

A Replication of “The Acquisition of Vocabulary in an Autonomous Learning Environment — the First Months of Beginning English” by Leni Dam and Lienhard Legenhausen

Michiko Ono

(Received October 22, 2004)

1 Introduction

The study is concerned with vocabulary acquisition in the first few months of the first year in senior high school. The main purpose of the study is to ascertain whether autonomous activities help students learn vocabulary better. Two kinds of tests were administered to learners — an informal recall test and a systematic vocabulary test with 80 items. This paper reports research that the writer carried out at her previous institution.

2 The Learner Groups

The subjects are Japanese first-year students at Niigata Chuo Senior High School. They are 15 or 16 years old. In Japan, senior high school is not compulsory, and entrance examinations are required. The school is one of the most prestigious schools in Niigata Prefecture, and the students' abilities are somewhat similar. Three groups were involved in the experiment. Each group consisted of 41 girls of mixed ability. English classes are taught six times a week, equally spread from Monday to Saturday. Each period is 50 minutes long. The students have so far taken three tests — an English ability test, the mid-term and final exams in English. The results are shown in Table 1.

Table 1: The Group Mean Scores of the English Ability Test and the Mid-Term and Final Exams in English (Full mark: 100, SD's in the parentheses)

Tests	Autonomous Group	Traditional Group 1	Traditional Group 2
Ability Test	85.32 (6.72)	83.34 (7.50)	85.34 (6.34)
Mid-term Exam	71.56 (9.70)	73.00 (10.18)	72.33 (12.44)
Final Exam	63.54 (9.68)	65.85 (13.65)	68.26 (13.38)

The three groups are taught English based on the grammar-translation method. Their learning back-

grounds are quite similar.

The autonomous learners experience autonomous learning only in their English classes, all the other subjects taught in the traditional way. Only this group was administered both the first and the second tests.

The first group of traditional learners is a control group. Unlike in the original project, this group did not take part in the first test, which was a spontaneous production, but in the second test alone.

The second group of traditional learners is a reference group, whose English course is also textbook-based, as is that of the other two groups, but this group is the best among the three as far as English exams are concerned. Since students are not differentiated by their abilities or aptitude in forming classes, the learning background of this group is practically the same as that of the other two groups. This group also took part only in the second test.

3 Vocabulary Acquisition in the Autonomous Classroom in the First Few Weeks

3.1 Learner Autonomy and Language Learning

An important task of the teacher is to try to raise learners' awareness of the many ways to gain new knowledge about language, and to help learners establish the situations where they can actuate and expand their knowledge.

3.2 Building Up a Shared Language

Following the original project, five methods were applied to autonomous learners.

- (1) The learners were asked to bring something on which were written some English words or sentences, such as T-shirts and pencil cases.
- (2) The learners were asked to write down in their own diaries some English sentences introducing themselves. They were allowed to use an English-Japanese dictionary and a Japanese-English dictionary. The expressions they used were, for example, "My name is ...," "I am...years old," "I like...," and so forth. However, these expressions did not seem quite adequate, for the subjects in the replication were in their fourth year of learning English while those in the original study were in their first year. So in this experiment the learners were encouraged to write other more complex English sentences in their diaries, using expressions they had learned from the textbook.
- (3) The students were asked to pick out at least 5 words they wished to remember from Lesson 4 of their English textbook. The lesson is about astronauts in space. The students were asked to write the words down in their notebooks and show them to their neighbor.
- (4) Language input was introduced by the instructor in the form of a well-known fairy tale read aloud —"Jack and the Beanstalk." Expressions and structures in the story that the instructor and learners found interesting or useful were typed to make a handout.

- (5) The learners produced language by writing their favorite words on simple wordcards each with L1 equivalents or drawings on the other side. The words were supposed to be the five words they had picked out and others from “Jack and the Beanstalk.” Those who had a picture book of the story were asked to bring it to class. (Nevertheless, no one did, so the instructor photocopied some pages from her own English fairy tale book with some pictures, and handed them out to each student.)

4 Size and Distribution of Vocabulary

In order to make clear the process of vocabulary acquisition in the first eight weeks, all the new and important words that were presented to learners in classrooms were typed out.

4.1 Number of Words

The typed-out vocabulary was classified into five groups—verbs, nouns, adjectives, adverbs, and others. The textbook used by the subjects was *Milestone English Course I*, published by Keirinkan Publishing Company (Michikazu Kaneda et al., 1993), which implemented the official English curriculum prescribed by the Ministry of Education of Japan. In this classification of vocabulary, semantic fields were given no consideration.

Table 2: Distribution of New Vocabulary across Parts of Speech

Parts of Speech	Total after 20 periods	Total after 32 periods	LESSON 4
Verbs	20	30	10
Nouns	24	41	26
Adjectives	7	13	5
Adverbs	4	5	6
Conjunctions	1	1	1
Others	28	46	26

The mid-term exam held after 20 periods covered Lessons 1 and 2. Lesson 4 was an experimental lesson.

4.2 Semantic Fields

To classify the words above in a semantic way may not sound too effective, because the four lessons students have learned are all concerned with people, the titles being “A Language for Everyone,” “Tara in Japan,” “Jim Abbot,” and “Upside Down in Space.” The lessons are about people who use English as a means of communication, an American exchange student in Japan, a professional baseball player, and astronauts, respectively. Nevertheless, following the original study, nouns will be categorized into seven semantic groups—people, animals, parts of the body, colors, numbers, food, and school items :

Table 3: Distribution of New Vocabulary across Semantic Fields

Fields	Total after 20 periods	Total after 32 periods	L. 4
People	18	32	14
Animals	0	0	0
Parts of the body	0	0	2
Colors	0	0	0
Numbers	0	0	0
Food	1	1	1
School items	0	1	0

There are some ten nouns which could not be classified in the groups above: "result," "series," "weightlessness," "gravity," "spot," "middle," "element," and "danger." Since it is obvious that most of the nouns are related to people, it is doubtful whether Table 3 is meaningful.

5 Receptive Knowledge and Availability of L2 Vocabulary

In order to examine the amount of vocabulary acquired by learners, two vocabulary tests were administered to the autonomous group.

5.1 Spontaneous Recall Test

The first test—which was administered after eight weeks (32 periods)—was an informal test, asking the learners to recall as many new words as they could and jot them down in their notebooks. The following cues were given:

1. Write down as many words as you can.
 - a. verbs
 - b. adjectives
 - c. adverbs
 - d. conjunctions
 - e. words for persons
 - f. things that you can eat

The students were given about five minutes. The instructor walked around the classroom to see how they were doing. Some students wrote more than 30 words, others less than 10. The average seemed somewhere between 10 and 20. The notebooks were not collected because the nature of the test was elicitation rather than the usual kind of test.

5.2 Vocabulary Test 1 Administered in June 1997 — Graded Strictly

Next, a vocabulary test on Lesson 4 was administered to all three groups on June 19. In this test were included 20 verbs, 36 nouns, 6 adjectives, 6 adverbs, and 12 others which are idiomatic phrases. In total there were 80 questions, worth one point each. The results are shown below. The figures indicate group mean scores.

Table 4: The Group Mean Scores of Vocabulary Test 1 (Inside the brackets are full marks ; inside the parentheses are SD's ; Gr stands for Group.)

Parts of Speech	Autonomous Gr	Traditional Gr 1	Traditional Gr 2
Verbs [20]	6.62	5.90	7.63
Nouns [36]	14.70	15.05	16.58
Adjectives [6]	1.32	1.13	1.53
Adverbs [6]	1.19	1.31	1.45
Others [12]	6.05	4.79	6.63
Total Score [80]	30.03 (8.88)	28.13 (9.85)	33.79 (12.76)

The students were not notified of the administration of the test in advance. The highest score was 68, which was obtained by a student in Traditional Group 2. The results are shown in percentages below :

Table 5: The Results of Vocabulary Test 1 in Percentages

Parts of Speech	Autonomous Gr	Traditional Gr 1	Traditional Gr 2
Verbs	34.9%	31.0%	38.2%
Nouns	40.8%	41.9%	46.1%
Adjectives	22.1%	18.8%	24.2%
Adverbs	19.8%	21.8%	24.2%
Others	51.0%	39.9%	55.3%
Total Score	37.2%	35.0%	42.2%

The distribution of each learner's total score follows :

Table 6: Distribution of Total Scores

Score									
	1 ~ 9	10~19	20~29	30~39	40~49	50~59	60~69	70~80	Total
Groups									
Autonomous Group	1	4	13	15	3	1	0	0	37
Traditional Gr 1	1	6	16	11	4	1	0	0	39
Traditional Gr 2	0	3	13	10	9	2	1	0	38

In each group a few students were absent on the day of the test because of sickness or inter-school sport tournaments.

5.3 Discussion

The results of the study show that the autonomous learning did work in some way. It is clear that Traditional Group 2, whose mean score was the highest in the English ability test, did best. But the Autonomous Group, which was the worst of the three in the two previous tests, did better than Traditional Group 1, the control group. The former was especially better than the latter in verbs, adjectives, and idiomatic phrases. However, there were some interesting but dispiriting findings. Some students, low-ability students in particular, were affected by the instructor's secondary explanations:

- (1) When the instructor taught the conjugation of the verb "win" with phonetic signs, she wrote the sign "[wʌn]" for "won" on the blackboard, reminding the students that its pronunciation is the same as that of "one," and giving other examples like "love," "fun," and "run." In the test, many students wrote "wun" instead of "won."
- (2) When the students were extremely busy preparing for the field day, which was the biggest school event of the year, few of them were prepared for the class, and none seemed to know the meaning of the verb "circle." So the instructor gave its meaning in Japanese and, by way of explanation, also gave the meaning of the noun form. As a result, many students wrote the noun meaning of the word when it was listed in the verb group in the test. Similar phenomena are found in other cases too.
- (3) When some extra examples are given, some students tend to remember the examples rather than the target vocabulary itself.

There were some other findings:

- (4) When students look up a word in their dictionary and pick up a wrong meaning, it is often retained in their memory for a long time even after it has been corrected in class, probably because of the intensity of the first input.
- (5) When two words appear together a few times, students tend to take the meaning of one word for the other. For example, many took the meaning of "lid" for "jar"; some mixed up "float" with "drift," perhaps because they are fairly close in meaning.
- (6) Some low-proficiency students mistook "dangle" for "danger," and "tray" for "try" because they look alike.

If similar tests would be given to the students in the future, perhaps more interesting facts would be discovered. The problems listed above will be discussed later.

5.4 Vocabulary Test 1 — Graded Leniently

First the instructor marked the papers strictly, as is always the case with exams in the school. So the answers with only one misspelled letter received no credit. Among these were "wun," "hight," "colege," and "space shutle." These errors, though wrong in spelling, cause very little confusion as far as communication is concerned because they can be pronounced correctly. It seems unfair to treat them in the same way as completely wrong answers or no answers. So the instructor marked the papers again leniently. The rules were as follows :

- (1) Words with only one misspelled letter are treated as right answers and are given one point each, as long as they can be pronounced correctly. Therefore, "hight" gets one point, whereas no point is given to "feight" which is supposed to mean "height." Similarly, "fload" for "floor" is all right, but "peanute" for "peanut." or "racket" for "rocket" are not.
- (2) In the test, vocabulary is classified into five groups—verbs, nouns, adjectives, adverbs, and others, as they appear in the textbook. If a student gives a noun meaning when a word is supposed to be a verb, no credit is given when the paper is graded strictly, but one point is given when graded leniently. So, even if a student gives a noun meaning to such verbs as "circle" and "toss," or a verb meaning to such nouns as "challenge" and "interview," the answer is treated as correct in the latter case.
- (3) Though most of the words in the test are polysemous, only one meaning is expected as a correct answer. However, when the paper is marked leniently, other meanings are also considered as correct.
- (4) Students are supposed to translate a few short sentences. If they use the wrong tense in translation, normally they do not get any credit, but when the paper is graded in a lenient way, they get one point.

The results of lenient grading were as follows :

Table 7: The Results of Vocabulary Test 1 — Graded Leniently as Compared with Those Graded Strictly (Inside the parentheses are SD's; full mark is 80.)

Groups		Graded Harshly	Graded Leniently
Group Mean Score	Autonomous Group	30.03 (8.88)	37.03 (9.43)
	Traditional Group 1	28.13 (9.85)	33.38 (11.68)
	Traditional Group 2	33.79 (12.76)	39.34 (12.80)

Table 8: The Extra Points Obtained from Grading Leniently (Inside the parentheses are SD's.)

Groups	Group Mean
Autonomous Group	7.05 (2.84)
Traditional Group 1	5.26 (2.26)
Traditional Group 2	5.53 (2.32)

5.5 Implication

Tables 7 and 8 indicate that the Autonomous Group got the most points by the lenient grading. Table 8 shows each group's difference in score between the two ways of grading. Extra points each student got from the lenient grading vary from 0 to 12. From these tables we infer that the Autonomous Group was superior to Traditional Group 1 not only in the number of correct answers but also in the number of nearly correct answers. The nearly correct answers are very useful in communication, for even though the students can not spell the words with complete accuracy, people can understand them when they are spoken, and vice versa. Therefore it may not be too much to say that autonomous learning has worked in a considerable way.

6 The First-Term Final Exam in English

Just before the final exam, students in each group were given an opportunity to work on the same vocabulary test on their own at school when the instructor was not around. The papers were not collected.

The final exam in English was held on July 9. There were 13 different kinds of questions, and about 10 of them could be classified as vocabulary questions. The following were the results :

Table 9: The Results of the First-Term Final Exam in English
(Inside the parentheses are SD's; inside the brackets are full marks.)

Groups		Vocabulary questions [70]	Total score [100]
Group Mean Score	Autonomous Group	44.88 (6.62)	63.54 (9.68)
	Traditional Group 1	46.63 (9.49)	65.85 (13.65)
	Traditional Group 2	47.77 (9.71)	68.26 (13.38)

6.1 Discussion

It is quite clear that, in terms of the total group mean score, Traditional Group 2 is the best, Traditional Group 1 the second best, and the Autonomous Group the worst. The same thing can be said about the results of the vocabulary questions. However, in terms of the proportion of the vocabulary score to the total score, the situation is quite the opposite. The Autonomous Group is the highest, Traditional Group 1 the second highest, and Traditional Group 2 the lowest. The percentages are 71.04%, 70.71%, and 69.99%, respectively. Therefore, we may safely say that the Autonomous Group learned vocabulary well in spite of their low scores.

7 Vocabulary Test 2 Administered in July 1997

The same vocabulary test (full mark:80) was administered to the three groups again without any announcement in advance. Students took the test on two different days because of the class schedule—the Autonomous Group and Traditional Group 2 on July 17, and Traditional Group 1 on the

next day. The results were as follows :

Table 10: The Results of Vocabulary Test 2—Graded Strictly and Leniently
(Inside the parentheses are SD's.)

Groups		Graded Strictly	Graded Leniently
Group Mean Score	Autonomous Group	37.67 (9.71)	45.87 (10.90)
	Traditional Group 1	38.41 (14.59)	47.88 (16.23)
	Traditional Group 2	38.69 (15.29)	53.13 (14.07)

The mean of the three scores is 38.26 when graded strictly, and 48.73 when graded leniently. The distribution of each learner's total score follows :

Table 11: Distribution of Total Score of Vocabulary Test 2—Graded Strictly

Score	Groups								
	1 ~ 9	10~19	20~29	30~39	40~49	50~59	60~69	70~80	Total
Autonomous Group	0	1	7	14	13	3	1	0	39
Traditional Group 1	1	3	8	11	6	10	2	0	41
Traditional Group 2	0	1	7	8	8	9	5	1	39

Table 12: Distribution of Total Score of Vocabulary Test 2—Graded Leniently

Score	Groups								
	1 ~ 9	10~19	20~29	30~39	40~49	50~59	60~69	70~80	Total
Autonomous Group	0	0	3	7	15	10	3	1	39
Traditional Group 1	1	2	2	6	10	9	7	4	41
Traditional Group 2	0	0	3	5	7	8	10	6	39

Extra points each student got from grading leniently vary from 0 to 20. Group means are shown below.

Table 13: Extra Points Obtained from Grading Leniently (Inside the parentheses are SD's)

Groups	Group Mean
Autonomous Group	8.82 (3.65)
Traditional Group 1	9.41 (3.99)
Traditional Group 2	9.41 (3.45)

7.1 Discussion

The results of Vocabulary Test 2 were better than those of Vocabulary Test 1, perhaps because students studied hard for the final exam which was held 10 days before Vocabulary Test 2.

The highest score was 72, when graded strictly, and 79, when graded leniently. They were obtained by the same student in Traditional Group 2 (TG 2). The lowest score was 9, and it belonged to a student in Traditional Group 1 (TG 1). This student got no extra points from the lenient grading.

Though Traditional Group 2 is still the best of all in the group mean score, Table 12 shows interesting facts:

- (1) Despite the fact that the Autonomous Group (AG) is the lowest in the group mean score, there are no students in this group who got credits less than 20, which is as good as the best group, TG 2. In addition, the lowest score of AG is 26, the same as that of TG 2.
- (2) The number of students whose scores are 40 or more and less than 70 are 28 in AG, 26 in TG 1, and 25 in TG 2. This implies that AG has more learners who are around or above the average than the other two groups though AG does not have very many excellent students like TG 2.

It must be added that the two students in AG who were absent on the day Vocabulary Test 2 was administered were very good students, while another two in TG 2 who did not take the test were very poor students. If the four students had taken the test, the results would probably have been somewhat different.

Similarly, Table 11 shows :

- (3) The number of students who got 30 points or more and less than 60 is also higher in AG than the other two groups, the number being 30 in AG, 27 in TG 1, and 25 in TG 2. Surprisingly, if the number of students who got 60 or more and less than 70 is added, the number would be 31 for AG, 29 for TG 1, and 30 for TG 2, which means that AG is the best as far as the number of good students is concerned. We may say that TG 2 is the best in the group mean score, because it has a greater number of students whose scores are very high. The results of Vocabulary Test 2 also reveal that the lowest score of AG is 19, while that of TG 2 is 17.

Other findings obtained from Vocabulary Test 2 are :

- (4) The lowest student did not get any extra points when graded leniently, and lower students did not get many extra points, either, whereas students whose scores were average or better received more than 10 extra points.
- (5) Some of the students who got very high scores in Vocabulary Test 2 had not got very high scores in Vocabulary Test 1. When the papers were strictly graded, there were almost 25 students whose scores in Vocabulary Test 2 were more than 20 points higher than the scores they had got in Vocabulary Test 1.

- (6) Concerning the extra points, AG's mean is the lowest in Vocabulary Test 2, the other two group means being the same. Nevertheless, the difference in group mean between AG and the other two is only 0.59. This figure is very low, compared with the differences in Vocabulary Test 1: the difference between AG and TG 1 was 1.79, and that between AG and TG 2 was 1.52.
- (7) Marking papers leniently can be very difficult in terms of deciding which items to accept and which to reject in a fair way. For example, in the vocabulary test, "roket" is accepted when graded leniently, but "rocet" is not, because "roket" can be pronounced like "rocket" but "rocet" is not, though both erred in only one letter. Nevertheless, the students who wrote "rocet" might pronounce it like "rocket." Then it would be unfair to grade them differently. In like manner, "sholder" is accepted but not "shorder," though those who wrote "shorder" might pronounce it like "shoulder."

8 Conclusion

As a whole, we may conclude that the autonomous approach was successful to a considerable extent as the results of the vocabulary tests show.

Regarding the problems listed in 5.3 and 5.4, if we attempt more experiments, we may find some feasible solutions.

It was not easy to try to follow each step of the original study in the replication, because sometimes this researcher did not quite understand why certain processes were necessary, or sometimes did not agree with the way the experiment was to be carried out. But this has been a good experience, and hopefully, more experiments concerning autonomous learning should be tried in the near future.

References

- Kaneda, Michikazu et al.. 1993. *Milestone English Course I*. Keirinkan Publishing Company.
- Dam, Leni and Lienhard Legenhausen. 1996. "The acquisition of vocabulary in an autonomous learning environment — the first months of beginning English" in *Taking Control: Autonomy in Language Learning*. Edited by Richard Pemberton, Edward S. L. Li, Winnie W. F. Or and Herbert D. Pierson. Hong kong University Press.