a dictogloss task modified for novice learners have an influence on their vocabulary learning?
2. How do the learners perceive the task involving listening, writing and discussion with their peers?

II. METHODOLOGY

A. Participants

26 second year students at a science and engineering university performed a dictogloss in a reading and writing class for 3 weeks. However, removing students who were absent from a number of the tasks, the final number of participants was 18. They performed the task for 30 minutes in a 90-minute class session every week. A textbook used in this course focuses on reading skills and vocabulary exercises. Their English proficiency was examined based on scores of a TOEIC Bridge practice test (Educational Testing Service [15]) administered at the beginning of the semester ($M=33.6$, $SD=4.74$), indicating that they were at CFER A2 proficiency level.

B. Materials

1) Task

In the current study, a dictogloss task was implemented based on a textbook, “Select Readings” (Oxford University Press [16]). In order to identify the appropriate length for the dictogloss text, a practice session was conducted one week prior to the treatment. Summaries of three texts of different length were created based on students’ comments on the practice dictogloss and teacher’s impression of students’ performance (45 words, 77 words, 81 words).

2) Target words

Target words were selected from the textbook based on the words introduced as new words each chapter. The words used are as follows:

amount, claim, concerned, depressed, describe, develop, disagree, encounter, happen, harmful, intense, participants, potentially, profession, questionnaire, radiation, range, regardless of, suffer, survey, tingly

These words are on the second or third level on the new general frequency list based on the vocabulary profiler¹. About 41% of the target words were known by the participants.

3) Mini dictionary

In order to reduce task complexity and provide equal opportunity for access to meanings of the target words, a mini dictionary was created for the current study based on Folse [17]. They were allowed to refer to the mini dictionary during the discussion of the dictogloss task. It contained target vocabulary with meaning and parts of speech in Japanese and two sample sentences.

C. Measurement Instruments

1) Pretest and immediate posttest and delayed posttest

To measure participants’ vocabulary knowledge, a modified version of the vocabulary knowledge scale (VKS), developed by Paribakht and Wesche [18], was adapted for the pretest, the immediate posttest, and the delayed posttest. The VKS was modified based on Folse [17]. Although the original version of VKS measured learners’ familiarity of vocabulary meaning with a five-point scale, this test consists

¹ <http://www.lexutor.ca/vp/>

of questions on three different word knowledge aspects as shown in Table I.

Scoring of the tests was conducted on two different perspectives of word recall knowledge: meaning and usage. Regarding meaning of the words, if Category I was reported, no point was given. For the self-reported Category II, two points were given if the correct meaning and part of speech of each item was demonstrated. If the correct meaning of a word was demonstrated with the wrong part of speech, one point was awarded. For the self-reported Category III, regarding usage of the words, two points were awarded if the correct meaning and part of speech of each item was demonstrated in a sentence. One point was awarded for syntactically incorrect use of the word in a sentence. Thus, the maximum score of each word was two points on both meaning and usage recall knowledge.

In addition, after the delayed posttest, the participants received an open-ended questionnaire in which they were asked how effective they considered the dictogloss task or how they perceived the task. They also replied to questions on the difficulty of the task, how they used the mini dictionary, and group work in general.

TABLE I: SELF-REPORTED CATEGORIES OF VOCABULARY KNOWLEDGE SCALE (MODIFIED VERSION)

I	I don't know what this word means.	(0 points)
II	I know the meaning of this word.	(2 or 1 point)
III	I can use this word in a good example sentence. Write the sentence that includes the word.	(2 or 1 point)
e.g.:		

D. Procedure

Before the treatment, the participants practiced performing a dictogloss in their regular class time to get used to the procedure. The treatment was conducted over a 3-week period. On the first day, the participants were told that they were taking part in research exploring how dictogloss affects their vocabulary learning. For the pretest, the participants were asked to complete a VKS. In the first stage of the lesson, the participants wrote the meanings of ten target English words before the target vocabulary was instructed. After that, they read a text and answered comprehension questions.

In the second stage, the participants performed the dictogloss task for 30 minutes. In the first step of the original dictogloss (the dictation step), learners listened to the text twice at normal speed, taking notes regarding words and phrases they heard. In contrast to the original step, in the current study, the participants listened to a summary of the text twice at normal speed but there was ten second break between each sentence. For the first listening, they were encouraged to listen to get the gist of the text. For the second listening, they performed a transfer information task in which they were asked to complete a table (Appendix B). Nation [19] points out that the information transfer task may allow learners to attend to word form and on the detail of the information used in listening, and thus would help their comprehension. After, the teacher promoted participants to check the words and phrases they heard and write them on a blackboard provided. Then, the participants worked in pairs or groups of three to reconstruct the passage.

In the second step (text reconstruction and analysis step) the participants had fifteen minutes to discuss with each other what they heard based on their notes. The participants were allowed to refer to the mini dictionary or their own dictionary during the discussion. They listened to the text for the third time and added any missing information to their sentences. They had ten minutes after the final listening to reconstruct the sentences from their shared information.

Upon task completion, they were asked to correct their passage by referring to the original passage.

In the third stage, the participants were asked to report their familiarity with the target words on the VKS as an immediate posttest.

Three weeks after the last treatment session, the participants were asked to report their familiarity of all twenty-one target words on the VKS as a delayed posttest.

E. Analysis of Data

Identical data of the VKS were scored by two experienced EFL instructors. The Pearson's r for inter-rater reliability was .98 for the pretest, posttest, and delayed posttest.

To test for statistically significant difference among three tests, the pretest, the immediate posttest, and the delayed posttest within group, a one-way repeated measures ANOVA was conducted as the number of this group was relatively small but the scores were normally distributed. The effect size of the pairwise comparisons were calculated using partial η^2 . Effect size values were interpreted as small = .01, medium = .06, or large = .14, as suggested by Takeuchi and Mizumoto [20].

Students' comments on the questionnaire were analyzed using percentages of their attitudes toward the task.

III. RESULTS AND DISCUSSION

In order to answer the first research question about the influence of a dictogloss task on vocabulary acquisition, the descriptive statistics of the VKS scores for all 21 items were analyzed. As shown in Table II, as expected, participants' knowledge of word meanings was higher than usage in the pretest. The scores for both meaning and usage increased in the immediate posttest but the difference between the two aspects of word knowledge was less pronounced. Twice as much as word meaning was gained and three times as much as usage was gained. Finally, scores for both meaning and usage knowledge reduced on the delayed posttest but still 60% of word meaning and usage was retained.

TABLE II: DESCRIPTIVE STATISTICS OF THE VKS SCORES

	Pretest (%)	Immediate posttest (%)	Delayed posttest (%)
Meaning	17.1 (41)	38.7 (91)	26.6 (63)
Usage	10.0 (24)	28.0 (67)	23.8 (56)
Total	27.1 (33)	66.7 (79)	50.4 (60)

Note. The total possible score for each aspect of word knowledge was between 0 to 42.

Data of mean scores on the three tests was submitted to a one-way repeated measures ANOVA to analyze the task effects on the vocabulary gains. The results reported in Table III and IV indicate a statistically significant difference among the results of the three tests with a large effect size (Meaning: $F(2, 34) = 90.34, p = .00$, partial $\eta^2 = .84$; Usage: $F(2, 34) =$

50.68, $p = .00$, partial $\eta^2 = .75$), therefore, a post-hoc analysis with a Bonferroni correction test was also applied to see the effects of time. The results revealed that knowledge of word usage was significantly more retained than the meaning of words. Regarding mean scores for meaning of the target words, there was a significant increase between the pretest and immediate posttest ($p < .001$) and then a significant decrease from the immediate posttest to the delayed posttest ($p < .001$). On the other hand, the mean scores for usage of the target words significantly increased from the pretest to the immediate posttest ($p < .001$) but there was no statistical significance in the difference between the immediate posttest to the delayed posttest ($p = .25$).

TABLE III: ANALYSIS OF VARIANCE FOR THE MEANING SCORES ON THE THREE TESTS

	df	Type III SS	Mean square	F value	Partial η^2
Time	2	4223.81	2111.91	90.34***	0.84
Error	34	794.85	23.38		

Note. *** $p < .0001$

TABLE IV: ANALYSIS OF VARIANCE FOR THE USAGE SCORES ON THE THREE TESTS

	df	Type III SS	Mean square	F value	Partial η^2
Time	2	3189.926	1594.96	50.68***	0.75
Error	34	1070.07	31.47		

Note. *** $p < .0001$

Laufer & Girsai [21] report that the most difficult aspect of word knowledge is form-meaning knowledge such as recalling word form and recalling word meaning. Additionally, word usage requires the learners to use not only knowledge of a word meaning but also knowledge of syntactic features of the word. Schmitt [22] suggests that incremental acquisition of words might proceed from learning meaning recall to building up other aspects of vocabulary knowledge although it has not been well understood how vocabulary is acquired yet. Therefore, it may take time to learn morphological and grammatical features of word knowledge through contextualized learning experience. These features would have been enhanced by the discussion during the dictogloss task. The learners examined the target words focusing on meaning and form of the words in context. Such cross-linguistic focus on form may have resulted in acquiring usage of the words, through *pushed output* and *noticing* as supported in earlier studies.

In summary, findings for the first research question indicated that the learners significantly improved scores for both meaning and usage of words immediately after the treatment. Mean scores for meaning of words were slightly higher than those for usage of words three weeks after the treatment but mean scores for usage of words retained more.

Table V includes the questionnaire. All participants were allowed to write their perception of the task freely. Their responses are classified into eight categories. With respect to the second research question, the results of the questionnaire revealed that 20% of the participants perceived the task as difficult, while 15% of them perceived it as enjoyable or found importance in the usage of language. However, the results of a separate questionnaire regarding the difficulty of the task indicated that approximately 70% of the participants perceived the task level to be of an appropriate difficulty level. Most of them used the mini dictionary in confirming the meaning of words during the discussion and their group

work was supportive with good feedback from peers.

TABLE V: THE PARTICIPANTS' PERCEPTION OF THE TASK

I had opportunities to practice listening skills.	13%
I had opportunities to practice spelling.	8%
I learned vocabulary.	18%
I enjoyed the task or learned important skills for conversation.	15%
I understood the contents of the text better.	15%
I learned chunks and expressions of English.	3%
I had opportunities to confirm my knowledge that wasn't for sure.	10%
It was difficult to perform the task.	20%
Total	100%

Furthermore, the participants seemed to perceive three factors for the effectiveness of the task. Eighteen percent of the students found it effective for practicing vocabulary, 8% of them mentioned spelling, and 13% of them mentioned listening practice. Some students referred to comprehension. Ten percent of the participants mentioned that this task helped them understand the text and 10% of them perceived the task efficient for hypothesis testing. In sum, the participants seemed to perceive the dictogloss activity as effective in promoting skills related to the productive knowledge of English. Interestingly, no one mentioned grammatical aspects of language learning. Nabei [23] points out that students perceive teacher talk as one of the most useful learning sources. The participants in the current study might have attended more to vocabulary since the teacher focused on their vocabulary knowledge during class discussions.

IV. CONCLUSION

The current study showed that the dictogloss task enhanced vocabulary learning. Results indicated that participants' vocabulary knowledge increased twice as much as the pretest. In particular, gains in knowledge regarding word meaning were greater than those for usage after the treatment. However, usage of the words was retained more than the meaning of words in a delayed posttest. The participants' word knowledge might have developed through a dictogloss task by using the target words in context. As a result, 60 % of the target vocabulary was learned productively on the delayed posttest. This study also showed that the participants tended to perceive the efficacy of the dictogloss task on their productive knowledge of English.

There are some limitations with the current study. Further practice of the task is needed as the novice learners of English tended to focus on the sound of each word, not the meaning. Consequently, they felt difficulty in listening. Moreover, there is a need for further studies to examine the efficacy of the task in which the learners are encouraged to have interaction with their peers with a large number of participants and a control group.

Overall, findings of the current study have pedagogical implications for the use of integrative tasks that promote oral interaction and written output in EFL classes, suggesting that even novice learners can benefit from collaborative work if they have extra support during the task performance.

APPENDIX A DICTOGLOSS TEXT

This dictogloss text is about mobile phone use.

These days, mobile phones are commonly used all around the world. The explosion in mobile phone use has some health professionals concerned that people may suffer health problems such as serious memory loss. What makes mobile phones potentially harmful is radiation. While mobile phone companies claim that the amount of radiation from mobile phones is too small to worry about, some scientists disagree. Whether mobile phone use is harmful or not, we should be careful using them.

APPENDIX B SAMPLE TRANSFER INFORMATION TASK

Direction: You are going to listen to a summary of the text. Complete the table below with the information from listening.

How many people answered the questionnaire?	
What did the questionnaire ask?	
What are the results of the questionnaire?	

REFERENCES

- [1] M. Kawakami, "The effect of text-based tasks on English vocabulary retention in less proficient Japanese learners of English: A case study," Ph.D. dissertation, Dept. Society of Humanities, Kanto Gakuin Univ., Japan, 2013.
- [2] P. Nation and K. Wang, "Graded readers and vocabulary," *Reading in a Foreign Language*, vol. 2, no. 2, pp. 355-380, 1999.
- [3] R. Warning and M. Takaki, "At what rate do learners learn and retain new vocabulary from reading a graded reader?" *Reading in a Foreign Language*, vol. 15, no. 2, pp. 130-163, 2003.
- [4] I. Elgort, "Deliberate learning and vocabulary acquisition in a second language," *Language Learning*, vol. 61, no. 2, pp. 367-413, 2011.
- [5] P. Nation, *Learning Vocabulary in Another Language*, Cambridge, U.K.: Cambridge University Press, 2001.
- [6] B. Laufer, "Comparing focus on form and focus on forms," *The Canadian Modern Language Review*, vol. 63, pp. 149-166, 2006.
- [7] S. Webb, "Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge," *Study of Second Language Acquisition*, vol. 27, pp. 33-52, 2005.
- [8] Y. Kim, "The contribution of collaborative and individual tasks to the acquisition of L2 vocabulary," *The Modern Language Journal*, vol. 92, pp. 114-130, 2008.
- [9] M. Swain and S. Lapkin, "Focus on form through collaborative dialogue: Exploring task effects," in *Researching Pedagogic Tasks, Second Language Learning, Teaching and Testing*, M. Bygate, P. Skehan, and M. Swain, Eds., pp. 99-118, Harlow: Longman, 2001.
- [10] M. Swain and S. Lapkin, "Focus on form through conscious reflection," in *Focus on Form in Classroom SLA*, Daugherty and William, Eds., Cambridge, U. K.: Cambridge University Press, 1998, ch. 3, pp. 64-81.
- [11] N. Storch, "Comparing second language learners' attention to form across tasks," *Language Awareness*, vol. 7, no. 4, pp. 176-191, 1998.
- [12] J. Williams, "The effectiveness of spontaneous attention to form," *System*, vol. 29, pp. 325-340, 2001.
- [13] R. Wajnryb, *Grammar Dictation*, Oxford, U.K.: Oxford University Press, 1990.
- [14] M. Swain and S. Lapkin, "Problems in output and the cognitive process they generate: A step towards second language learning," *Applied Linguistics*, vol. 16, pp. 371-391, 1995.
- [15] Educational Testing Service, *TOEIC Bridge Official Guide & Question Collection*, Tokyo, Japan: Institute for International Business Communication, 2007.
- [16] L. Lee and E. Gundersen, *Select Readings*, Oxford, U.K.: Oxford University Press, 2011.
- [17] K. Folse, "The effect of type of written exercise on L2 vocabulary retention," *TESOL Quarterly*, vol. 40, no. 2, pp. 273-293, 2006.
- [18] T. Paribakht and M. Wesche, "Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition," in *Second Language Vocabulary Acquisition*, J. Coady & T. Huckin, Eds., Cambridge, U.K.: Cambridge University Press, 1993, pp. 174-200.
- [19] P. Nation, *Teaching ESL/EFL Listening and Speaking*, N.Y.: Routledge, 2009, pp. 47-49.
- [20] O. Takeuchi and A. Mizumoto, *The Handbook of Research in Foreign Language Learning and Teaching*, Tokyo: Shohakusha, 2012.
- [21] Laufer and Girsai, "Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation," *Applied Linguistics*, vol. 29, no. 4, pp. 694-716, 2008.
- [22] N. Schmitt, *Researching Vocabulary*, U.K.: Macmillan, 2010.
- [23] T. Nabei, " 'Collage students' reports on an English as a foreign language class: Output hypothesis, output activities, and noticing," *Kansai University Faculty of Foreign Language Studies Kiyo*, vol. 7, pp. 45-59, 2012.



Mutsumi Kawakami serves as a lecturer at Tokyo Denki University. She completed her M.A. in TESOL at San Francisco State University and her PhD in English literature at Kanto Gakuin University. Her research interests include second language acquisition, strategy instruction, L2 vocabulary instruction, and task-based language teaching.