Can Contextualization Improve the Efficiency of Productive Learning Activities in Comparison to Receptive Learning Activities for Vocabulary Acquisition?

Lance Burrows

#### **ABSTRACT**

Vocabulary acquisition has been deemed by some as the most important part of any English as a Foreign Language (EFL) program. However, the limited amount of class time and exposure to English, in an EFL setting, has prompted many educators to try to improve the efficiency of their vocabulary learning activities in class. Research has shown that decontextualized productive style vocabulary activities can help improve productive, as well as, receptive vocabulary knowledge. However, these productive activities often take substantial amounts of time to complete. Time that is often not available to EFL educators. This study investigated the role that contextualization may have in the vocabulary acquisition process. By contextualizing productive tasks, can the educators expedite the vocabulary acquisition process? And once the vocabulary terms have been learned in a productive activity, does receptive vocabulary knowledge also improve?

Keywords: receptive vocabulary knowledge, productive vocabulary knowledge, contextualization, decontextualization, vocabulary learning activities

語彙習得において文脈化が受容的学習活動と比較して産出的学習活動の 効率化に与える影響について ランス ブロース

要旨 語彙習得は英語を外国語として学ぶ環境 (EFL)において最も重要な学習内容の一つと考えられてきた。しかしながら EFL では授業時間や学習者が英語に触れる機会が限られており、指導者は限られた時間の中での効果的な語彙学習活動を模索してきた。先行研究によると、非文脈化された産出的語彙学習活動は、受容的語彙知識だけでなく産出的語彙習得にも効果があるとしている。しかしながらこのような非文脈化された語彙学習活動は多くの時間を必要とするため現場での実践は困難である。このような状況をふまえ、本研究は文脈化が語彙習得過程に与える影響について検証した。具体的には、語彙の産出的学習活動を文脈化することは語彙習得を促進するのか、また産出的活動における語彙学習は受容的語彙知識の促進にもつながるのか、ということについて検証した。

キーワード:受容的語彙知識、産出的語彙知識、文脈化、非文脈化、語彙学習活動

Since Krashen (1985) introduced his Input Hypothesis in the mid-1980's, there has been a steady push to include more receptive learning activities in the foreign language (L2) classroom. In this hypothesis, Krashen advocates that by offering learners comprehensible input through receptive learning tasks (RLTs, i.e., reading and listening tasks), teachers can help stimulate L2 acquisition. He further states that RLTs are all that is necessary in promoting L2 learning; that in fact productive learning tasks (PLTs, i.e., speaking and writing tasks) are unnecessary for learner L2 acquisition.

This extreme approach to L2 learning has been criticized by some who believe that PLTs are also essential to L2 development (Swain, 1993; Nation, 1990; Joe, 1998). One type of PLT, generative tasks, for example, requires learners to use newly presented language in a personal and creatively original way (e.g., creating a sentence using a new vocabulary word). This "generation" (i.e., creating a new sentence) allows the learner to associate new language with language that has already been acquired. By integrating language that has been previously acquired with newly encountered language, the learner theoretically, is forced to more deeply process the new information and will therefore maintain longer retention and a richer, fuller knowledge of the new language (Joe, 1998).

The idea of complementing RLTs with PLTs in the L2 classroom seems to be most ideal. However, an abundance of time and ample exposure to input is not a luxury that English as a Foreign Language (EFL) classrooms share with English as a Second Language (ESL) classrooms. The need to expedite the lexical acquisition process is real and problematic in most EFL settings. So, what is the solution? Which type of task should be more heavily focused on in the L2 classroom, PLTs or RLTs?

It stands to reason that if words are learned receptively, then learners are likely to gain more receptive knowledge, whereas words that are learned productively will add to the learner's productive knowledge. But due to the limitations on time and exposure to the L2, it may be difficult to allot ample time for both types of tasks in the EFL classroom. In most practical situations, teachers do not expect students to acquire only one type of knowledge (i.e., productive or receptive). Therefore, the teacher may be forced to decide which type of task would be more beneficial and efficient for her students' L2 acquisition. The teacher may struggle with choosing which type of learning should be focused on in class.

Assuming that the teacher wishes to promote both productive and receptive language acquisition in the EFL classroom, it seems logical that PLTs would offer more practical benefits to the learner. For example, if a learner can write a new word and use it correctly in context, can we assume that he can also read it, if encountered in a reading passage? In other words, is it possible for the learner to get a "2-for-1 deal" with language? If students learn through PLTs, can they simultaneously improve their receptive abilities?

Research to date has largely focused on either receptive and productive vocabulary size (Laufer, 1998; Laufer & Paribakht, 1998; Waring 1997) or whether receptive knowledge gains lead to productive knowledge gains (Aitchison, 1994). Although little research has been done to compare receptive and productive learning, there are a few studies that have brought some important points to light.

Some researchers assert that as one learns a word, the receptive knowledge comes first and the productive follows behind. "Items pass from the learner's receptive vocabulary store to his smaller productive one all the time, some items perhaps moving back again from the productive to receptive one, when the learner forgets items he has once mastered but has not met very frequently or recently" (Ringbom, 1985: 168, as cited from Waring, 1997). This means that if a word has been acquired productively it would logically be known receptively as well. It would also mean conversely, that a word could never be produced if it had not first been receptively acquired. Furthermore, a word that could be produced could never "not" be recognized in listening or reading. This clearly is not the case. Although these occurrences may be few, it is entirely possible that a word may be available for speech but may not be recognized or recalled in reading or listening.

Waring (1997) contends that receptive and productive knowledge is not serial in nature, as Ringbom postulates. He claims that the nature of these two types of knowledge is much more complex; involving associations, and having multi-dimensional aspects, at least in part, a network-like phenomenon. Through his study, he supports his hypothesis that receptive and productive knowledge are not necessarily linearly-linked; that some words learned receptively will also be available productively and some words learned productively may be recognizable in listening or reading. In reference to the time-spent-on-task factor, he also showed that it took the students on average 26% more time to productively learn the 15 prescribed vocabulary in his study than to learn them

receptively. However, he also discovered that the students on average could only produce in written form 6 of the 15 words that had been receptively acquired, while 11 of the 15 words that had been productively acquired could be recognized and comprehended in a receptive test. There is obviously a trade-off involved. More time had been consumed by the productive learning side but in the end, it allowed for more vocabulary that could be utilized both receptively and productively.

Webb (2004) found similar results in two studies that he conducted on Japanese students studying English as a foreign language in Kyushu, Japan. He was testing the effects of receptive and productive vocabulary learning on word knowledge. Students were required to learn target words in three glossed sentences (receptive learning) and in a sentence production task (productive), at different times. Following these learning scenarios, receptive and productive tests were conducted to measure five aspects of vocabulary knowledge-orthography, syntax, association, grammatical factors, and meaning and form. Overall, the words were tested in ten different ways.

The design of Experiment 1 entailed dividing the participants into two groups: the receptive task group and the productive task group. The participants in the receptive task group were given 10 nonsense words and shown the L1 equivalent and three example sentences for each one. In each of the three example sentences, the target vocabulary were underlined and written in bold. For this task, the participants were told to simply learn the words. In the other half of the group, the participants in the productive task group were given the same 10 vocabulary items and L1 equivalents, but not the three example sentences. Instead, the productive task group was required to generate 10 original sentences using the 10 prescribed vocabulary items. Both the receptive and productive task groups were given 12 minutes to complete their tasks. They were also informed that they would be tested on these items after completing the tasks. However, they were not told how they would be tested.

In his second study, he organized a completely new group of participants (different from the participants in Experiment 1) and designed the treatment so that the new group received both the receptive and productive treatments. Of course, in these treatments different nonsense vocabulary words were used (i.e., ten words were used for the receptive treatment and ten different words were used for the productive treatment). In this study, the students were only given enough time to complete the task at hand and were instructed to continue on to the next task as soon as they had

finished the first. In other words, the students were not allowed to review their work after they had completed their tasks. They were also not told that they would be tested at the end of the two tasks. The instructions from Experiment 2 were the same for the tasks in Experiment 1: in the reading task, learners were told to simply learn the terms, while in the production task, they were told to write each one in a sentence.

In the end, the first experiment showed that when time was limited and equally set for both productive and receptive tasks, the receptive learning task was superior on all levels of evaluation. However, in the second experiment, where participants were allowed enough time to completely finish their tasks, it was found that the productive tasks were more superior. Webb postulates that if the second experiment represents authentic learning, productive tasks are more practical.

These two studies which are very similar in many ways both emphasize the element of time as a crucial factor in the discussion of receptive/productive learning. Furthermore, both support the notion that productive learning may be more efficient than receptive learning when an increase in productive as well as receptive knowledge is the goal of a learning scenario and students are given ample time to practice productively. One element that seems to be missing in these two studies is the subject of contextualization of the prescribed vocabulary. In both studies, the vocabulary set has been presented, learned and tested in a de-contextualized manner.

Nation (2001: 353) tells us, "The value of context may be to orient the learner to the correct part of speech and, by more closely resembling conditions of normal use, encourage normal access to the meaning." Studies of guessing from context have shown that there are strong relations between guessing skills and vocabulary knowledge, reading skill (Herman, Anderson, Pearson, and Nagy, 1987), reading comprehension and verbal IQ (Hafner, 1967). However, little research has shown how contextualization during productive vocabulary acquisition tasks may serve to provide concepts to which new vocabulary may be associated. This possibility of increased association may prove beneficial in productive as well as receptive vocabulary acquisition.

The following hypotheses directed this study:

- 1. Contextualization of vocabulary will result in higher vocabulary acquisition in both receptive (reading) and productive (writing) tasks.
- 2. With cognitive connections that are made from the contextualization of vocabulary in the receptive and productive tasks, productive tasks will become

superior in creating productive and receptive vocabulary knowledge to receptive tasks, even if time-on-task is held constant.

#### The Study

# **Participants**

The participants in this study were 34 Japanese native speakers from two first-year English as a foreign language (EFL) classes at a co-ed university in Osaka, Japan. Their mean score on the Vocabulary Levels Test (Schmitt, 2000) at the second 1,000-word level was 28.1/30, which indicated that they were in control of almost all the 2,000 most frequent words (Schmitt, 2000).

# Definition of receptive and productive learning

Because this study was conducted using a contextualized setting, to demonstrate "receptive knowledge" of a word, the subject had to provide a specific L1 translation of a particular target English word which in essence, showed an ability to recognize the word and recall the meaning learned (L2 to L1). "Productive knowledge" was tested by requiring the participants to retrieve a L2 equivalent for a particular target English word, which in essence, showed an ability to recall the equivalent L2 word and have control over its spelling (L1 to L2). In both tests, the participants were required to provide the target vocabulary within a contextualized setting. Because the activities leading up to the testing were conducted with contextualized situations, the tests were also conducted that way.

### Study Design

There were four treatments in the study: a de-contextualized receptive activity, a de-contextualized productive activity, a contextualized receptive activity, and a contextualized productive activity. The de-contextualized receptive and productive activities closely resembled the test design of Waring (1997). This study utilized 15 words taken from one vocabulary word level above the participants' current level (i.e., a subject who tested into the 2000 word level group, was tested using words from the 3000 word level). A pre-test, of 20 words, was also conducted to insure that the participants were encountering the test vocabulary for the first time. There were two words which 4 of the participants were familiar with so those words were omitted in

the activities and further tests.

Word cards were utilized to conduct the de-contextualized activities and tests. All 15 words for the receptive test and all 15 words for the productive test will be printed with the L1 translations on the opposite side. The L1 side will be pink and the L2 side will be blue. For the receptive activity, participants will be learning from L2 (blue side)  $\rightarrow$  L1 (pink side) and L1 (pink side)  $\rightarrow$  L2 (blue side) for the productive activity.

Because the study was a counterbalance study, in the first week of the study, Class A was given the de-contextualized receptive task of memorizing vocabulary from  $L2 \rightarrow L1$ . This task was conducted with participants learning from word cards, with no contextualization for assistance. They were given the word cards and told only to look at the blue side of the cards first and then the pink side of the cards second as they tried to memorize each word. The participants were told to try to learn the words until they were able to go through the 15 word set 2 times without mistakes. Learning the words also meant that students would be able to write the L1 and L2 translations, so they were encouraged to learn them by writing. The students were never told that they would be tested on the target words. There was no limit set on the amount of time that they could take to complete the task. After all students had finished memorizing, participants ceased their memorizing and were given a short, 10-minute activity completely unrelated to the study or the words which they had been exposed to. This was meant to serve as a distracter. After this 10-minute interval, the participants were given surprise receptive and productive vocabulary knowledge tests. The test was of a simple design. First, a list of the words in the L1 from the activity was provided on a paper and the participants were required to write the L2 meaning. Then, the receptive test was given. Participants were provided with the L2 and they were asked to offer the L1 translation. The productive test was given first because it allowed for less exposure to the words than the receptive test. Additional exposure to the words from a receptive test might have further reminded the participants of the terms and ruined the integrity of the study design. A follow-up test was given at 2 weeks to further track the participants' ability to recall. For each of these subsequent tests, the order in which the words were placed on the page was changed to protect against testing effects.

Also within the first week, Class B was given the de-contextualized productive task. They were given the cards and told only to look at the pink side of the cards first

and then the blue side of the cards second as they tried to memorize each word. The participants were told to try to learn the words until they were able to go through the 15 word set 2 times without mistakes. Again, participants were given as much time as they needed to complete the task. There was no time limit. When all participants claimed to have finished, they were asked to cease their memorizing and were given a short, 10-minute activity completely unrelated to the study or the words which they had been exposed to. After this 10-minute interval, the participants were given surprise receptive and productive vocabulary knowledge tests. The test were of simple design. First, a list of the words in the L1 from the activity were provided on a paper and the participants were required to give the L2 meaning. Then, the receptive test was given. Participants were provided with the L2 and they had to offer the L1 translation.

The productive test was given first because it allowed for less exposure to the words than the receptive test. Additional exposure to the words from a receptive test might have further reminded the participants of the terms and ruined the integrity of the study design. A delayed test was given 2 weeks after the initial test to further track the participants' ability to retrieve. Again, for each of these subsequent tests, the order in which the words were placed on the page was changed to protect against testing effects.

In the second week, Class A was given the de-contextualized productive treatment and tested on it, and Class B was given the de-contextualized receptive treatment and tested on it, using the same formats as the above mentioned.

After finishing the initial de-contextualized receptive and productive tasks/tests, participants in Class A and B began the contextualized activities and tests. In the third week, participants of Class A were shown a series of 12 pictures on a page for the contextualized receptive activity. The pictures described a familiar topic to most Japanese students, "How to make okonomiyaki". Below the set of pictures were 12 sentences explaining the pictures in English. Within these sentences, the target vocabulary terms were bolded and underlined to promote salience to the learner. Also, under each of the target words, the Japanese translation was given. The participants were asked to first put the sentences in order. Then there was a set of ten multiple-choice comprehension questions that the participants had to complete. Again, there was no time limit. Then again, after a brief, 10 minute activity in class they were given the same kind of receptive/ productive tests that were offered in the de-contextualized

versions of this study.

In the same week, Class B began the contextualized productive section. In this situation, participants were also asked to talk about food, but this time, it was about "How to make curry rice". In this setting, the participants again were shown a set of pictures leading them through the recipe of curry rice and were asked to make their own sentences. They were required to use the target words that had been given below the pictures. Each target vocabulary word was listed with the L1 translation under it. They were asked to make sentences using commands. For example, "First, peel the onion. Then, slice the carrots into small pieces." After all students had finished, students were asked to stop the activity and participate in another distracter activity for ten minutes. After the ten minutes, they were given the same type of receptive/productive test that they have received for all other parts of the study. These tests for the contextualized parts of the test were also given two weeks after the initial test to further track the participants' vocabulary knowledge.

In the fourth week, Class A was given the contextualized productive treatment and tested on it, and Class B was given the contextualized receptive treatment and tested on it, using the same formats as the above mentioned.

### Results

The descriptive statistics for both the productive and the receptive test scores are reported in Tables 1 and 2, respectively. A repeated-measures analysis of variance (ANOVA) was used to analyze the variance among the different treatments. The scores reported here are from the delayed tests, which were taken 2 weeks after the initial treatment. It is thought that the delayed tests offer a more accurate reading of the participants' vocabulary acquisition.

Table 1

Descriptive statistics of contextualized and de-contextualized production activities

Treatments	DPP	DPR	CPP	CPR
M	5.18	7.09	4.91	5.62
95% LB	4.51	6.21	4.04	4.90
95% UB	5.84	7.97	5.79	6.33
SD	1.90	2.53	2.47	2.05
Skewness	0.352	0.685	0.143	-0.010
SES	0.403	0.403	0.403	0.403
Kurtosis	-0.146	0.244	-0.698	-0.790
SEK	0.788	0.788	0.788	0.788

Note. DPP = de-contextualized production activity, tested on productive knowledge; DPR = de-contextualized production activity, tested on receptive knowledge; CPP = contextualized productive activity, tested on productive knowledge; CPR = contextualized production activity, tested on receptive knowledge; 95% LB = 95% confidence interval lower bound; 95% UB = 95% confidence interval upper bound; S = skewness; K = kurtosis; SES = standard error of skewness; SEK = standard error of kurtosis.

Table 2

Descriptive statistics of contextualized and de-contextualized receptive activities

Treatments	DPP	DPR	CPP	CPR	
M	2.65	5.91	3.97	5.85	
95% LB	2.15	4.93	3.43	5.23	
95% UB	3.15	6.89	4.51	6.48	
SD	1.43	2.80	1.55	1.79	
Skewness	-0.052	0.642	0.208	-0.334	
SES	0.403	0.403	0.403	0.403	
Kurtosis	-0.792	0.591	-0.720	-0.777	
SEK	0.788	0.788	0.788	0.788	

As was mentioned above, a repeated-measures ANOVA was conducted with the factor being de-contextualization/contextualization and the dependent variable being the vocabulary test scores for the receptive/ productive tests. The productive measures will first be focused on here. The results for the ANOVA indicated a significant effect for productive activities, Wilk's  $\Lambda$  =.44, F(3, 31) = 12.97, p < .01, multivariate  $\eta^2 = .56$ . Follow-up polynomial contrasts indicated a non-significant effect for linearity, but a significant effect for quadratic and cubic indices. The original hypothesis that CPP > CPR > DPP > DPR was not confirmed. Instead, the mean scores were DPR > CPR > DPP > CPP (see Table 1 *Notes*).

Post hoc analyses to the ANOVA for the recall scores consisted of conducting pairwise comparisons to find which treatment affected vocabulary acquisition most strongly. Using the Bonferroni adjustment to control for Type I error, the pair-wise comparisons were tested at .01 divided by 6 or .0017 level. It was found that in paired-wise t-tests, two comparisons were significant. The de-contextualized productive activities when testing for receptive knowledge (M = 7.10, SD = 2.53) had significantly more effect over the de-contextualized productive activities when testing for productive knowledge (M = 5.18, SD = 1.90), t(33) = -6.38, p < .01. In these comparisons, the 95% confidence interval for the mean difference between the two ratings was -2.52 to -1.30. Furthermore, the de-contextualized productive activities when testing for receptive knowledge (M = 7.10, SD = 2.53) were also significantly more effective over the contextualized productive activities when testing for productive knowledge (M = 4.91, SD = 2.50), t(33) = 3.70, p < .01. These finding are not consistent with the hypothesis, that contextualization should boost productive and receptive acquisition more than dedecontextualization.

Again a repeated-measures ANOVA was conducted with the factor being decontextualization/ contextualization and the dependent variable being the vocabulary test scores for the receptive/ productive tests. Next the receptive side shall be analyzed. The results for the ANOVA indicated a significant effect for receptive activities, Wilk's  $\Lambda$  =.15, F(3, 31) = 60.10, p < .01, multivariate  $\eta$  <sup>2</sup>= .85.

Follow-up polynomial contrasts indicated a significant effect for linearity, F (1, 33) = 79.20, p < .01, partial  $\eta^2 = .71$ . The original hypothesis that CRP > CRR > DRP > DRR was not confirmed. Instead, the mean scores were DRR > CRR > CRP > DRP (Table 2 *notes*).

Post hoc analyses to the ANOVA for the recall scores consisted of conducting pair-wise comparisons to find which treatment affected vocabulary acquisition most strongly. Using the Bonferroni adjustment to control for Type I error, the pair-wise comparisons were tested at .01 divided by 6, or .0017 level. It was found that in paired-wise t-tests, all but one comparison were significant (Table 3). These finding are not consistent with the hypothesis that contextualization should boost productive and receptive acquisition more than de-contextualization.

Table 3

Paired sample t test results for receptive activities

Pairs	DRP/DRR	DRP/CRR	DRP/CRP	DRR/CRP	CRR/CRP
M	-3.26	-3.21	-1.32	1.94	1.88
95% LB	-4.15	-3.84	-1.85	.94	1.52
95% UB	-2.38	-2.57	80	2.94	2.24
SD	2.53	1.82	1.51	2.87	1.04
t value	-7.54	-10.26	-5.10	3.94	10.58
df	33.00	33.00	33.00	33.00	33.00

*Note.* DRP = de-contextualized receptive activity, tested on productive knowledge; DRR = de-contextualized receptive activity, tested on receptive knowledge; CRP = contextualized receptive activity, tested on productive knowledge; CRR = contextualized receptive activity, tested on receptive knowledge; 95% LB = 95% confidence interval lower bound; 95% UB = 95% confidence interval upper bound; for all scores p < .01.

### Discussion

When considering the discrepancies between the original hypotheses and the actual results it is important to consider both the productive and receptive activities scores. If we line the full set of eight score on one continuum, it would look something like this: DPR (M=7.08, SD=2.53) > DRR (M=5.91, SD=2.80) > CRR (M=5.85, SD=1.79) > CPR (M=5.62, SD=2.05) > DPP (M=5.18, SD=1.90) > CPP (M=4.91, SD=2.50) > CRP (M=3.97, SD=1.55) > DRP (M=2.65, SD=1.43). First of all, we can see that the receptive tests were all higher. That would stand to reason, that receptive tests would seem to be easier for participants to do well on. The productive tests scores were, logically lower on the scale. The fact that the decontextualized scores were higher for the receptive tests, shows that Japanese students are extremely adept

at memorizing. This delayed check was only given two weeks after the initial exposure to the vocabulary terms. This short interval may have had an effect on the outcome. For future research, a more lengthy interval might possibly produce different results. Also, the fact that the participants were not asked to actually use the vocabulary (just recall it) also prompts a question about how well the participants would be able to actually utilize the vocabulary had it been in a real-life situation. They seem to be able to memorize the vocabulary fairly well, even in the decontextualized situation, but could they also at some point put that knowledge to use. This remains a question.

### **REFERENCES**

- Aitchison, J. (1994). Words in the Mind. Oxford: Blackwell.
- Hafner, L. E. (1967). Using context to determine meanings in high school and college. *Journal of Reading*, 10, 491-498.
- Herman, P. A., Anderson, R.C., Pearson, P.D., and Nagy, W.E. (1987). Incidental acquisition of word meaning from expositions with varied text features. *Reading Research Quarterly*, 22, 263-284.
- Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, 51, 539-558.
- Joe, A. G. (1995). Text-based tasks and incidental vocabulary learning. *Second Language Research*, 11, 149-158.
- Joe, A. G. (1998). What effects do text-based tasks promoting generation have on incidental vocabulary acquisition? *Applied Linguistics*, 19, 357-377.
- Krashen, S. (1985). The Input Hypothesis: issues and implications. London: Longman.
- Laufer, B. (1998). The development of passive and active vocabulary in a second language: same or different? *Applied Linguistics 19*: 255-271.
- Laufer, B., & Paribakht, S. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning*, 48, 365-391.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Boston: Heinle & Heinle Publishers.
- Nation, P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the vocabulary levels test. *Language Testing*, 18, 55-88.
- Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Review*, 50, 158-164.
- Waring, R. (1997). A study of receptive and productive vocabulary learning from word cards. *Studies in Foreign Language and Literature*, 21, 94-114.
- Webb, S. (2004). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. Studies in Second Language Acquisition, 27, 33-52.