Lexical Attrition in Japanese University Students: A Case Study

OKAMOTO, Mayumi
Setsunan University

Abstract

This paper aims to see how university students' English lexical ability changes over time by examining three lexical dimensions of breadth, depth, and retrievability. The research spanned half a year, yielding the following three results: first, students' average receptive vocabulary contained 5,895 words, whereas the words beyond the 2,000 level they produced accounted for only 19.8% of their total production. Second, their performance was poorest on the collocation test. Third, their lexical knowledge suffered marked attrition in all three dimensions, with 25% attrition in the receptive vocabulary, 41% in the collocational knowledge, and 15% in retrievability. This study raises some questions concerning language curriculum development.

Introduction

It is a commonplace assumption that university students' English proficiency is at its highest at the time of the entrance examination, after which it is said to decline rapidly. If this is the case, it is a serious problem for teachers of English and curriculum planners. Despite the seriousness of the issue, there does not seem to be much research addressing how university students' English proficiency changes over time or how one should measure it.

Among many assessment measures of language proficiency, vocabulary tests are recognized as reliable indicators of second language proficiency (L2 proficiency, hereafter), which could provide a multidimensional grasp of language knowledge and use (Coady & Huckin, 1997). The Vocabulary Levels Test (VLT), for example, is a receptive test designed by Nation (2001), which has five sections representing five frequency levels of word families: the 2,000 (2k) word level, 3,000 (3k) word level, Academic Word level (AWL), 5,000 (5k) word level, and 10,000 (10k) word level.[1] The VLT has a few productive versions; for example, the Lex30 (Meara & Fitzpatrick, 2000) is an association test assessing how many words are accessed or retrieved by test-takers at four frequency levels. In light of Levelt's (1989) observation that proper lexical access and its speed are highly dependent on the occurrence frequency of words, the words produced in the Lex30 can be taken as representing test-takers' high-frequency productive vocabulary. As for the assessment of depth, there are a word association test, a collocation test, and a test of self-report and performance. Lewis (2000), for instance, notes that the collocation tests used in Cambridge EFL Examinations offer a good measure of English proficiency, ascribing its effectiveness to a range of collocations appropriately selected for each level.

Although many vocabulary tests have been developed, there has been, to the best of my knowledge, no study which kept track of how university students' knowledge changes

in breadth, depth, and use over time. Thus, in an attempt to get a real picture of it, this paper aims at a semi-longitudinal investigation of Japanese university students' lexical knowledge. Specifically, this paper addresses the following questions:

- How many words do university students know in terms of reception and production (breadth and use problems)?
- How well do they know words (depth problem)?
- Do they retain or lose their lexical knowledge over time (change problem)?
- Do these three dimensions are related to each other (relation problem)?

Method

Participants

The participants were 146 freshmen and 129 sophomores at Kyoto University. The freshmen were from three faculties, whereas sophomores were from ten faculties (Table 1). The 275 participants were relatively advanced learners, whose mean score was 184.5 on the test administered by the National Center for University Entrance Examinations (Center Examination, hereafter), on which the overall mean score was 128.5.[2]

Table 1 Participants Cross-Tabulated between Faculties and Years

	Agriculture	Engineering	Science	Integated Human Studies	Economics	Law	Education	Letters	Medicine	Pharmaceutical Science	Total
Freshmen	0	80	0	0	30	0	0	36	0	0	146
Sophomores	12	31	11	10	17	24	2	8	9	5	129
Total	12	111	11	10	47	24	2	44	9	5	275

Materials

This study used four materials: the VLT, the Lex30, a collocation test adapted from the Cambridge CAE, and a questionnaire. The VLT was used for three reasons. First, it shares a vocabulary list with productive vocabulary tests, such as the Lex30. Second, it can provide a sufficient number of test items since it is available in two versions. Third, it seems the nearest thing available to a standard vocabulary test (Meara, 1996). The Lex30 is basically an association test, in which test-takers are asked to write down more than two words they associate with prompt words such as "dirty," "trade," and "close." The words elicited were entered into the Vocabprofile program to sort them into four frequency categories, and the number of word families beyond the 1k level was calculated for the Lex30 score. The original versions of the VLT and the Lex30 were remade into three tests, with each one shortened to a third of the original test (shortened versions, hereafter) because of the limited test time. The shortened versions of VLT were composed of four clusters with three test items at the 3k, the AWL, the 5k, and the 10k level, each version consisting of 48 test items. The *t*-test among the shortened versions showed no statistically significant difference, with Cronbach's alpha standing at .890 and .920. The shortened version of Lex30 consists of ten prompt words. On the basis of the

correlation coefficient between prompt words and the total, 30 prompt words in the shortened versions were modified so that they would be of approximately the same difficulty in eliciting advanced words. For the Lex30, the Cronbach's alpha in the two shortened versions was .728 and .810.

The collocation test of the Cambridge CAE was used because it is appropriate in providing test items at the same level of difficulty. The test was also modified to reduce the effects that students' reading ability may have on their test scores: the passages were divided into short sentences with blanks to be filled in. The questionnaire asked for information on ID number, faculty, grade, sex, age, overseas experience, self-estimated scores on the Center Examination.[3] It also asked for their autonomous learning time, self-assessed proficiency in English, motivation for learning English, and future need for English.[4] In the final session, it asked whether they were exposed to English during the summer vacation.[5]

Procedures

In this study, there were three test sessions: the first session (Test 1) at the beginning of the academic year in May; the second (Test 2) in July; and the last (Test 3) in October. Students took the Lex30, the collocation test, the VLT, and the questionnaire in that order. The total amount of time was approximately 30 minutes: five for the Lex30, five for the collocation test, ten for the VLT, and five for the questionnaire.

In each session, the three scores were analyzed in addition to the mean scores: Lexical Sophistication (LS) beyond 1k; LS beyond 2k; and a receptive vocabulary size. The LS is the percentage of "advanced" words in the text (Laufer & Nation, 1995, p.309). Laufer (1995) suggests that the percentage of the beyond-2000 words in learners' sample (i.e., the condensed profile "beyond 2000") represents lexical richness in free production. Since the score of Lex30 is calculated by adding words beyond the 1k level, the percentage of the words beyond 1k as well as that beyond 2k were calculated. The formula used in estimating the receptive vocabulary size was adapted from Laufer (1998).[6] This study avoided estimating productive vocabulary size, attempting instead to analyze the percentage of words found at each frequency level because of the open-ended character of the Lex30. In order to see how students' orthographic knowledge might change, spelling errors were also counted. In a difference analysis, the data from the students identified across the sessions were used.

Results

This section reports the results of the three tests. Test 1 focused on the cross-sectional analysis between freshmen and sophomores as well as between students and native speakers of English; Tests 2 and 3 focused on the semi-longitudinal analysis over the May, June, and October sessions.

In Test 1, freshmen were compared with sophomores in terms of mean scores on the three tests, LS, and the estimated receptive vocabulary size (Table 2). The comparison was based on the assumption that the difference in lexical knowledge would be affected by the additional year of study (Laufer, 1998). On the cross-sectional analysis between

freshmen and sophomores, sophomores averaged higher on the VLT (Table 2), though a *t*-test did not show a significant difference between them.

Table 2 Mean Values in Years and Total

Year		LEX30 score	LS beyond 1k	LS beyond 2k	collocation	VLT	Estimated receptive vocabulary
freshmen	Mean	12.06	42.36	19.75	5.67	32.92	5811.28
N=146	SD	4.11	10.80	9.23	1.81	5.87	1302.26
sophomores	Mean	11.65	42.29	20.00	5.67	33.71	5989.84
N=129	SD	4.51	12.09	9.50	1.71	5.77	1449.33
Total	Mean	11.87	42.33	19.86	5.67	33.29	5895.04
N=275	SD	4.30	11.40	9.34	1.76	5.83	1373.58

A comparison among students from different faculties, however, exhibited a different aspect (Table 3). Although the comparison did not reveal any significant difference, the mean VLT scores did show a declining trend as the school year advanced. The results on the collocation test showed a very low mean score. The correct answers accounted for 56.7%, with sophomores generally not performing better than did freshmen. The results on Lex30 and LS also exhibited a declining trend as the school year advanced. This is true of Engineering and Economics, but not Letters, in which sophomores averaged higher than did freshmen.

Table 3 Cross Tabulation of Mean Values between Freshmen and Sophomores in Each Faculty

Faculty		N	LEX30	LS beyond 1K	LS beyond 2k	collocation	VLT	Estimated receptive vocabulary
Engineering	Freshmen	79	12.16	43.23	21.05	5.57	32.57	5680.40
	sophomores	31	10.42	40.87	21.23	5.39	32.26	5595.06
Economics	Freshmen	30	10.90	39.07	16.72	5.60	32.93	5905.30
	sophomores	17	10.65	38.90	16.78	5.71	31.65	5692.96
Letter	Freshmen	36	12.92	43.56	19.69	6.00	33.83	6023.78
	sophomores	8	14.00	48.11	23.01	5.38	32.13	5374.05

The results on Test 1 were also compared with the results on the Lex30 and the VLT given to eight returnees and two native speakers of English (NS, hereafter) in order to set up a frame of reference.[7] The students averaged lower than did NSs at every level except the AWL of the VLT (Table 4). The *t*-tests showed a significant difference at 10k (t = -4.05, $p < .000$), total score (t = -2.40, $p < .017$), and estimated receptive vocabulary (t = -3.29, $p < .001$).

Table 4 Comparison between University Students and NSs in VLT

		3k	AWL	5k	10K	Total	Estimated receptive vocabulary
University students N=275	Mean	9.71	10.74	8.89	3.95	33.29	5895.04
	SD	2.00	1.40	2.14	2.33	5.83	1373.58
NS N=8	Mean	10.71	10.29	10.14	7.57	38.71	7640.69
	SD	1.98	2.14	2.91	2.51	8.32	1734.54

* A full mark at each level is 12.

Table 5 shows that the percentage of the words produced decreased as the frequency-level advances, though the percentage at the AWL remains low. Tested against these criteria, NSs and students showed a significant difference ($p < .001 \sim .02$)

Table 5 Mean Percentage of Produced Words at Different Frequency Levels

		1k	2k	AWL	Off-list
University students	Mean	57.69	22.47	3.86	16.00
N=275	SD	11.41	8.20	3.93	8.22
NS	Mean	46.13	23.11	3.82	26.94
N=10	SD	10.24	7.18	4.20	7.13

In the semi-longitudinal analysis (Table 6), a slight change was observed between Test 1 and Test 2: the mean score improved on the Lex30 and declined on the VLT and the collocation test. The comparison between Test 1 and Test 2 in the VLT showed a significant decline ($t = -2.98$, $p < .00$). In addition, Table 6 shows the decline in all the dimensions on Tests 1, 2, and 3. The percentage of attrition is 15% for the Lex30, 41% for the collocation, and 25% for the VLT. The LS beyond 1k decreased by 6% and LS beyond 2k by 5%. The receptive vocabulary levels on Tests 1 and 3 declined from 5,894 to 4,293 words, that is, an average loss of 1601 words brought about by a space of approximately five months. The *t*-tests showed that the difference between Test 1 and Test 3 were statistically significant in all the dimensions except LS beyond 2k ($p < .000 \sim .012$, *t*-test).

Table 6 Change in Mean Scores from Test 1 to Test 3

	VLT			Estimated receptive vocabulary		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
Freshmen	32.93	29.89	24.93	5818.60	5076.42	4272.40
Sophomores	33.66	31.61	25.43	5980.34	5296.72	4315.69
Total	33.27	30.56	25.18	5894.08	5161.70	4293.89
	Collocation			Lex30		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
Freshmen	5.70	3.95	3.37	12.08	12.39	9.92
Sophomores	5.67	4.25	3.31	11.70	12.36	10.47
Total	5.69	4.06	3.34	11.90	12.38	10.20
	LS beyond 1k			LS beyond 2k		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
Freshmen	42.46	41.03	39.01	19.80	19.57	18.67
Sophomores	42.50	43.23	41.60	19.98	19.89	19.23
Total	42.48	41.88	40.31	19.88	19.69	18.95

Table 7 shows a change in the percentage of correct answers at four levels. At the AWL, 5k, and 10k levels, there was a consistent decrease, whereas at the 3k level, there was a slight increase.

Table 7 Percentage of Correct Answers at Four Levels in VLT

	3k	AWL	5k	10k	Total
Test 1	80.57	89.16	74.10	32.72	69.14
Test 2	79.98	81.88	74.18	18.93	63.74
Test 3	83.45	70.38	48.58	7.40	52.45

Table 8 shows percentages of the words produced at the four levels of the Lex30. There was a decline at the 2k and the off-list level, and a rise at the 1k and the AWL level. This accounts for the result that in LS beyond 1k, there was a significant decline, whereas in LS beyond 2k, there was not. In brief, the increase at the AWL level offsets the decrease at the off-list level.

Table 8 Change in Word Production Percentage at Four Levels

	1k	2k	AWL	Off-list
Test 1	57.5	22.63	3.86	16.03
Test 2	58.2	22.33	3.88	15.59
Test 3	59.5	21.41	4.99	13.75

The number of spelling errors found in the Lex30 was also compared between Test 1 and Test 3. It averaged 1.02 on Test 1 and 1.37 on Test 3; the paired t-test exhibited a significant increase in the amount (t = -2.37, p = .019).

The questionnaire provided some data relevant to the change mentioned above. As shown in Table 9, the correlation between scores on the Center Examination and those on the three tests becomes weaker over time. Since scores on the Center Examination represent students' initial proficiency in this study, the result indicates that the advantage in initial proficiency diminishes over time.

Table 9 Change in Correlation between Scores on Center Examination and Three Tests

	Test 1	Test 2	Test 3
Lex30	0.34	0.32	0.07
VLT	0.45	0.38	0.28
Collocation	0.13	0.19	0.06

According to the questionnaire, students' confidence in English proficiency declined slightly, whereas their future need for English and their motivation to learn English remained unaltered (Figures 1, 2, and 3).

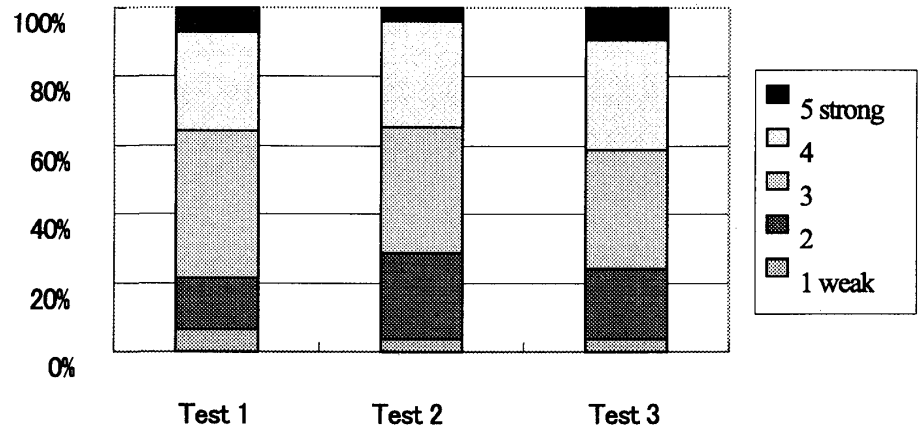

Figure 1 Change in Motivation for Learning English

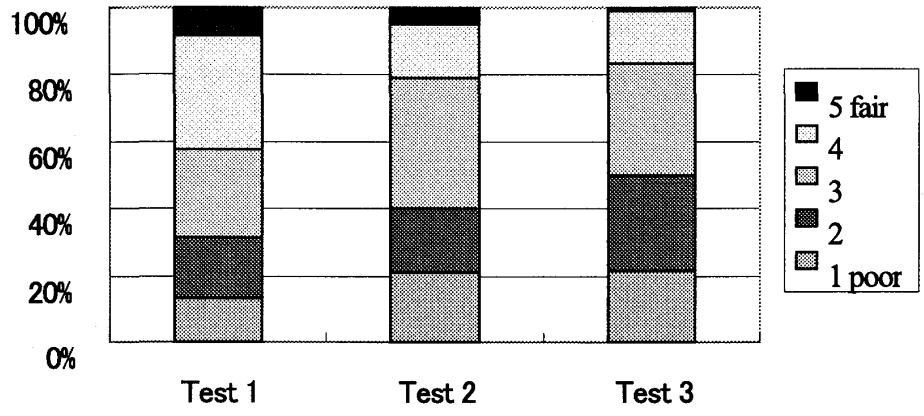

Figure 2 Change in Self-assessed Proficiency

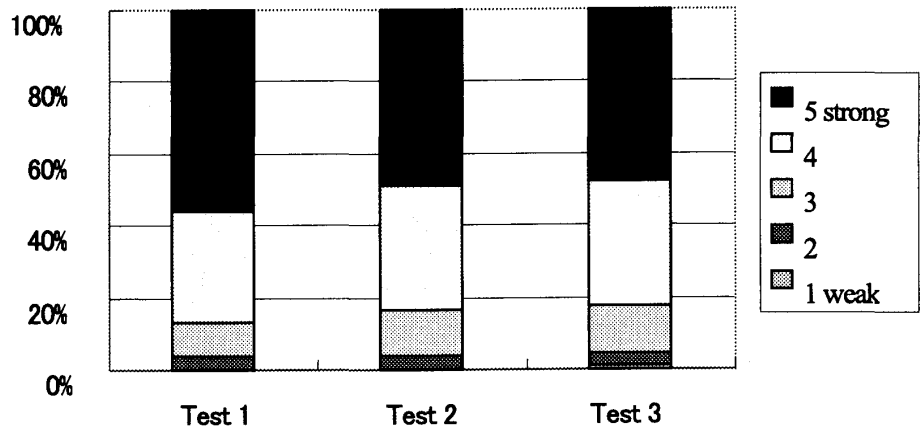

Figure 3 Change in Future Need for English

The mean autonomous learning time also showed a declining trend, with approximately 84 minutes in Tests 1 and 2, and 52 minutes in Test 3. The paired *t*-test showed that the differences in the amount of learning time between Test 2 and Test 3, and between Test 1 and Test 3 were statistically significant ($p < .005 \sim .000$).

Discussion

This section discusses the findings and attempts to answer the questions raised at the outset of this paper. It mentions in passing some other findings which were not presented in the previous sections.

Breadth and Use Problem

The results from Test 1 show that students have a receptive vocabulary of 5,894 word families. While it is smaller than the average NS's vocabulary of 7,640 words, it is larger than the average university student's vocabulary of 3,700 words (Shimamoto, 2000), 4,230 words (Shmitt & Meara, 1997), and 3,773 words (Nonaka, 2004). Interestingly, students performed slightly better than did NSs at the AWL level. In light of the assumption in Schmitt (2000) and Hirsh and Nation (1992) that 2,500-5,000 word families is large enough to provide initial access to authentic texts, students' receptive vocabulary seems sufficient in this regard. However, for the sake of advanced learning, many students would need a richer vocabulary, as some studies suggest. Laufer (1992), for instance, claims that with the knowledge of 6,000 word families, one can only attain 76% of reading comprehension. Similarly, Hazenberg and Hulstijn (1996) argue that in order to read materials such as university textbooks, one would need 10,000 word families. Given that many university students learn from academic materials, they need to expand their vocabulary and surpass the 5k and the AWL level.

Productive vocabulary is generally smaller than receptive vocabulary, and this was true of this study as well. Despite an accuracy rate of 69.1% on the VLT in which the target words went beyond the 2k level, the words beyond the 2k level students produced accounted for only 19.8% of the total production on the Test 1 (Tables 7 and 8). This indicates that receptively, they know many words beyond 2k, whereas productively, they

can retrieve very little of them. It is generally assumed that a productive vocabulary is consolidated largely through speaking and writing. However, the questionnaire shows that the students' ordinary forms of autonomous learning are reading (45.08%) and listening (34.28%), with hardly any opportunity to speak or write English. Having a poor productive vocabulary may be a natural result because presumably most Japanese universities do not provide students with an opportunity to speak or write English. The small percentage at the AWL level in the Lex30 could be due to a characteristic difference between the AWL and the other levels: the AWL contains 570 word families which constitute a specialized vocabulary most frequently used in academic texts (Coxhead, 2000), and retrieving academic words would require an academic context.

The Cross-sectional analysis reveals that sophomores seem to differ from freshmen in some respects, depending on which faculty they belong to. With Engineering and Economics, for instance, Lex30 and VLT scores showed a declining trend as their school year advanced. A comparison in Letters shows that sophomores averaged higher than did freshmen on the Lex30 and LS, which is natural in that their majors are closely related to languages. Sophomores' lower mean on the VLT might indicate that there was little additional learning of vocabulary over a year. This study has also made a comparison between students and NSs, showing that NSs performed much better than did students, a natural result. The difference was statistically significant in all the dimensions and was greater in the Lex30 than in the VLT. NSs produced 30.7% of the words beyond, whereas students did 19.8%, which is reflected in the LS beyond 2k. Interestingly, on both the Lex30 and the VLT, students averaged as high as did NSs at the AWL, possibly because students have been exposed to the AWL words very frequently.

Depth Problem

Compared with the results on the VLT, the results on the collocation test were very poor even at the beginning of the semester, exhibiting a sharp decline as time passed by: correct answers accounted for 56% on Test 1 and 33% on Test 3. These data show that students did not know many collocations. As pointed out by Decarrico (2001), collocational knowledge is more advanced knowledge left to higher-level learners who are enriching the vocabulary they have acquired. At the same time, however, collocational knowledge is necessary for using words in natural combinations and enhancing their fluency (Lewis, 1997). Meara (1993), underscoring the importance of the internal lexical network, points out that "the responses of L2 learners in a variety of simple psycholinguistic tasks would depend to a large extent on how closely their L2 lexicon continued to behave like an unstructured list of words and how far an internal structure had developed within it" (p.293). Seen in this light, the students in this study know the meanings of words but have yet to develop their collocational knowledge.

Change Problem

The semi-longitudinal assessment demonstrates that there was consistent attrition of lexical knowledge over five months in the three dimensions. Receptive vocabulary and collocational knowledge exhibited almost a linear decline. In the VLT, most of the receptive vocabulary at the 3k level was stable, whereas beyond the 3k level, it is not. In

other words, advanced words were the first to be forgotten. The declining pattern seems consistent with the general idea of language attrition, "last learned, first forgotten" (Freed, 1980).

A rise in the Lex30 in Test 2 seems to show that the retrievability of words is readily enhanced. Since a rise in the Lex30 was observed after the spring semester and was commensurable to a rise in the autonomous learning time, it will be justifiable to assume that the rise in Lex30 was due to an increased involvement in learning English. However, the sharp decline in the LS suggests that the rise in Lex30 could be ascribed mostly to that at the 1k level. The inference from this is that the rise in Test 2 does not mean that students expanded their productive vocabulary, but that they retrieved more of their existing vocabulary at the 1k level. Test 3 showed, however, that the retrievability declined during the two-month vacation.

In a typical Japanese environment of foreign language learning, lexical attrition seems to begin after the entrance examination. This is, as many would believe, most probably because students spend less time on English after the entrance examination, which is reflected in the reduced autonomous learning time. As claimed by de Bot and Kroll (2002), availability of words depends on recency and frequency of use (with many exceptions, though). The question that immediately arises is whether their lexical knowledge will remain stable as the core vocabulary that competent speakers of the language should be equipped with (Carter, 1987). Although further research is still needed, a prior pilot study of those students whose years ranged from the second to the fifth year makes it likely that students' lexical knowledge will continue to deteriorate.

The attrition found in this study was different in four respects compared with previous research. The first involves the effect of the initial proficiency on lexical attrition. Some studies on L2 attrition claim that one of the most important factors is the highest level of proficiency in L2 attained (Nagasawa, 1999; Russell, 1999). However, as shown in Table 9, the correlation between performance on each test and scores on the Center Examination, the latter being the initial proficiency, declined over time. This means that on Test 1, the students whose scores on the Center Examination had been higher performed better; on Test 3, however, this was not necessarily so. It shows that their advantage in initial proficiency was not retained. Moreover, given students' considerably high mean scores on the Center Examination, they are relatively advanced learners. Nevertheless, this study shows that they failed to maintain the level of knowledge they had attained.

The second factor involved in lexical attrition is students' motivation and attitudes. The loss of motivation causes language attrition by reducing the use of the language (Nagasawa, 1999). However, the result in this study seems to exhibit a more complex pattern. There were little change observed in motivation and future need, a gradual decline in perceived English proficiency, and a marked decline in learning time. Thus, it may be the case that while students might know that they are losing their English knowledge, they do not seem to do anything particular to attain it, despite the fact that they want and need to acquire high proficiency

The third is the attrition of orthographic knowledge. Reporting their investigations into

the attrition of receptive lexical knowledge, Welten and Grendel (1993) state that orthographic knowledge is not subject to attrition. However, this study, which examined spelling errors in the productive test, showed that there is a certain amount of attrition observable in orthographic knowledge as well. A closer look at errors such as "vertiaal" and "habitural" for "virtual" and "habitual" suggests that they may have poor knowledge of pronunciation. Their poor knowledge of orthography and pronunciation might cause them to hesitate to use them, which is likely to lead to further attrition.

The fourth is an attrition curve. Beginning early in the spring semester, this study observed sharp attrition in most of the dimensions. However, the kind of initial plateau which has been documented in many studies on L2 attrition was not observed in this study. Russell (1999) confirmed that there was an initial plateau by examining 80 native speakers of English who had stayed in Japan for two years learning Japanese. Meara (2004) developed computer models of lexical networks and examined how the networks might be affected by attrition. His theory of attrition is based on the disorganization of the lexical network with a gradual loss of its nodes. The patterns he presents are varied but do share a feature in that they all had a plateau before a rapid decline. Thus, the linear attrition of the receptive vocabulary presented in this study raises some additional questions. Given that freshmen stop intensive learning in March, should the four months before T2 represent a plateau? Given their poor collocational knowledge, is it plausible to say that the plateau was not there because of their frail lexical network? In order to understand why the attrition proceeds in linear fashion, further research will be needed.

The comparison between freshmen and sophomores was based on the assumption that the difference in lexical knowledge between the groups is attributable to the additional year of study. However, the results of this study raise a question about the assumption. This study observed no significant difference between freshmen and sophomores; however, it found remarkable attrition in both groups by keeping track of the individual students. Moreover, on Test 3, freshmen performed less well in all the dimensions than sophomores did on T1. If freshmen and sophomores form a developmental continuum, it should be expected that freshmen's proficiency would improve greatly by the following May, which did not seem to have occurred in this study. There are two possible explanations for this. One is that it was due to inadvertently uneven sampling of sophomores. On a cross-sectional analysis, the difference between the faculties should perhaps be taken into consideration, as may be seen in Table3; large and even sampling from each year in each faculty would be necessary. It is also possible that freshmen's average initial proficiency level is considerably different from year to year. Unfortunately, there was no survey conducted which, by using the same set of tests every year, examined freshmen's English proficiency at the time of their matriculation. These things indicate that in order to grasp the dynamics of lexical change, it will be essential to survey a larger sample of juniors and seniors, keeping track of the individuals over a longer period of time.

Relation Problem

The three lexical dimensions show a synchronized declining pattern, indicating that they are rooted in the same ground. Many researchers claim that there is a relationship

among breadth, depth, and use. Henriksen (1999), for instance, claims that depth plays a crucial role in understanding and producing words. Although this study only examined the collocation as the depth of knowledge, the low mean scores indicates that students' lexical network is not sufficiently formed. Furthermore, according to Henriksen (1999), a sparse network leads to little activation of words, causing sharp attrition in productive and receptive vocabulary.

Conclusion

This study has attempted to explore the real picture of university students' lexical knowledge of English by investigating the three dimensions of breadth, depth, and retrievability over a period of six months. In spite of the temporal limitations of this study, some implications may be drawn with respect to university education of English. First, it is necessary to maintain students' existing lexical knowledge and to facilitate additional learning. As pointed out by de Bot and Kroll (2002, p.144), "[t]hrough non-use of a language, the level of activation of knowledge in that language decreases, even to the point that that knowledge is considered lost." In order to develop an effective curriculum, it will be necessary to make a more in-depth analysis of students' vocabularies and develop an effective way to maintain the knowledge they have acquired. In order to retain new words for a long period of time, learners need to repeat them at increasingly longer intervals, according to Pimsleur's memory schedule (Nation, 2001). Since each word assumedly has its own schedule of repetition, the more words students have to retain, the more often they have to repeat them. By the same token, in order to retain the existing vocabulary of 6,000 words and expand it, students need to be repeatedly exposed to English over a long period of time.

Second, in this study there was no opportunity to examine a large number of juniors and seniors; however, the pilot study involving some juniors and seniors suggests that attrition is likely to proceed as the year advances. In many universities, the language program is intended primarily for freshmen and sophomores, but it might be worthwhile to consider the possibility and advantage of dividing the program so that juniors and seniors may have an opportunity to study English.

References

Carter, R. (1987). Is there a core vocabulary?: Some implications for language teaching. *Applied Linguistics,* 8, 2, 178–193.

Coady, J., & Huckin, T. (1997). Series editors' preface. In J. Coady & T. Huckin, (Eds.), *Second language vocabulary acquisition: a rationale for pedagogy.* New York: Cambridge University Press.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly,* 34, (2), 213–235.

de Bot, K., & Kroll, J. F. (2002). Psycholinguistics. In N. Schmitt, (Ed.), *An introduction to applied linguistics.* London: Arnold.

Decarrico, J. S. (2001). Vocabulary learning and teaching. In M. Celce-Murcia, (Ed.), *Teaching English as a second or foreign language.* Boston: Heinle & Heinle.

Freed, B. (1980). *The problem of language skill loss.* Paper presented at the Annual Meeting of the Modern Language Association, New York, December 1980, adapted from Weltens, B. (1987). The attrition of foreign-language skills: A literature review. *Applied Linguistics,* 8,1, 22–38.

Hazenburg, S., & Hulstijn, J. (1996). Defining a minimal receptive second-language vocabulary for non-native university students: An empirical investigation. *Applied Linguistics,* 17 (2), 145–163.

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Aquisition,* 21, 303–317.

Hirsh, D., & Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language,* 8, 2, 689–696.

Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. J. L. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics.* London: Macmillan Academic and Professional.

Laufer, B. (1995). Beyond 2000: A measure of productive lexicon in a second language. In L. Eubank, L. Selinker & M. Sharwood Smith (Eds.), *The current state of interlanguage: Studies in honor of William E. Rutherford.* Amsterdam: John Benjamins

Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics,* 19, 255–271.

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics,* 16, 307–322.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* Cambridge: MIT Press.

Lewis, M. (1997). *Implementing the lexical approach: Putting theory into practice.* Hove: Language Teaching Publications.

Lewis, M. (2000). *Teaching collocation: Further developments in the lexical approach.* Hove: Language Teaching Publications.

Meara, P. (1993). The bilingual lexicon and the teaching of vocabulary. In R. Schreuder & B. Weltens, (Eds.), *The Bilingual Lexicon.* Amsterdam: John Benjamines.

Meara, P. (1996). The dimensions of lexical competence In G. Brown, K. Malmkjaer & J. Williams (Eds.), *Performance and competence in second language acquisition.* Cambridge: Cambridge University Press.

Meara, P. (2004). Modelling vocabulary loss. *Applied Linguistics,* 25, 2, 137–155.

Meara, P., & Fitzpatrick, T. (2000). Lex 30: an improved method of assessing productive vocabulary in an L2. *System,* 28, 19–30.

Nagasawa, S. (1999). Learning and losing Japanese as a second language: A multiple case study of American university students. In L. Hansen (Ed.), *Second Language Attrition in Japanese Contexts.* New York: Oxford University Press.

Nation, I.S.P. (2001). *Learning Vocabulary in Another Language.* Cambridge: Cambridge University Press.

Nonaka, T. (2004). Nihonjin daigakusei no eigo goi size: English vocabulary size of Japanese university students. *Niigata Seiryou Daigaku Tankidaigakubu Kenkyu Houkoku,* 34, 25–34.

Russell, R. A. (1999). Lexical maintenance and attrition in Japanese as a second language. In L Hansen (Ed.), *Second language attrition in Japanese contexts.* New York: Oxford University Press.

Schmitt, N. (2000). *Vocabulary in language teaching.* Cambridge: Cambridge University Press.

Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework: word associations and verbal suffix. *Studies in Second Language Acquisition,* 19, 17–36.

Shimamoto, T. (2000). An analysis of receptive vocabulary knowledge: Depth versus breadth. *JABAET,* 4, 69–90.

Weltens, B., & Grendel, M. (1993). Attrition of vocabulary knowledge. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon.* Amsterdam:John Benjamins.

Footnotes

[1] A word family consists of a headword, its inflected forms, and its derived forms.

[2] The freshmen averaged 183.1 and the sophomores 185.84 on the Center Examination. The overall means was 126.8 in 2003 and 130.1 in 2004. Of the 275 students who answered the questionnaire, 250 volunteered to report their scores on the Center Examination with the knowledge that their scores would be made anonymous use of in the researcher's future work.

[3] In the questionnaire, students were asked to write down part of their mobile phone numbers, because this information was essential when identifying the individuals and seeing how the individuals' performance would change.

[4] In this study, "autonomous learning time" means a period of time students spend learning English spontaneously outside their English class.

[5] Seven students who had declared overseas experience were eliminated from the diachronic analysis because the *t*-test between the seven and the other students showed significant differences in VLT (t = -4.87, $p < .000$) and in Lex30 (t = -5.43, $p < .000$).

[6] Estimated receptive vocabulary = (VLT3k × 3 + (VLT3k + VLT5k) / 2 + VLTAWL + VLT5k + VLT10K × 5) × 10000 / 132.

[7] The eight returnees had lived in English-speaking countries for more than four years ranging from four years to 12 years, and returned to Japan about six months or a year before the test. They took the Lex30 and VLT which was used in Test 1. The two native speakers were teachers of English. They were recognized as a group in this study since their scores on these tests were not significantly different. They also took the Lex30 used in Test 1.