

RECEPTIVE AND PRODUCTIVE VOCABULARY ACQUISITION: EXAMINING PROCESSING TIME AND MEMORY THRESHOLD

Andrew Oberg

Language Education Center, Surugadai University, Hanno City, Japan

email: a.r.oberg@surugadai.ac.jp

Abstract: In the present research a picture card-based vocabulary study method was used to determine the pace of learner acquisition in terms of both receptive and productive knowledge. Fifty-eight first-year Japanese university students in two classes were used in the study. The subjects were placed into a single experimental group and both classes were taught by the researcher. During the treatment sessions the subjects studied fifteen vocabulary items chosen from the course textbook (*Fifty-Fifty: A speaking and listening course, Book One*, W. Wilson and R. Barnard, 2007, Hong Kong: Pearson Longman) using self-produced picture cards. The same post-treatment test was given to the subjects following each treatment session and the results were analyzed using Paired Samples T-tests supported by non-parametric Wilcoxon matched pairs signed ranks tests by comparing one session's data with the data of the previous session. The post-treatment data showed a significant difference between the pre-treatment and first sessions in terms of receptive knowledge, and between all sessions except the eighth and ninth in terms of productive knowledge.

Keywords: Vocabulary learning; receptive knowledge; productive knowledge; processing time; memory threshold

Abstrak: Dalam penelitian ini, teknik pembelajaran kosakata dengan menggunakan kartu bergambar digunakan untuk menentukan kecepatan pemerolehan bahasa pembelajar baik dalam pengetahuan reseptif maupun produktif. Subjek penelitian ini terdiri atas lima puluh delapan mahasiswa universitas Jepang tingkat satu yang terbagi ke dalam dua kelas. Mereka ditempatkan dalam satu kelompok eksperimental, dan kedua kelas diajar oleh peneliti sendiri. Selama tahap eksperimen berlangsung, para subjek mempelajari lima belas item kosakata yang dipilih dari sebuah buku teks mata kuliah (*Fifty-Fifty: A speaking and listening course, Book One*, W. Wilson and R. Barnard, 2007, Hong Kong: Pearson Longman) dengan menggunakan kartu bergambar yang mereka buat sendiri. Pos tes setelah eksperimen berakhir sama diberikan pada subjek setiap selesai satu sesi dan hasilnya dianalisis dengan menggunakan Uji-t sampel berpasangan yang didukung oleh Uji Pemeringkatan Tanda Padanan-Pasangan Wilcoxon nonparametrik dengan membandingkan data satu sesi dengan data dari sesi sebelumnya. Data dari pascaperlakuan menunjukkan perbedaan yang signifikan antara sesi praperlakuan dan sesi pertama dalam hal pengetahuan reseptif, dan antara semua sesi, kecuali sesi ke delapan dan ke sembilan, dalam hal pengetahuan produktif.

Kata Kunci: Pembelajaran kosa kata, pengetahuan reseptif, pengetahuan produktif, waktu pengolahan, ambang batas ingatan

The repetitious study of a pre-determined list of items has repeatedly been shown to have great advantages for long-term memory formation within both SLA research and other related fields (Bahrick, Bahrick, Bahrick & Bahrick, 1993; Cowan, 2000; Ellis & Beaton, 1993; Ellis, 2002; Ericsson & Kintsch, 1994; Gass & Mackey, 2002; Gupta & MacWhinney, 1997; Henriksen, 1999; Hulstijn, 2002; Knowles, 2008; Lewis, 1993; Mohensi-Far, 2008a, 2008b; Nakata, 2008; Nation, 2001, 2002; Papagno & Vallar, 1992; Schmitt, 2000; Segler, 2002; Tarone, 2002; Wei, 2007; Weil, 2008). It is, therefore, important to determine at what point, if a specific point can be ascertained at all, of study items are receptively (or passively) known, at what point they are productively (or actively) known, and at what point the items move from being receptively to productively known. If such required processing times and a memory threshold between the types of knowledge can be established and generalized, their application to teaching methodologies and classroom practice could contain many potential advantages for learners. Few studies have attempted to directly determine where these points may occur; however, most memory studies focus instead on the structure and limits of the working memory and/or the transition to long-term memory storage (Byrne & Bovair, 1997; Cowan, 2000; Ericsson & Kintsch, 1994; Gupta & MacWhinney, 1997; Just & Carpenter, 1992; Papagno & Vallar, 1992; Parks & Yonelinas, 2009; Petrides, Alivisatos, Meyer & Evans, 1993)—related but not identical issues with that of receptive/productive knowledge. Therefore, the current research attempts to address this concern by analyzing subjects' results on a repeated measures test that was designed to measure both receptive and productive knowledge of fifteen vocabulary items. By analyzing the results of each test compared with those of the previous treatment session's test a basis for determining these points was established. Cognitive psych-

ology findings regarding memory were taken into account here, as were ELT findings regarding types of knowledge, picture/word association, the phonological loop, and test design.

Working Memory, Processing, and Memory Thresholds

For a new piece of information to achieve storage in a person's long-term memory, it must first pass through the working memory, and therefore, a deeper understanding of the working memory may be of some interest. Research has shown that processing and storage in the working memory are limited by activation, and that the amount of activation available varies by individual (Just & Carpenter, 1992). Furthermore, such activation, or attention on the items currently held in the working memory, is thought to be a necessary aspect not only of later storage of the item in a person's long-term memory, but also in the efficient retrieval of said item (Schmidt, 1993). The available attention that can be given to an item appears to be capacity-limited (Cowan, 2000); however, and if such a limit does vary by individual as Just and Carpenter (1992) indicated, then the issue of creating activities that will work for all learners and interest them enough to maintain the level of focus required for further storage is an important one. Moreover, if an individual's capacity is constrained in some way, say through over-taxation, then processing is likely to also be constrained, as is the communication between processing resources. Such boundaries are the results of resource limitations, and not architectural structures (Just & Carpenter, 1992). But do architectural boundaries, as such, exist in the memory? Meara (1996) proposed that the movement from a receptive to a productive knowledge of a vocabulary item (or passive to active, in his terms) is likely to entail crossing a threshold rather than movement on a continuum. If such a threshold did exist, it would fit with Schank's model of memory

(1980) that proposed four layers that form a hierarchy of increasingly specific (and therefore, predictive) situational or informational templates. Other research (Parks & Yonelinas, 2009) has shown evidence of a bottom threshold, a point at which movement beyond leads to memory failure and further indicated that memory traces may remain in the mind for all items or stimuli encountered. If such findings hold out to be true, then the implications for foreign language learning could be vast, particularly if every new language item was retained in some way—even in a network that was purely receptive. Therefore, being able to ascertain just how much processing, on average, of an item is needed to move it from a receptive knowledge to a productive one, allowing for individual differences in capacity and applicable activation could be quite advantageous to an educational setting.

From Working Memory to Long-term Memory

Knowledge of the physical processes involved in long-term memory formation can also be beneficial to the language teacher, and recent findings from research done in experimental psychology have shed considerable light on this matter. When a piece of information is encountered and engaged, be it linguistic or otherwise, the brain first stores it in the working memory, which is thought to contain three primary components: "...a visuo-spatial short-term memory, a verbal short-term memory, and a central executive, which controls the flow of information to and from the other components" (Gupta & MacWhinney, 1997, p. 270). If the piece of information in question is completely new, it will fail to activate the central executive (a mass of differentiated tissue referred to as a 'chunk node'); this, in turn, will lead to the brain creating a new "chunk node" that releases an associated context signal. The connection between the context signal and new piece of

information is initially very fragile; however, remaining activated for only two to thirty seconds, after which time it is subject to decay if not reactivated by an additional stimulus or thought process (Cowan, 2000). One of the simpler ways to achieve this reactivation is through repetition of the material, a method that has enjoyed widespread and longstanding support from both within and without SLA literature (Bahrick, Bahrick, Bahrick & Bahrick, 1993; Cowan, 2000; Ellis & Beaton, 1993; Ellis, 2002; Ericsson & Kintsch, 1994; Gass & Mackey, 2002; Gupta & MacWhinney, 1997; Henriksen, 1999; Hulstijn, 2002; Knowles, 2008; Lewis, 1993; Mohensi-Far, 2008a, 2008b; Nakata, 2008; Nation, 2001, 2002; Papagno & Vallar, 1992; Schmitt, 2000; Segler, 2002; Tarone, 2002; Wei, 2007; Weil, 2008). In addition to strengthening the specific relationship between the new piece of information and its context signal, such repetition also helps to secure the context signal into the wider neural network (Henriksen, 1999). Once complete, this process leads to information being stored in the long-term memory. An intermediary state, termed the "long-term working memory," has also been suggested by Ericsson and Kintsch (1994, p. 3), who state that "information in LT-WM [long-term working memory] is stored in stable form, but reliable access to it may be maintained only temporarily by means of retrieval cues in ST-WM [short-term working memory]." Much more research in this area is needed, but the broader lesson that can be drawn in relation to foreign language learning is the need for multiple exposures and repetitious interactions with the target material.

'Knowing' a Word and the Receptive/Productive Distinction

Of primary importance in framing the argument that follows is the definition of what "knowledge" of a word entails. If there is a definable point at which a word can be

said to be receptively or passively known, then a necessary part of ascertaining that point will be establishing what we mean by saying that someone “knows” a given term. Schmitt (2000) has written that a word’s meaning is taken from its referent, but due to the lack of uniformity in our world and the need for a broader language than a simple system of referents would allow (if such were the case we would have little other than proper nouns), the meaning of a word can more accurately be said to be its relationship with its associated concept. To “know” a word would, therefore, be to understand the relationship between that word’s concept and its phonetic expression in sound or visual expression in symbols. In other words, it would mean having the information of this relationship (between the word and its concept) stored in one’s long-term memory in an associative way with its context signal, either hearing the word spoken or seeing it written (Henriksen, 1999). These points should be fairly clear and further discussion is necessary. Following from this, however, is the important distinction of receptive versus productive knowledge. Nation (2001) has stated this as being rooted in the common differentiation between listening and reading (receptive skills) and speaking and writing (productive ones). He noted that “passive” and “active” are sometimes used as synonyms for these terms (referencing Corson, 1995; Laufer, 1998, Meara, 1990). Nation (2001) also cited Meara as expressing the difference in terms of active vocabulary being stimulated by other words, but passive only by things external; they are, thus, both types of associative knowledge (Meara, 1990). Corson (1995), on the other hand, is said to base his distinction on use rather than knowledge, creating a boundary between active and passive vocabulary, where the latter includes active terms as well as words that are partly known, of low-frequency, not easily available for use, and whose use is avoided. The aforementioned definitions are lacking in

practical applicability in an L2 sense however, and so, in the present paper the following criteria have been applied (based on a stripped-down version of Nation’s (2001, pp. 26-28) “[S]cope of the receptive/productive distinction”): 1) Receptive knowledge entails knowing the word’s L1 equivalent, being able to recognize the word when heard and/or seen and/or written, and knowing the word’s associated concept; 2) Productive knowledge entails being able to use the word to express its meaning, being able to say and/or write the word, and being able to use the word correctly in an original sentence.

Image Association and the Phonological Loop

The use of imagery as a means of providing a deeper mental processing via meaningful association to strengthen learning (Mohseni-Far, 2008a; Papagno & Vallar, 1992; Schmitt, 2000) is an important one and is borne out in much research. In a review of a study comparing two types of vocabulary item annotations—verbal only and verbal with visual information—the annotations with a visual element were found more helpful than those with only a verbal element (Son, 2001). Other research has stated that “vocabulary acquisition is enhanced if the verbal information is accompanied by pictorial information” (Nikolova, 2002, p. 103), which seems natural given the working model of memory described by Gupta and MacWhinney (1997, see above). Moreover, other data suggest that students move from a formal to a semantic knowledge of L2 words (McNeill, 1996), and as such, classroom exercises and other activities might be more helpful if they focused more fully on meaning, an element that image-based representations can assist in providing (Nation, 2002).

Another area of primary concern to vocabulary acquisition in language learning is known as the Phonological or Articulatory Loop. Research done in experimental

psychology has shown that, “overt repetition—i.e. re-cycling material through the phonological loop component of short-term memory lead[s] to . . . better long-term representations” (Ellis & Beaton, 1993, p. 553; see also Papagno & Vallar, 1992), particularly when said aloud (Ellis & Beaton, 1993). Such verbal repetition of an item has been repeatedly shown elsewhere to have positive results with both acquisition and retention (Baddeley, 1997; Cowan, 2000; Ellis, 1995; Ellis & Beaton, 1993; Gupta & MacWhinney, 1997; Hulstijn, 2001; Segler, 2002). Papagno and Vallar (1992), however, cautioned that phonologically similar words will tend to interfere with one another and have detrimental effects on the learning of both.

METHOD

The testing instruments to be used will need to be matched to the information sought (Nation, 2001; Schmitt, 2000). Towards this end, Nation (2001, p. 372; the following are the researcher’s paraphrases) has listed four questions to consider: 1) Is the knowledge needed to answer the test questions similar to that the teacher wants to test?; 2) Is it easy to make enough questions for all the vocabulary that the teacher wants to test?; 3) Will the questions be easy to mark?; and 4) Will answering the question provide repetition of the vocabulary and possibly extend learners’ knowledge?” (Schmitt, 2000, p. 173; the following are also the researcher’s paraphrases); furthermore, he has detailed three dimensions to assist in deciding on a format: 1) Discrete (independent knowledge or use) versus Embedded (part of a larger assessment); 2) Selective (specific items are the focus) versus Comprehensive (takes account of the whole content of an answer); and 3) Context-independent (the subject can produce the answer without referring to context) versus Context-dependent (looks at

the subject’s ability to take account of context to answer the question). Based on the above and the criteria stated regarding receptive and productive knowledge, the test used in the current research (see Appendix A) was designed to be of the discrete, comprehensive, and context-type (or the passive and free active types, according to Nation’s terminology (2001)). There has been some argumentation that questions of the controlled productive type are preferable (Meara & Fitzpatrick, 2000; Schmitt, 2000); however, the researcher’s purpose was to see if the subjects could actively produce samples that demonstrated knowledge of the items’ meaning in an unaided and unprompted way, and hence, the free sentence composition element was included. Based on the research in the field reviewed above, the following research question was formed: 1) How much processing of material is required for acquisition of vocabulary items on a receptive level, and how much is required for acquisition on a productive level? and 2) Is there any basis for a threshold of processing time after which the majority of subjects move from a receptive knowledge to a productive knowledge of the material?

Based on the results of previous research, acquisition of the material using the method employed was assumed to occur (Oberg, 2011); however, no prediction was made regarding at what point either receptive or productive knowledge of the material would emerge, or if a basis for a threshold could be found, and therefore, 2-tailed T-tests are used in the analysis of the learners’ test data (Woods, Fletcher & Hughes, 1986).

Participants

Two classes comprising fifty-eight first-year university students were used in the study. The students were divided into classes based on their major of study and separated further by level based on their scores on the G-TELP test (General Tests of English

Language Proficiency) taken by all first-year students upon entrance to the university. One class consisted of twenty-seven Psychology majors, and the other class consisted of thirty-one Economics majors. Both groups were streamed at Level 2 (the second highest) for English ability within their respective majors. These two classes comprised the single group used in the study.

Variables

Care was taken to limit the variables in the present study. Previous research showed the card-based study method used here to be effective for acquisition (Oberg, 2011), and therefore, a control group was deemed unnecessary in this case. This was due to the focus of the present study being on the point at which receptive and/or productive knowledge may emerge, rather than on whether the method used would yield acquisition of the material. Furthermore, all of the participants involved in the research studied the same fifteen vocabulary items for ten minutes per treatment session. In non-research settings such as the one described below (see Procedure), varying degrees of exposure to the material can occur, and thus influencing the rate at which material is acquired; however, in the present study subjects were instructed not to review the vocabulary items outside of the treatment sessions, and hence, any extra exposure would have been incidental and unlikely to affect the outcomes of the statistical analyses applied. The same vocabulary test (see Appendix A) was used for data collection in each session; thus, making any overall improvements shown between sessions indicative of acquisition of the material. It could be argued that such improvements were merely results of the so-called practice effect and not actual acquisition, but in the present study acquisition is sought precisely through repetitious use of the target items, and the two-fold nature of the test (a matching section showing receptive gains and a

sentence writing section showing productive gains) was designed to reflect this. Furthermore, supporting the absence of any practice effect, the matching section's scores were not uniformly perfect, or entirely consistent, even at the very end of the study, which is where one would expect to see a practice effect if it were indeed taking place. The same criteria were used when marking each test and all of the marking was done by the researcher. Each test question was worth one point and either fully correct or fully incorrect. In the sentence writing section of the test, where judgments would have to be made as to what determines "correct," the following two sub-criteria were used: 1) Is it demonstrative of meaning? (Does the sentence indicate that the meaning of the item is understood? e.g. "I like my toolbox" would be incorrect, as there is nothing describing what exactly a "toolbox" is, whereas, "I took my hammer from my toolbox" would be correct.), and 2) Is it accurate? (Is the sentence reasonably grammatically correct?) All aspects of grammar were included in this second sub-criterion with the exception of correct article use; some leniency was also given for prepositions (e.g. "I played soccer on my backyard" was considered to have met this criterion despite the incorrect "on"). Correct spelling of words other than those being studied, however, was not a criterion as students were not allowed to use dictionaries during the tests and had the further pressure of a time limit (fifteen minutes) in completing the test. The justifications for the above controls are as follows: Although the researcher circulated the room during each treatment session helping students construct original sentences using the target items, it was naturally not possible for the researcher to be everywhere at once, and therefore, instructions were given for improving sentence writing to each class as a whole at the beginning of each treatment session, as otherwise students would have limited opportunities for acquiring the material productively. The time limit on each test

reduced the possibility of subjects arriving at a correct answer merely by chance, as the pressure to finish the test required the subjects to move quickly through the sections—either they knew the material or they did not. Finally, the scores on each test were not counted towards the subjects' class grades, meaning that the subjects would be less likely to study the material outside of the treatment sessions and also less likely to cheat on the tests themselves, improvements shown could, therefore, be assumed to indicate acquisition of the material through the methods employed.

Procedure

The present study was conducted over the course of one semester, comprising of fourteen ninety-minute class periods during which the pre-treatment session and ten treatment sessions took place. In the pre-treatment session the subjects were given a surprise vocabulary pre-test of the items to be studied in order to ascertain their initial level of knowledge. This same test was, then, used for all subsequent data collection (see Appendix A). Following the surprise pre-test, a list of the vocabulary items (see Appendix B) was distributed to each subject and the L1 equivalent for each item was verbally provided, and subjects were encouraged to take notes on the L1 meaning of each item but were not required to do so. All of the items used in the study were taken from the course textbook (*Fifty-Fifty: A speaking and listening course, Book One*, W. Wilson and R. Barnard, 2007, Hong Kong: Pearson Longman). The subjects were, then, provided with a blank sheet of paper and instructed to divide it into sixteen equal squares and draw whatever pictorial representation of each vocabulary item they thought best fit the item in question, with one square used per item. The subjects were not allowed to write either the L2 target word or the L1 meaning on the cards, only the picture they chose to draw. The squares were then cut into cards which

became each subject's card set. In the first through the tenth treatment sessions, the following pattern was employed: 1) The subjects made groups of three or four and used all of the group members' cards to play either the "Memory" or "Go Fish" game(s) for ten minutes. 2) While playing those games, the subjects were required to verbally state the vocabulary item represented by each card, and in the case that a pair was made, they were further required to verbally state an original sentence using the vocabulary item in question. Subjects were allowed to use their word lists and/or dictionaries during the playing of the game(s). 3) Immediately following the game time, the same vocabulary test was given to each subject, and all subjects were allowed fifteen minutes to complete the test. Subjects were not allowed to use their word lists or dictionaries while taking the test. During the time spent playing the games, the researcher circled the room to encourage subjects to state the items on the cards and help with sentence creation as needed. Additionally, prior to the first through tenth sessions, a brief reminder by the researcher on the need to make sentences that demonstrate meaning was given along with some examples. No model sentences were provided by the researcher in any format other than verbal. Subjects were not allowed to keep any of the vocabulary tests administered. The test results were then analyzed using Paired Samples T-tests (for repeated measures) supported by non-parametric Wilcoxon matched pairs signed ranks tests. For each of the analyses administered, the previous tests' results were used as one "group" and compared with the current tests' results as the other "group" (also comprising the pairs used in the Wilcoxon matched pairs analysis). This was done to discover the point at which, if any, a significant difference emerged both for receptive knowledge (as measured by the matching portion of the test) and productive

knowledge (as measured by the sentence writing portion of the test), and further, whether a basis could be found for a threshold at which most subjects passed from a receptive knowledge to a productive one. Appendix C contains a chart indicating how this analysis was organized (note that ‘Test 0’ is the surprise pre-test); scores were removed for subjects that were absent on either test day in the pairing prior to any analysis.

RESULTS AND DISCUSSION

The mean score on the receptive pre-test was 74.58%, and following the first treatment session the mean scores stayed

very close to one-hundred percent for the duration of the study. The receptive scores for tests 1-10 were as follows: 96.42%, 97.70%, 99.49%, 98.89%, 99.23%, 99.29%, 99.76%, 99.15%, 98.59%, and 99.36%. The mean scores on the productive tests showed much more variation; however. The mean score on the pre-test was 4.63%, with the productive scores for the treatment sessions (tests 1-10) as follows: 11.85%, 20.69%, 27.05%, 34.69%, 39.62%, 43.81%, 51.31%, 56.61%, 57.05%, and 66.54%. As a visual inspection of this distribution shows, the mean scores continued to rise after each treatment session, and the subjects demonstrated a large improvement overall in their ability to productively use the material tested (see Table 1, below).

Table 1: Receptive and productive test means (%)

Test	Receptive Mean (%)	Productive Mean (%)
0 (pre-test)	74.58	4.63
1	96.42	11.85
2	97.70	20.69
3	99.49	27.05
4	98.89	34.69
5	99.23	39.62
6	99.29	43.81
7	99.76	51.31
8	99.15	56.61
9	98.59	57.05
10	99.36	66.54

Paired Samples T-tests (for repeated measures) were performed on the data gathered from the ten treatment sessions. The data were not normally distributed in all cases however, and so additional non-parametric Wilcoxon matched pairs signed ranks tests were also performed to corroborate the 2-tailed significance scores from the T-tests. These analyses were applied to determine how much processing was needed for receptive and productive acquisition of the material, and at which point—that is, after how many treatment

sessions—a significant difference may emerge indicating a sudden and noticeable gain in productive acquisition of the material, i.e. a memory threshold being crossed from a receptive knowledge to a productive one. However, the results of both analytical tests used did not support the existence of such a threshold (see Interpretation, below). The alpha levels set for each test were .05, and all of the Paired Samples T-test scores were supported by the Wilcoxon matched pairs signed ranks test scores: where the T-tests showed a significant difference, the

Wilcoxon tests did as well, and vice versa (see Tables 2 and 3, below).

On the receptive test data, the T-test 2-tailed significance scores (p) and Wilcoxon matched pairs scores (W) were as follows: For the pre-test and test 1 $p=.000$ with $W=.000$; for tests 1 and 2 $p=.207$ with $W=.181$; for tests 2 and 3 $p=.105$ with $W=.098$; for tests 3 and 4 $p=1.000$ with $W=1.000$; for tests 4 and 5 $p=1.000$ with $W=1.000$; for tests 5 and 6 $p=.322$ with $W=.276$; for tests 6 and 7 $p=.289$ with $W=.285$; for tests 7 and 8 $p=.200$ with $W=.163$; for tests 8 and 9 $p=.533$ with $W=.458$; and for tests 9 and 10 $p=.160$ with $W=.102$. As can be seen from these results, a significant difference only emerged after the first treatment session (with all other scores above the alpha level), the questionable meaning of which is discussed more fully below (see Interpretation).

The productive test data were equally uniform with the receptive but in the opposite direction, with all pairs, save one, (tests 8 and 9) showing a significant difference. The productive test T-test 2-tailed significance and Wilcoxon matched pairs scores were as follows: For the pre-test and test 1 $p=.000$ with $W=.000$; for tests 1 and 2 $p=.000$ with $W=.000$; for tests 2 and 3 $p=.011$ with $W=.008$; for tests 3 and 4 $p=.000$ with $W=.000$; for tests 4 and 5 $p=.025$ with $W=.037$; for tests 5 and 6 $p=.015$ with $W=.011$; for tests 6 and 7 $p=.000$ with $W=.001$; for tests 7 and 8 $p=.000$ with $W=.001$; for tests 8 and 9 $p=.541$ with $W=.360$; and for tests 9 and 10 $p=.000$ with $W=.000$. It should be noted that in the cases where either p or W equaled .000 there was an associated value of some kind, however miniscule, but the statistical software used in the analysis only generated a score out to three decimal points.

In response to the original research question and sub-question, some potentially interesting results were found regarding both receptive and productive acquisition of the

material, as well as the presence or absence of a possible memory threshold, discussed separately below.

Regarding the amount of processing needed to receptively acquire knowledge of the material studied, it was found that a significant difference emerged between the pre-test and the first treatment session. On the face of it, it would appear from this result that only a single study session was necessary for the subjects to gain a receptive knowledge of the items, but the initially high scores on the pre-test (mean of 74.58%) seem to indicate otherwise. Although the subjects' mean scores moved from 74.58% to 96.42%—a considerable improvement—after the first session, the researcher cannot help but feel that this improvement was due to the subjects brushing up on pre-existing vocabulary knowledge rather than acquiring new knowledge. That the mean scores thereafter remained very close to 100% does appear to indicate solid acquisition of the material, but it is the researcher's judgment that stating from this that only one treatment session is sufficient processing time for acquisition to occur would be inappropriate.

As for the amount of processing required for productive knowledge of the material to be acquired, the results found here are slightly more problematic. Although the subjects' mean scores demonstrated a steady increase (starting at a pre-test mean of 4.63% and then going from 11.85% after the first treatment session to 66.54% after the tenth), the consistent occurrence of significant differences in the statistical analysis of the data makes it difficult to pinpoint just where a cut-off line could be placed. If one were to suggest the point at which the mean exceeded 50%, then that point would be after the seventh session (51.31%), but such a distinction seems arbitrary and is not supported by either the Paired Samples T-test or Wilcoxon matched pairs results, since each grouping of test data revealed a significant difference except for

the eighth and ninth. The ninth session itself presents its own set of problems, as the mean score increased by the smallest amount after that session (only 0.44% above the eighth session's mean), and as mentioned, there was no significant

difference found between the eighth and ninth sessions alone in all of the treatment sessions' data.

Table 2: Receptive test Paired Samples T-test 2-tailed significance and Wilcoxon matched pairs signed ranks test scores

Test	Paired Samples T-test 2-tailed significance scores (p)	Wilcoxon matched pairs signed ranks test scores (W)
0 (pre-test) and 1	.000	.000
1 and 2	.207	.181
2 and 3	.105	.098
3 and 4	1.000	1.000
4 and 5	1.000	1.000
5 and 6	.322	.276
6 and 7	.289	.285
7 and 8	.200	.163
8 and 9	.533	.458
9 and 10	.160	.102

The reasons for this anomaly may have been environmental. As the researcher recalls, it was extremely hot and humid that day and the university at which the present study took place had been refraining from using the air conditioning in an effort to save power, the resulting discomfort may have affected subjects' motivation and caused them to be less willing to put effort into their sentence writing. Such reasoning is merely intuitive, however, and not supported by any data. Nevertheless, the overall trend of improvement and the consistent emergence of significant differences indicate that, again, no set processing time can be said to definitively be required for productive acquisition to take place.

Finally, regarding the research sub-question, whether or not there is a basis for a threshold at which the majority of subjects move from a receptive knowledge to a productive one, the results taken together point to the absence of such a threshold. Although the subjects quickly

demonstrated strong receptive knowledge of the material, both the steady improvement in productive test scores and repeated statistically significant differences between tests point to productive knowledge of the material being acquired on a continuum rather than crossing a definable threshold. Moreover, the lack of any sudden and noticeable jump in the mean scores themselves appears to support this.

The findings of the present study would benefit from similar research done in the future, particularly if done over a longer period of time, with a wider variety of subjects, and with a set of vocabulary items that were completely unknown at the outset of the study. A data collection device that included more than one type of testing mechanism each for receptive and productive knowledge may also help strengthen the results reported here. Finally, a study that could ensure the attendance of each subject for each treatment session would also likely prove to be advantageous.

Table 3: Productive test Paired Samples T-test 2-tailed significance and Wilcoxon matched pairs signed ranks test scores

Test	Paired Samples T-test 2-tailed significance scores (p)	Wilcoxon matched pairs signed ranks test scores (W)
0 (pre-test) and 1	.000	.000
1 and 2	.000	.000
2 and 3	.011	.008
3 and 4	.000	.000
4 and 5	.025	.037
5 and 6	.015	.011
6 and 7	.000	.001
7 and 8	.000	.001
8 and 9	.541	.360
9 and 10	.000	.000

The study had its limitations. A similar study done over a longer period of time may yield more robust data than that presented here. The number of ten vocabulary tests used to collect subjects' data (excluding the standard-setting pre-test ('Test 0')) was chosen arbitrarily to maximize the potential possible within the confines of a single semester, a limitation stemming from the quasi-experimental setting. Had it been possible to extend the study, treatment sessions organized on an optimal spacing basis may also have resulted in more reliable data; studies have shown such a schedule to be advantageous for learning and long-term retention of a variety of materials, including L2 vocabulary (Baddeley, 1997; Bahrck, Bahrck, Bahrck & Bahrck, 1993; Hulstijn, 2001; Mizuno, 1996, 2003). Moreover, as the mean receptive score of 74.58% on the pre-test ('Test 0') indicates, the subjects had an initially high receptive knowledge of the material, a factor that limited the usefulness of the analysis applied in response to the research question and sub-question. The final two limitations were also results of the non-laboratory setting and included some subjects forgetting their vocabulary card sets,

lists, or both, and the unexpected absences of some subjects. In the case of the former, subjects were instructed to share the sets and/or lists with the other members of their group, and in the latter, their scores were removed from the data before any analysis was done.

CONCLUSION

In the research reported here a significant difference was found between the pre-test and first treatment session test for receptive knowledge of the material, and between all paired tests, save the eighth and ninth sessions', for productive knowledge on a repeated measures study of vocabulary acquisition. The experimental method made use of recent findings on memory formation, distinctions in types of vocabulary knowledge, image association, the phonological loop, and aspects of test design in foreign language study. Although acquisition of the material was demonstrated for both receptive and productive knowledge, the analysis did not reveal specific processing times required for such gains or

any basis for a threshold between receptive and productive knowledge. It is, therefore, argued that these findings point to vocabulary acquisition occurring on a continuum and in the absence of a definable threshold.

REFERENCES

- Baddeley, A. D. (1997). *Human memory: theory and practice* (Revised ed.). East Sussex: Psychology Press.
- Bahrack, H. P., Bahrack, L. E., Bahrack, A. S., & Bahrack, P. E. (1993). Maintenance of foreign language vocabulary and the spacing effect. *Psychological Science, 4*(5), pp. 316-321.
- Byrne, M. D., & Bovair, S. (1997). A working memory model of a common procedural error. *Cognitive Science, 21*(1), pp. 31-61.
- Corson, D.J. (1995). *Using English words*. Dordrecht: Kluwer Academic Publishers.
- Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, pp. 87-114.
- Ellis, N., & Beaton, A. (1993). Factors affecting the learning of foreign language vocabulary: Imagery keyword mediators and phonological short-term memory. *The Quarterly Journal of Experimental Psychology Section A, 46*(3), pp. 533-558.
- Ellis, N. C. (1995). The psychology of foreign language vocabulary acquisition: implications for CALL. *Computer Assisted Language Learning, 8*(2), pp. 103-128.
- Ellis, N.C. (2002). Frequency effects in language processing. *Studies in Second Language Acquisition, 24*, pp. 143-188.
- Ericsson, K. A., & Kintsch, W. (1994). Long-term working memory. *Institute of Cognitive Science Technical Report #94-01, University of Colorado at Boulder, 2-56*.
- Gass, S. M., & Mackey, A. (2002). Frequency effects and second language acquisition. *Studies in Second Language Acquisition, 24*, pp. 249-260.
- Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language, 59*, pp. 267-333.
- Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition, 21*, pp. 303-317.
- Hulstijn, J.H. (2001). Intentional and incidental second language vocabulary learning: a reappraisal of elaboration, rehearsal, and automaticity. In: P. Robinson (Ed.), *Cognition and second language instruction*, (pp. 258-286). Cambridge: Cambridge University Press.
- Hulstijn, J.H. (2002). What does the impact of frequency tell us about the language acquisition device? *Studies in Second Language Acquisition, 24*, pp. 269-273.
- Just, M.A., & Carpenter, P.A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), pp.122-149.
- Knowles, L. (2008). Mind blocks. *Language, 7*(12), pp. 28-34.
- Laufer, B. (1998). The development of passive and active vocabulary: same or different? *Applied Linguistics, 19*, pp. 255-271.
- Lewis, M. (1993). *The lexical approach: The state of ELT and a way forward*. London: Heinle, Cengage Learning.
- McNeill, A. (1996). Vocabulary knowledge profiles: evidence from Chinese-speaking ESL teachers. *Hong Kong Journal of Applied Linguistics, 1*, pp. 39-63.
- Meara, P. (1990). A note on passive vocabulary. *Second Language Research, 6*, pp. 150-154.

- Meara, P. (1996). The vocabulary knowledge framework. *The Lognostics Virtual Library*, pp. 1-11. Retrieved July 3, 2009, from <http://www.lognostics.co.uk/vlibrary/meara1996c.pdf>
- Meara, P., & Fitzpatrick, T. (2000). Lex30: an improved method of assessing productive vocabulary in an L2. *System*, 28(1), pp. 19-30.
- Mizuno, R. (1996). *Bunsankouka: Tenboto aratana setsuno teian* [Theories of spacing effects: Review and proposal of a new theory]. *Bulletin of the Shizuoka Institute of Science and Technology*, 5, pp. 181-197.
- Mizuno, R. (2003). *Gakusyukoukano ninchishinrigaku* [The cognitive psychology of learning effects]. Kyoto: Nakanishiya Publishing.
- Mohseni-Far, M. (2008a). In search of the best technique for vocabulary acquisition. *Estonian Papers in Applied Linguistics*, 4, pp. 121-138.
- Mohseni-Far, M. (2008b). A cognitively-oriented encapsulation of strategies utilized for lexical development: In search of a flexible and highly interactive curriculum. *Porta Linguarum*, 9, pp. 35-52.
- Nakata, T. (2008). English vocabulary learning with word lists, word cards, and computers: implications from cognitive psychology research for optimal spaced learning. *ReCALL*, 20(1), pp. 3-20.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, P. (2002). Teaching vocabulary. *Asian EFL Journal*, 7(5), pp. 1-9. Retrieved May 21, 2009, from http://www.asian-efl-journal.com/sept_05_pn.pdf.
- Nikolova, O. R. (2002). Effects of students' participation in authoring of multimedia materials on student acquisition of vocabulary. *Language Learning & Technology*, 6(1), pp. 100-122.
- Oberg, A. (2011). Comparison of the effectiveness of a CALL-based approach and a card-based approach for vocabulary acquisition and retention. *CALICO Journal*, 29(1), pp. 118-144.
- Papagno, C., & Vallar, G. (1992). Phonological short-term memory and the learning of novel words: The effect of phonological similarity and item length. *The Quarterly Journal of Experimental Psychology Section A*, 44(1), pp. 47-67.
- Parks, C. M., & Yonelinas, A. P. (2009). Evidence for a memory threshold in second-choice recognition memory responses. *PNAS*, 106(28), pp. 11515-11519.
- Petrides, M., Alivisatos, B., Meyer, E., & Evans, A.C. (1993). Functional activation of the human frontal cortex during the performance of verbal working memory tasks. *Proceedings of the National Academy of Sciences (USA)*, 90, pp. 878-882.
- Schank, R. C. (1980). Language and memory. *Cognitive Science*, 4, pp. 243-284.
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- Segler, T. M. (2002). Second language vocabulary acquisition and learning strategies in ICALL environments. *Computer Assisted Language Learning*, 15(4), pp. 409-422.
- Son, J.-B. (2001). CALL and vocabulary learning: A review. *English Linguistic Science*, 7, pp. 27-35.
- Tarone, E. (2002). Frequency effects, noticing, and creativity. *Studies in Second Language Acquisition*, 24, pp. 287-296.
- Wei, M. (2007). An examination of vocabulary learning of college-level learners of English in China. *Asian EFL Journal*, 9(2), pp. 93-114.

Weil, N. (2008). Vocabulary size, background characteristics, and reading skill of Korean intensive English students. *Asian EFL Journal*, 10(4), pp. 26-59.

Woods, A., Fletcher, P., & Hughes, A. (1986). *Statistics in language studies*. Cambridge University Press.

Appendix A: Vocabulary test of items studied

English II Semester 1: Vocabulary Test

Name: _____ Student Number: _____

A. Matching: Match the words with their Japanese meanings

A. ~分かかる	B. 引き出し	C. 富	D. 独身である	E. 主婦/主夫
F. 暖炉	G. 作業台	H. 工具箱	I. ちょうど	J. 裏庭
K. 一人で	L. 外食する	M. 十二宮図	N. 独立	O. 窓台

1. be single _____
2. zodiac sign _____
3. independence _____
4. takes ~ minutes _____
5. backyard _____
6. windowsill _____
7. fireplace _____
8. toolbox _____
9. homemaker _____
10. fortune _____
11. eat out _____
12. alone _____
13. drawer _____
14. on the dot _____
15. workbench _____

Name: _____ Student Number: _____

B. Sentence writing: Write one original sentence for each of the vocabulary items from the list.

be single zodiac sign independence takes ~ minutes backyard
windowsill fireplace toolbox homemaker fortune eat out
alone drawer on the dot workbench

1. _____
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____
9. _____
10. _____
11. _____
12. _____
13. _____
14. _____
15. _____

Appendix B: List of vocabulary items studied

English II Semester 1: Vocabulary List

1. be single
2. zodiac sign
3. independence
4. takes ~ minutes
5. backyard
6. windowsill
7. fireplace
8. toolbox
9. homemaker
10. fortune
11. eat out
12. alone
13. drawer
14. on the dot
15. workbench

Appendix C: Paired Samples T-test and Wilcoxon matched pairs signed ranks test data analysis organization chart

Data analysis sequence	Receptive results compared	Productive results compared
1 st analysis	Test 0 (pre-test) and Test 1	Test 0 (pre-test) and Test 1
2 nd analysis	Test 1 and Test 2	Test 1 and Test 2
3 rd analysis	Test 2 and Test 3	Test 2 and Test 3
4 th analysis	Test 3 and Test 4	Test 3 and Test 4
5 th analysis	Test 4 and Test 5	Test 4 and Test 5
6 th analysis	Test 5 and Test 6	Test 5 and Test 6
7 th analysis	Test 6 and Test 7	Test 6 and Test 7
8 th analysis	Test 7 and Test 8	Test 7 and Test 8
9 th analysis	Test 8 and Test 9	Test 8 and Test 9
10 th analysis	Test 9 and Test 10	Test 9 and Test 10

Schmidt HG. Foundation of problem-based learning: some explanatory notes. *Med Educ* 1993;27:422-32.