Comparison of the Effectiveness of a CALL-Based Approach and a Card-Based Approach to Vocabulary Acquisition and Retention

Andrew Oberg
Surugadai University

ABSTRACT
In this study, two study methods for the support of acquisition and retention of 10 vocabulary items were compared; one method used representative picture cards and the other a CALL interface. Seventy-one first-year Japanese university students comprising two classes participated in the study. The students studied a practice set of 10 vocabulary items using both of the two methods and then a treatment set of 10 different items using only one of the methods to which the students were randomly assigned. A t test done on the groups’ vocabulary pretest scores showed no significant difference between the two groups in terms of knowledge of the items at the outset of the experiment. The analysis of the posttreatment data showed no significant difference between the groups. Finally, a posttreatment survey revealed a slight preference for the CALL method among the students.

KEYWORDS
Vocabulary Learning, CALL, Picture Cards, Short-term Memory, Long-term Memory

INTRODUCTION
Comprehensive second language (L2) study has been taking place in human societies since at least the second century B.C.E. when Roman children were being instructed in Greek (Schmitt, 2000). Through the centuries, and through the various regional and panregional lingua franca that the world has seen, the acquisition of new vocabulary in an L2 has been a major focus of learners and teachers alike. Vocabulary knowledge is seen as so essential that it has been described as a prerequisite for successful communication (Nation, 2001). Communicative success is achieved more effectively by learners with a more expansive general vocabulary than by those with a smaller, but more controlled, vocabulary (Meara & Fitzpatrick, 2000). Despite the importance given to vocabulary and its central place in language learning, there have been relatively few studies directly comparing the use of word cards and CALL interfaces for vocabulary acquisition; studies tend to tout either the efficacy of word cards (Mondria & Mondria-de Vries, 1994; Nation, 2001; Schmitt & Schmitt, 1995; Waring, 2004) or that of computers (notably, in both cases, as compared to word lists) (Ellis, 1995; Hulstijn, 2001; Nation, 2001) rather than compare the two. What is needed, therefore, is a direct empirical comparison of these two methods of vocabulary study for effectiveness in vocabulary acquisition and retention. A recent study has made such a comparison (along with word lists; Nakata, 2008); however, that study failed to take into account certain insights into memory from experimental psychology (Cowan, 2000; Ellis & Beaton, 1993; Ericsson & Kintsch, 1994; Gupta & MacWhinney, 1997; Papagno & Vallar, 1992), advances made in vocabulary pedagogy through the lexical approach (Lewis, 1993; Nattinger & DeCarrico, 1992; Schmitt, 2000), and the use of imagery to aid in the forming of associations and processing.
depth (Mohseni-Far, 2008a; Nikolova, 2002; Papagno & Vallar, 1992; Schmitt, 2000). The present study takes into consideration the insights into the ways in which memory is structured and works, aspects of lexically based pedagogy, and the effective application of imagery in the acquisition and retention of vocabulary items.

PREVIOUS RESEARCH

Human Memory and Study Methods

When acquiring an L2, or any new information, each individual’s capacity for memory quickly comes to the fore as an important concern. How memory is structured in the mind and how information moves from working memory (WM) to long-term memory (LTM) are still not completely known, but recent research in cognitive psychology has shed some light on the matter. In one study (Gupta & MacWhinney, 1997), a subject, suffering from a ‘pure STM’ deficit in which reduced auditory-verbal short-term memory (STM) is affected independently of other major language or cognitive skills, was asked to associate an unknown L2 word with a known L1 meaning and an arbitrary, but known, L1 phonological form with a different, but known, L1 meaning. Both conditions required mapping a new form to a known meaning, but only the first required the learning of a new phonological form. The subject was unable to successfully map the L2 word onto the L1 meaning, but was able to associate the new L1 phonological form with the L1 meaning. The authors conclude that, “immediate serial recall and vocabulary acquisition may involve common processing mechanisms” (p. 272). The authors further suggest a WM model that has three main components, namely: “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” (p. 270). Thus, when a new word is encountered, it fails to activate the central executive (termed a ‘chunk node’ in the brain), and so a new ‘chunk node’ is activated and releases a context signal. Once activated, the connection between the specific context signal and the new word’s form is established, beginning at the phoneme level and then automatically parsed by the brain’s syllable template. This connection, however, remains fragile and subject to decay and therefore must be strengthened through multiple encounters with the word before it can enter a learner’s LTM; hence the wide and long-standing support for repetitious study in the research (Bahrick, Bahrick, Bahrick & Bahrick, 1993; Cowan, 2000; Ellis & Beaton, 1993; Ericsson & Kintsch, 1994; Gupta & MacWhinney, 1997; Lewis, 1993; Mohseni-Far, 2008a, 2008b; Nakata, 2008; Nation, 2001, 2005; Papagno & Vallar, 1992; Schmitt, 2000; Segler, 2002; Wei, 2007; Weil, 2008). Such repetition reinforces the application of the new meaning to the new form, strengthens the connection between the broader concept of the word and its textual/audio representation, and helps in further cementing the related context signal in the network that is being built in the brain (Henriksen, 1999). Knowles defines this neural network as “[the] specialized ‘cortical columns’ which are hypothesized to structure all modalities of sensory input” (2008, p. 30). The new relationship between the concept and the form becomes increasingly complex until full automatization has been achieved. The pedagogies that most effectively and efficiently support this process should therefore be explored in pedagogical approach to study methods to be employed by students.

The Phonological/Articulatory Loop

One such pedagogical approach is known as the phonological, or sometimes the articulatory, loop. Research done in experimental psychology has shown that, “overt repetition—i.e. recycling material through the phonological loop component of short-term memory lead to ... better long-term representations” (Ellis & Beaton, 1993, p. 553; see also Papagno & Vallar,
1992). In effect, the application of the phonological loop would simply mean using rote repetition—and it should be stressed here that this repetition has its best effects when students actually say the words aloud (Ellis & Beaton, 1993)—to promote the strengthening of the connection that is being developed between the newly created “chunk node’s” specific context signal and the vocabulary item’s form. Gupta and MacWhinney (1997) expanded on this idea to show that the same kind of rehearsal could be applied to multiple chunks of information with equal benefits. Gupta and MacWhinney caution, however, that there is a limit to the number of different chunks that can be learned this way, and that the “rehearsal rate ... depends not only on articulation rate, but on the number of chunks (items) ... to be recalled” (p. 276). Elsewhere, Cowan (2000) has argued that this limit is around four chunks, where ‘chunk’ is defined as, “a collection of concepts that have strong associations to one another and much weaker associations to other chunks concurrently in use” (p. 89).

There seems to be convincing reason to employ repetition in the classroom at least at some level when learning new vocabulary items. However, a potential negative side effect of repetition is student boredom, which can cause some students to simply fail to repeat the items with their peers in whatever form that repetition takes place. Nevertheless, the multiple exposures can lead to some form of passive learning, an area of vocabulary which other research has shown could potentially be exploited to a much greater extent than it has been (Laufer, 1998). Moreover, verbal repetition of an item has also been shown to have positive results in both acquisition and retention (Baddeley, 1997; Cowan, 2000; Ellis, 1995; Ellis & Beaton, 1993; Gupta & MacWhinney, 1997; Hulstijn, 2001; Segler, 2002). The application of rote repetition to stimulate the phonological loop and thus strengthen the nascent ‘chunk nodes,’ and subsequent ‘cortical columns,’ does indeed seem to be justified regardless of the concerns about boredom. A caution here, though, is given by Papagno and Vallar (1992), whose research suggests that phonologically similar words will tend to interfere with one another and thus have detrimental effects on the learning of both. This has been echoed by Nation (2001) and Schmitt (2000) regarding vocabulary items with similar or directly opposite meanings.

The Keyword Method and Other Uses of Imagery

Still another pedagogy that is often promoted in the study of vocabulary learning is termed the keyword method (or sometimes the keyword technique). This method is said to involve deep mental processing and thereby facilitate learning (Schmitt, 2000). To use the method, learners first find a word in their L1 that is phonologically similar to the word they wish to remember in the L2. Schmitt gives the example of trying to learn the Japanese word katana (sword) by matching it to the English cat and then picturing a Samurai cat waving a sword. When learners again hear the word katana, they are reminded of the image that they conjured for themselves which then leads them to the meaning of ‘sword’ (Nation, 2001; Schmitt, 2000). This, and other ways of creating a meaningful association (whether through imagery or semantic linking) are thought to aid in long-term retention (Mohseni-Far, 2008a).

There are problems with the keyword method though, especially as concerns the current study. For example, research has not shown that the technique can be applied equally well to lexical chunks as to single words, and even then the technique has been criticized as only applicable to concrete nouns and rarely to abstract ones (Segler, 2002). The direct use of the keyword method would therefore require ignoring the copious amount of research which underscores the advantages of learning items as lexical phrases (see the section on the lexical approach below). In addition, it may not always be possible for students to find a phonologically similar word in their L1 to match the L2 word they wish to learn. In the case of the imagery in Schmitt’s Samurai cat, katana and cat are not actually as phonologically similar as
they appear. The phonemes in *katana* cannot be split into separate consonants and vowels; the written word in Japanese would either appear as a single character of Chinese origin (刀) or as a series of hiragana (or possibly katakana) that cannot be broken down further (かたな or カタナ, respectively). Both of these versions of *katana* would appear romanized as ka-ta-na. The matching of *cat* can be applied to the first phonemes in *katana*, namely, *kata*, but the phonological similarity becomes somewhat vague after that. In addition to these concerns of adequate phonological matching are concerns of imageability and parts of speech. Ellis and Beaton (1993), in a study of 47 psychology major undergraduates learning German, found that "the effectiveness of the keyword method depends upon the part of speech and/or the imageability of the keyword and, further that part of speech, and/or imageability, of the foreign word to be learned influences recall performance" (p. 554). It is hard to imagine a student coming up with an appropriate keyword image for many purely grammatical items, and the issue regarding parts of speech echoes earlier misgivings about using the technique with lexical chunks. In summary, although the keyword method has been shown to be effective for learning individual words, it may not be universally applied with full confidence.

Although the keyword method has its limitations, the use of imagery as a means of providing meaningful association—and therefore deeper mental processing 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 plus visual information), the annotations with a visual element were found more helpful than those with a verbal element only (Son, 2001). Similarly, Yoshii and Flaitz (2002), in a study of 151 adult ESL learners, found that the group that studied vocabulary with a combination of text and picture annotations (as compared to groups that studied vocabulary with text only and picture only) consistently outperformed the other two groups in both immediate and delayed testing. 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). Moreover, other data suggest that students move from a formal to a semantic knowledge of L2 words (McNeill, 1996); as such, classroom exercises and other activities might be more helpful if they focused on meaning, an element that image-based representations can help in providing (Nation, 2005).

The Lexical Approach

Despite having first been widely publicized as far back as 1992, the lexical approach is still making inroads into many language classrooms. Put simply, this pedagogy involves learning vocabulary in groups of regularly co-occurring words. In the seminal work on the subject, Nattinger and DeCarrico (1992) define lexical phrases as being either strings of set lexical items that cannot be altered or general frames that allow some or much substitution/rearrangement of their parts. Within these two broad categories, however, are four subdivisions: polywords, institutionalized expressions, phrasal constraints, and sentence builders. In another book on the subject, Lewis (1993) defines lexical items as 'socially sanctioned independent units’ and goes on to state that "Many are words, but many consist of multi-word units” (p. 90). He also subcategorizes lexical items as consisting of: words, multiword items (also referred to as multiword units), polywords, collocations, and institutionalized expressions.

However lexical items are defined and categorized, research shows a very strong basis for the notion that bits of language often come together and that it is this togetherness that allows speakers to quickly organize their thoughts and express themselves with a high degree of automatization (Boers, 2000; Knowles, 2008; Lewis, 1993; Nation, 2001; Nattinger &
DeCarrico, 1992; Schmitt, 2000; Wible, Kuo, Chen, Tsao, & Hung, 2006). The main reason given for the automatization allowed by lexical items is that in whatever form they occur, they represent conventionalized and high-frequency forms of language. Since they are remembered wholesale by both speaker and hearer, they spare both sides a great deal of processing burden (Nattinger & DeCarrico, 1992). Moreover, because many lexical items actually contain slots that are subject to change (e.g., see you later/tomorrow/next Tuesday), they also have the advantage of fluency, an aspect that language learners can benefit from employing. In fact, Nattinger and DeCarrico (1992) point out that naturally occurring lexical phrase use is common among children learning both their L1 and an L2. Based on these advantages, both Lewis and Nattinger and DeCarrico stress the need for L2 learners to be equipped with prefabricated language that they can employ or substitute within a chunk’s allotted slots as a first stepping-stone towards fluency and communicative competence. In defending their position, Nattinger and DeCarrico state that “Many earlier researchers thought these prefabricated chunks were distinct and somewhat peripheral to the main body of language, but more recent research puts this formulaic speech at the very centre of language acquisition and sees it as basic to the creative rule-forming processes which follow” (p. xv). Supporting this view, Wible et al. found that, “users store and retrieve chunks as single multiword units rather than by rule-governed composition in real time” (p. 869). If these advantages of lowered processing burden and fluency prove to be valid, and current research indicates that they will, then teaching based on these ideas has many potential benefits.

The final advantage of approaching vocabulary learning from a lexical perspective involves aspects of how the human memory works and the receptive/productive axis of vocabulary knowledge. Within general neural functioning of memory some points related specifically to lexical chunking should be underscored. As mentioned above, one theorized aspect of memory is that there is a limit of about four separate chunks which can be held in the STM at any one time (Cowan, 2000). Items remain activated in STM for only about 2 to 30 seconds before decaying sets in, unless of course that item is re-stimulated by an additional presentation or thought. This is one reason why rehearsal is so important in vocabulary acquisition. It has also been similarly argued elsewhere that the WM allows for the short-term maintenance of sequenced information and that rehearsal of those sequences (linguistic or otherwise) promotes their cementing in the LTM (Ellis & Sinclair, 1996).

Again, the application of lexical chunking to vocabulary study seems to be a natural choice. Given the limit on item activation in WM and STM, however, it is important for learners to be able to process and integrate material in their LTM to fully secure the storage of that item. One proposed route for material to follow is that of long-term working memory (LT-WM; Ericsson & Kintsch, 1994). According to this model, information that has entered the LT-WM 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]” (p. 3; emphasis added). It is unclear what the exact difference is between STM and ST-WM, but it may be that information in ST-WM is currently being activated and therefore not heading towards the kind of decay normally found in the STM. However, LT-WM is not a replacement for ST-WM, but rather a way of supporting ST-WM to help deal with very specific demands on memory and is therefore relevant only to certain skilled activities (Ericsson & Kintsch, 1994). In essence, the proposed LT-WM would function as a further intermediary step reducing processing demands in real time. This model fits in well with Meara’s (1996) own ‘multistate model of vocabulary acquisition’ in which vocabulary knowledge is in a state of flux traveling between any of the various interrelated stages of knowledge. This is the basis for a continuum along the receptive/productive axis, although Meara holds that there is a clear threshold between when a vocabulary item is known receptively and when it is known productively. Vocabulary that is taught based on the lexical approach can therefore take advantage of WM’s ability to maintain sequenced information, the capacity to hold multiple chunks concurrently in ST-WM (or more briefly in the STM), sup-
ported by repetition through LT-WM and finally into the LTM, all the while reducing the need for processing of the item and assisting in the quick development of fluency and productive vocabulary knowledge.

**CALL and ICALL: Vocabulary and Other Applications**

Although not yet universal, personal computers clearly play a larger and larger role in the classroom and offer many advantages to both the language learner and teacher. Some of the potential advantages in vocabulary acquisition are the aforementioned uses of imagery to provide meaningful context (Mohseni-Far, 2008a; Papagno & Vallar, 1992; Schmitt, 2000) and the ease of mechanical repetition to aid in memory (Segler, 2002). Computers can also provide structured practice, training, and testing in ways that best contribute to vocabulary acquisition (Ellis, 1995). Studies done on spaced learning in which items are rehearsed over a longer period of time at set intervals have been shown to be more effective than mass learning for a variety of materials, including L2 vocabulary (Baddeley, 1997; Hulstijn, 2001; Mizuno, 1996, 2003) and a computer program could be used to facilitate such study intervals. Another possible advantage in the use of a CALL program in vocabulary acquisition is its “ability to meet unobtrusively learners’ individual needs [as] an inherent feature of multimedia CALL; it allows learners to select the options which best assist them in deriving correct word meanings” (Grace, 1998, p. 38). Furthermore, the use of CALL tools like readily available dictionaries and thesauri can give students not only the correct definitions of terms, but also examples in context plus other information, all of which helps to stimulate vocabulary acquisition (Ellis, 1995).

Other possible applications of CALL activities include potentially improved understanding of multidimensional language constructs, more useful testing (Chapelle, 2001), corpus samples used during a learning activity or as feedback (Desmet & Paulussen, 2007), smart phone systems and associated personalizable functions (Pemperton, 2007), or many other uses. Further examples are the algorithm developed by Wible et al. (2006) to locate chunks automatically within the millions of words in the British National Corpus or the sophisticated ICALL programs that make use of artificial intelligence and can therefore respond to users in ways that is not explicitly programmed (O’Brien, 1994). Another application of computers in language learning is to link all the computers in a classroom to provide not only learner-computer interactions but also learner-learner interactions (Chapelle, 2001) in the form of synchronous and asynchronous computer-mediated communication.

**RESEARCH QUESTIONS**

Based on the literature review above, the project described here addresses two research questions. In the general context of comparing a technology application approach and a traditional flashcard approach for vocabulary acquisition and retention,

1. Does a CALL image-based method done in small groups favor the acquisition of 10 lexical chunk items over a paper card image-based method done in small groups?

2. Does a CALL image-based method done in small groups favor the retention of 10 lexical chunk items over a paper card image-based method done in small groups?
METHOD

Participants

A total of 71 first-year university students participated in the study. The students involved in the study were all streamed at the university's 'A' level and were further divided into two classes based on their major field of study. One class consisted of 38 Management majors and the other of 33 Environmental Engineering majors. A pretest of 30 potential items to be used for the two vocabulary study methods was given to these students. To determine whether the two groups differed in their knowledge of the vocabulary items, a t test was applied to their pretest scores (see Table 1).

<table>
<thead>
<tr>
<th>Group</th>
<th>M</th>
<th>SD</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Picture card</td>
<td>14.79</td>
<td>2.663</td>
<td>0.145</td>
<td>.886</td>
</tr>
<tr>
<td>CALL</td>
<td>14.70</td>
<td>2.721</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The figures in Table 1 show that the mean scores of the two groups on the pretest were not significantly different and that the two groups can be assumed to be equal at the beginning of the project.

Variables

Care was taken to minimize the effect of extraneous variables in the project. The students in both groups had the same instructor for the duration of the experiment and had equal practice time in two study methods. The same sentence types were used in the pre- and posttests and differed from the sentence types used in the treatment. When testing, both the CALL group and the picture card group were allowed to view the images that had become associated with the vocabulary items (Mohseni-Far, 2008a; Nikolova, 2002; Papagno & Vallar, 1992; Schmitt, 2000). This was done for the CALL group via a projected composite of the images used in the program (see Figure 1) and for the card group by allowing each student to use the cards that they had created themselves.
Students were actively discouraged from studying the vocabulary items outside of the treatment sessions. The only uncontrolled, and arguably uncontrollable, variable was that of incidental exposure to the items being studied. This could have been a factor because the institution where the study took place preferred that the vocabulary items used in the study be taken from the course textbook. Thus, students could have had extra practice with the items while studying for the course in general. However, this factor is not likely to have made much of a difference in the project because all students had the same opportunities for extra exposure as the course progressed.

**Procedure**

**Instructions**

Before the study began, a full explanation of the details of the study was given to the potential subjects in the two classes in both English and Japanese. Consent forms were then distributed. All 71 students in the two classes agreed to participate in the study.
Choosing the vocabulary items

Thirty potential vocabulary items were chosen using the guidelines outlined in the lexical approach (Lewis, 1993; Nattinger & DeCarrico, 1992; see item list in Appendix A). Students were instructed to write the Japanese translation of each item as a pretest. Following this, the researchers’ translations of each of the items were triangulated with a native speaker of Japanese. The pretests were then marked, and the number and percent known of each item in the list were calculated. The least well known items were separated into two groups of 10 items each; the first 10 were to be used for practice with each of the two methods, and the remaining 10 for the treatment itself. All of the items chosen, the percentages known of each, and the differences between percentages known by quarter are listed in Table 2.

Table 2
Vocabulary Items Chosen for Practice and Treatment Sessions

<table>
<thead>
<tr>
<th>Practice items</th>
<th>Treatment items</th>
<th>Percent of practice items known</th>
<th>Percent of treatment items known</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stick things together</td>
<td>Find the circumference of</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00</td>
</tr>
<tr>
<td>Move around something</td>
<td>Store media on an iPod</td>
<td>0.00%</td>
<td>1.39%</td>
<td>Treatment &gt; practice = +1.39</td>
</tr>
<tr>
<td>Hang out with friends</td>
<td>Measure the voltage of a battery</td>
<td>1.39%</td>
<td>1.39%</td>
<td>0.00</td>
</tr>
<tr>
<td>Insert into a slot</td>
<td>Step over something</td>
<td>12.50%</td>
<td>2.78%</td>
<td>Practice &gt; treatment = +9.72</td>
</tr>
<tr>
<td>Commute to school</td>
<td>Capture sound</td>
<td>20.83%</td>
<td>16.67%</td>
<td>Practice &gt; treatment = +4.16</td>
</tr>
<tr>
<td>Spend money on</td>
<td>Make/receive a call</td>
<td>37.50%</td>
<td>40.28%</td>
<td>Treatment &gt; practice = +2.78</td>
</tr>
<tr>
<td>Recognize a face</td>
<td>Check my email</td>
<td>44.44%</td>
<td>50.00%</td>
<td>Treatment &gt; practice = +5.56</td>
</tr>
<tr>
<td>Study mechanical engineering</td>
<td>Climb up and down stairs</td>
<td>54.17%</td>
<td>55.56%</td>
<td>Treatment &gt; practice = +1.39</td>
</tr>
<tr>
<td>Re-charge batteries</td>
<td>Stay up all night</td>
<td>56.94%</td>
<td>62.50%</td>
<td>Treatment &gt; practice = +5.56</td>
</tr>
<tr>
<td>Play video games</td>
<td>Get my driver’s license</td>
<td>68.06%</td>
<td>63.89%</td>
<td>Practice &gt; treatment = +4.17</td>
</tr>
</tbody>
</table>

Practice sessions

*Picture card study method.* Following the pretest, both groups completed 10 practice sessions for each study method. The first method practiced was the picture card method. At the beginning of the first session of this method, a list of the items to be studied was distributed to all
the students, and the Japanese translations of each item was projected onto a large screen in keeping with Nation’s recommendations on teaching vocabulary (Nation, 2005). After the students had copied down the L1 meanings on their item list of items, they were given a blank sheet of paper. They were instructed to use the next 10 minutes to divide their blank sheets into 10 separate cards and draw a picture that they thought best represented each of the vocabulary items. At the end of 10 minutes, the students took a sentence fill-in-the-blank test on the course website (see sample screenshot in Figure 2 and complete test in Appendix B; note that the items are from the list of treatment vocabulary items).

Figure 2
Online Vocabulary Test

Students were allowed to use their cards while taking the test. In the following four study sessions, also of 10 minutes each, the students were divided into groups of three or four and used their cards to play a ‘Memory’ matching game, saying each item aloud as they flipped over the cards. Students were allowed to consult their vocabulary item lists as they played the game but were not allowed to write the actual vocabulary item directly onto their cards. After the fifth and final practice session, students were again directed to take another online vocabulary test (see test in Appendix C).

CALL study method. After the students finished the picture card practice sessions, they completed sessions using the CALL study method with the same 10 items. This time, however, students were separated into pairs and instructed to use the matching program on the course website to study the items; they were reminded to say the items aloud as they matched them (see Figure 3; again the items from the list of treatment vocabulary items). The drop menu next to each image contained the 10 vocabulary items as possible answer choices. Students studied in this way for five sessions of 10 minutes each.
Treatment sessions
Following the practice sessions, the treatment sessions for each group began. The Management majors’ class used the picture card method to study the 10 treatment vocabulary items, and the Environmental Engineering majors’ class used the CALL method to study the same 10 items. Each class was given the list of new vocabulary items, and once again the L1 meaning of each item was provided for the students using the classroom projector. The students in each group used their specific method to study the vocabulary items for five sessions of 10 minutes each. An initial test was given after the first session (see Appendix D) and a post-test after the fifth session (see Appendix E). Students studying using the picture cards were again allowed to use the cards during the test, and the CALL group had the images used in the program projected for use as memory aids. After all of the students took the posttest in the final study session, the students in each group completed a separate Likert-scale survey (see Appendices F and G). All of the survey questions and answer choices were explained in the students’ L1 to ensure that there were no issues of misunderstanding obstructing their responses. Finally, 2 weeks after the final study session, a surprise delayed retention test was given (see test in Appendix H).

RESULTS
Scores
A visual inspection of the distribution of scores on the treatment initial tests compared to those on the posttests indicates improvement for both the picture card group and the CALL group. The mean scores on the treatment tests for the picture card group increased from 5.44
to 6.72, while the scores for CALL group increased from a mean of 5.86 to 6.57. Additionally, both groups showed a high rate of retention on the delayed posttest with means of 5.16 and 5.45.

The surveys given to each group after the final treatment showed a remarkable balance of opinion (see Appendix I). The results, taken together, indicate that the CALL method was slightly more popular and considered more useful than the picture card method, despite the higher percentage of students who stated that they would want to study vocabulary with the picture cards again rather than with the CALL method.

**Analysis**

A t test was performed on the data gathered from the treatment initial tests and treatment posttests for both groups (see Table 3).

<table>
<thead>
<tr>
<th>Group</th>
<th>Treatment initial test</th>
<th>Treatment posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Picture card group (n = 36)</td>
<td>5.44 1.557 1.130 .263</td>
<td>6.72 0.454 1.108 .272</td>
</tr>
<tr>
<td>CALL group (n = 28)</td>
<td>5.86 1.297</td>
<td>6.57 0.634</td>
</tr>
</tbody>
</table>

The figures in Table 3 show no significant difference between the groups on either the treatment initial test or treatment posttest. One interesting point to note in the t-test results is the tightening of the standard deviations between the initial tests and posttests. Although both groups experienced this change, it was more pronounced for the picture card group in which the standard deviation went from 1.557 to 0.454, compared to the CALL group in which the standard deviation went from 1.297 to 0.634. This point will be further explored below.

However, the data were not normally distributed; to corroborate the results of the t test a nonparametric Mann-Whitney Test was also used (see Table 4).

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean rank</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment initial test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture card group (n = 36)</td>
<td>30.03</td>
<td>.209</td>
</tr>
<tr>
<td>CALL group (n = 28)</td>
<td>35.68</td>
<td></td>
</tr>
<tr>
<td>Treatment posttest</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture card group (n = 36)</td>
<td>33.89</td>
<td>.402</td>
</tr>
<tr>
<td>CALL group (n = 28)</td>
<td>30.71</td>
<td></td>
</tr>
</tbody>
</table>

The results of this analysis show no difference between the groups.
A *t* test was also performed on the delayed test data (see Table 5).

### Table 5

**T-Test Results of Delayed Posttest**

<table>
<thead>
<tr>
<th>Group</th>
<th>M</th>
<th>SD</th>
<th><em>t</em></th>
<th><em>p</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Picture card group (n = 37)</td>
<td>5.16</td>
<td>1.500</td>
<td>-0.808</td>
<td>.422</td>
</tr>
<tr>
<td>CALL group (n = 31)</td>
<td>5.45</td>
<td>1.434</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The *t*-test results show no difference between the groups.

Once again, the data were not normally distributed, and a Mann-Whitney test was used to support the results of the *t* test (see Table 6).

### Table 6

**Mann-Whitney Test Results of Delayed Posttest**

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean rank</th>
<th><em>p</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Picture card group (n = 37)</td>
<td>32.74</td>
<td>.412</td>
</tr>
<tr>
<td>CALL group (n = 31)</td>
<td>36.60</td>
<td></td>
</tr>
</tbody>
</table>

Again, the results of this analysis show no difference between the groups.

### DISCUSSION OF RESULTS AND CONCLUSION

**Interpretation**

In response to the original research questions, the results of the study showed no significant difference between the two methods used for both vocabulary acquisition and retention. This is not to say that the methods were not effective, however, because both the picture card group and the CALL group showed improvement from the pretest to the treatment initial test and from the treatment initial test to the treatment posttest. Although the change from the pretest scores to the treatment initial test scores can largely be explained by the decrease in the number of items tested from 30 to 10, which formed a narrower bandwidth and much coarser scale, the increase in the scores from the treatment initial test to the treatment posttest represents the acquisition of material. That this acquisition occurred is demonstrated by the higher means and tightening of the standard deviations. Therefore, although the null hypothesis must be accepted in regards to the research question, it can be stated that both methods are equally valid ways of studying vocabulary. Moreover, a high percentage of the items were retained on the delayed retention test from the treatment posttest, 76.79% of the items for the picture card group and 82.95% of the items for the CALL group. Finally, the survey results showed that both groups enjoyed studying with their respective methods and found them useful for learning.

**Methodological Implications and Future Research**

The survey results which showed the CALL method to be more popular and considered more useful overall indicate that where possible the application of an image-based CALL vocabulary study program would be beneficial to students in an L2 classroom. Teachers can readily produce such a program by using a simple matching program like the Java Hot Potatoes program
used in this study (http://hotpot.uvic.ca). Moreover, students can be allowed to choose the images for the programs themselves, thereby encouraging creative experiences with student-authored materials (Nikolova, 2002). Given the right CALL tools (e.g., art or graphics software), students can even produce files of images that they had imagined and apply them to the matching program. In situations in which CALL development is not possible, the results of the present study indicate that student-authored picture cards are a good alternative to CALL programs. Both methods involve deep mental processing of the material involved, and the context provided by the images helps promote LTM storage of the items.

Future research could well extend the findings of the present study. Of particular importance for future research would be the sole use of vocabulary items that were completely unknown by all subjects. This was not possible in the current study due to the quasi-experimental setting and accompanying institutional constraints. The inclusion of more subjects to accentuate the detail of the scale and allow for more robust statistical findings would also benefit future research, as would lengthier testing devices.

**Limitations**

The most striking limitation in the present study is that not all of the items were completely unknown at the outset of the project. This was an unfortunate consequence of the fact that an experimental setting was not possible. Another limitation of the quasi-experimental setting was that the subjects could not be randomly assigned to the two groups; both groups also suffered from the unexpected absences of subjects. Yet another limitation was the size of the tests themselves; almost certainly longer and more varied tests would have given a clearer picture of just how well the items were learned and retained.

**Conclusion**

In the research reported above no significant difference was found in the acquisition and retention of 10 vocabulary items between the group that studied the items using a picture card-based method and the group using a CALL-based method. The study methods incorporated recent findings on human memory, the phonological/articulatory loop, uses of imagery, and the lexical approach. The data showed the acquisition of the material studied by both groups from an initial test done after the first treatment session and a posttest done after the fifth session. In addition, the results of a surprise delayed test of retention revealed that 76% of the items were retained in the picture card group and 82% in the CALL group. Finally, the participants also generally responded positively on a posttreatment survey regarding the methods employed. It is therefore argued here that both the picture card method and the CALL method can be efficient and effective means of vocabulary acquisition and retention.

**REFERENCES**


**APPENDIX A**

Potential Vocabulary Study Items

1. insert into a slot
2. play video games
3. get my driver’s license
4. commute to school
5. study mechanical engineering
6. study in the library
7. watch videos
8. stick things together
9. recognize a face
10. capture sound
11. mix it together
12. re-charge batteries
13. hike in the mountains
14. work at my part-time job
15. climb up and down stairs
16. turn the lights on/off
17. take photos
18. check my email
19. stay up all night
20. move around something
21. step over something
22. eat breakfast/lunch/dinner
23. meet my friends
24. hang out with friends
25. find the circumference
26. store media on an iPod
27. send messages from my mobile phone
28. make/receive a call
29. measure the voltage of a battery
30. spend money on
APPENDIX B

Picture Card Practice Test

Here the answer choices are listed for the first question only, but that in the online version of the test all seven questions contained all 10 items as answer choices.

1. The ability to know who someone is by looking at them is called being able to ________.
   
   A. stick things together  
   B. move around something  
   C. hang out with friends  
   D. insert into a slot  
   E. commute to school  
   F. spend money on  
   G. recognize a face  
   H. study mechanical engineering  
   I. re-charge batteries  
   J. play video games  

2. If I have some extra cash I usually __________ buying new CDs.

3. On weekends I like to ___________. Sometimes we play sports

4. When you are driving, if you want to ____________ then you need

5. Here at KUT, many students want to _____________. An example of how to use the skills learned this way would be to design and maintain industrial equipment.

6. After classes I sometimes ____________ with my friends. They have an Xbox and PS3 system, and I have a DS.

7. To put a memory stick into a PC, you need to find the USB portals and then __________ (it) __________.
APPENDIX C
Picture card Posttest

Here the answer choices are listed for the first question only, but that in the online version of the test all seven questions contained all 10 items as answer choices.

1. It is easy to ___________ if you use a powerful glue or other adhesive.
   A. stick things together
   B. move around something
   C. hang out with friends
   D. insert into a slot
   E. commute to school
   F. spend money on
   G. recognize a face
   H. study mechanical engineering
   I. re-charge batteries
   J. play video games

2. Many KUT students ___________ by bicycle, but some students take the train or bus.

3. With a lot of use mobile phones can lose their power quickly, so it is important to know when you need to ___________ (the) ___________.

4. If I do not ___________ at a party I feel very nervous because I do not know anyone there.

5. You can start that slot machine if you take a coin and ___________ (it) ___________ near the handle.

6. Many people ___________ their hobbies, especially if their hobbies involve using expensive machines like motorcycles or cars.

7. A new kind of robotic vacuum cleaner can automatically ___________ that gets in its way.
APPENDIX D

Treatment initial test

Here the answer choices are listed for the first question only, but that in the online version of the test all seven questions contained all 10 items as answer choices.

1. A microphone is a very good way to ___________. If the microphone is attached to a video camera you can also get an image to go with it.
   A. find the circumference
   B. store media on an iPod
   C. measure the voltage of a battery
   D. step over something
   E. capture sound
   F. make/receive a call
   G. check my email
   H. climb up and down stairs
   I. stay up all night
   J. get my drivers license

2. On New Years Eve many people like to ___________ and celebrate the changing of the calendar year.

3. To ___________ of a circle you need to multiply pi by the diameter.

4. To exercise their legs, some people who enjoy running also ___________. Often they will do this at sports stadiums.

5. Junichi didn’t like jumping, so if there was an object at his feet he would always _____ rather than to jump across it.

6. Everyday when I wake up, the first thing I do is turn on my computer and ___________.

7. A mobile phone is useful for communicating. You can use it to ___________ or to send and get email messages.
APPENDIX E
Treatment posttest

Here the answer choices are listed for the first question only, but that in the online version of the test all seven questions contained all 10 items as answer choices.

1. After practicing on the course and roads for weeks, I was finally able to ____________.
   A. find the circumference
   B. store media on an iPod
   C. measure the voltage of a battery
   D. step over something
   E. capture sound
   F. make/receive a call
   G. check my email
   H. climb up and down stairs
   I. stay up all night
   J. get my drivers license

2. You can use this Voltmeter to ____________, but be careful not to accidentally shock yourself.

3. My friend is a musician, and he likes to ____________ whenever he goes around the city. He has an interesting audio collection of everyday things he has heard and recorded.

4. You can ____________. In fact, you can put music, videos, or even photos on one. The most common memory sizes are 8 GB, 16 GB, and 30 GB, but other sizes are available as well.

5. Pi x D is the formula to ____________ of a circle. It can also be written as pi times the diameter.

6. My brothers’ son wanted to ____________ watching movies, but my brother said he could not because he needs sleep to stay healthy.

7. The first telephones could only ____________, but modern phones have many more features such as speed-dial and caller ID.
APPENDIX F
Picture card group survey

Please choose only one option for each statement below.

1. I enjoyed studying words with the picture cards.
   □ Strongly agree
   □ Agree
   □ Neither agree nor disagree
   □ Disagree
   □ Strongly disagree

2. Studying words with the picture cards was useful for remembering vocabulary.
   □ Strongly agree
   □ Agree
   □ Neither agree nor disagree
   □ Disagree
   □ Strongly disagree

3. I want to study words with the picture cards again.
   □ Strongly agree
   □ Agree
   □ Neither agree nor disagree
   □ Disagree
   □ Strongly disagree

4. Studying words with the picture cards is more useful for remembering vocabulary than studying with the computer program.
   □ Strongly agree
   □ Agree
   □ Neither agree nor disagree
   □ Disagree
   □ Strongly disagree

5. Studying words with the computer program is more useful for remembering vocabulary than studying with the picture cards.
   □ Strongly agree
   □ Agree
   □ Neither agree nor disagree
   □ Disagree
   □ Strongly disagree
APPENDIX G

CALL group survey

Please choose only one option for each statement below.

1. I enjoyed studying words with the computer program.
   - □ Strongly agree
   - □ Agree
   - □ Neither agree nor disagree
   - □ Disagree
   - □ Strongly disagree

2. Studying words with the computer program was useful for remembering vocabulary.
   - □ Strongly agree
   - □ Agree
   - □ Neither agree nor disagree
   - □ Disagree
   - □ Strongly disagree

3. I want to study words with the computer program again.
   - □ Strongly agree
   - □ Agree
   - □ Neither agree nor disagree
   - □ Disagree
   - □ Strongly disagree

4. Studying words with the computer program is more useful for remembering vocabulary than studying with the picture cards.
   - □ Strongly agree
   - □ Agree
   - □ Neither agree nor disagree
   - □ Disagree
   - □ Strongly disagree

5. Studying words with the picture cards is more useful for remembering vocabulary than studying with the computer program.
   - □ Strongly agree
   - □ Agree
   - □ Neither agree nor disagree
   - □ Disagree
   - □ Strongly disagree
APPENDIX H

Delayed retention test

Here the answer choices are listed for the first question only, but that in the online version of the test all seven questions contained all 10 items as answer choices.

1. His father hurt his leg so badly that he could not ____________ and always had to use an elevator.
   A. find the circumference
   B. store media on an iPod
   C. measure the voltage of a battery
   D. step over something
   E. capture sound
   F. make/receive a call
   G. check my email
   H. climb up and down stairs
   I. stay up all night
   J. get my drivers license

2. The salesperson said that to ____________ was very simple. You only had to plug it into your computer and then choose what to put into its memory.

3. Among other things, you will need to know how to accurately ____________ if you want to keep your car running smoothly. If there are problems with the alternator, your car could lose all power.

4. Before I can ____________, I have to prove to my parents that I will be very safe and careful when I am on the roads.

5. Because she thought it was unlucky, she always refused to ____________, and instead walked around whatever was in front of her.

6. I tried to ____________ yesterday but realized that I had completely forgotten both my address and password. So if you need to contact me, then please just call me.

7. My geometry teacher told me to ____________ of the circle, but I forgot the formula so I could not solve the problem.
**APPENDIX I**
Survey results

<table>
<thead>
<tr>
<th>Statement</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I enjoyed studying words with the picture cards.</td>
<td></td>
</tr>
<tr>
<td>Strongly agree</td>
<td>15</td>
</tr>
<tr>
<td>Agree</td>
<td>19</td>
</tr>
<tr>
<td>Neither agree or disagree</td>
<td>2</td>
</tr>
<tr>
<td>Disagree</td>
<td>0</td>
</tr>
<tr>
<td>Strongly disagree</td>
<td>0</td>
</tr>
<tr>
<td>2. Studying words with the picture cards was useful for remembering vocabulary.</td>
<td></td>
</tr>
<tr>
<td>Strongly agree</td>
<td>11</td>
</tr>
<tr>
<td>Agree</td>
<td>19</td>
</tr>
<tr>
<td>Neither agree or disagree</td>
<td>6</td>
</tr>
<tr>
<td>Disagree</td>
<td>0</td>
</tr>
<tr>
<td>Strongly disagree</td>
<td>0</td>
</tr>
<tr>
<td>3. I want to study words with the picture cards again.</td>
<td></td>
</tr>
<tr>
<td>Strongly agree</td>
<td>14</td>
</tr>
<tr>
<td>Agree</td>
<td>14</td>
</tr>
<tr>
<td>Neither agree or disagree</td>
<td>8</td>
</tr>
<tr>
<td>Disagree</td>
<td>0</td>
</tr>
<tr>
<td>Strongly disagree</td>
<td>0</td>
</tr>
<tr>
<td>4. Studying words with the picture cards is more useful for remembering vocabulary than studying with the computer program.</td>
<td></td>
</tr>
<tr>
<td>Strongly agree</td>
<td>11</td>
</tr>
<tr>
<td>Agree</td>
<td>18</td>
</tr>
<tr>
<td>Neither agree or disagree</td>
<td>7</td>
</tr>
<tr>
<td>Disagree</td>
<td>0</td>
</tr>
<tr>
<td>Strongly disagree</td>
<td>0</td>
</tr>
<tr>
<td>5. Studying words with the computer program is more useful for remembering vocabulary than studying with the picture cards.</td>
<td></td>
</tr>
<tr>
<td>Strongly agree</td>
<td>6</td>
</tr>
<tr>
<td>Agree</td>
<td>12</td>
</tr>
<tr>
<td>Neither agree or disagree</td>
<td>9</td>
</tr>
<tr>
<td>Disagree</td>
<td>6</td>
</tr>
<tr>
<td>Strongly disagree</td>
<td>3</td>
</tr>
</tbody>
</table>
### CALL Group

<table>
<thead>
<tr>
<th>Statement</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
</tr>
<tr>
<td>1. I enjoyed studying words with the computer program.</td>
<td></td>
</tr>
<tr>
<td>Strongly agree</td>
<td>7</td>
</tr>
<tr>
<td>Agree</td>
<td>20</td>
</tr>
<tr>
<td>Neither agree or disagree</td>
<td>1</td>
</tr>
<tr>
<td>Disagree</td>
<td>0</td>
</tr>
<tr>
<td>Strongly disagree</td>
<td>0</td>
</tr>
<tr>
<td>2. Studying words with the computer program was useful for</td>
<td></td>
</tr>
<tr>
<td>remembering vocabulary.</td>
<td></td>
</tr>
<tr>
<td>Strongly agree</td>
<td>7</td>
</tr>
<tr>
<td>Agree</td>
<td>17</td>
</tr>
<tr>
<td>Neither agree or disagree</td>
<td>3</td>
</tr>
<tr>
<td>Disagree</td>
<td>1</td>
</tr>
<tr>
<td>Strongly disagree</td>
<td>0</td>
</tr>
<tr>
<td>3. I want to study words with the computer program again.</td>
<td></td>
</tr>
<tr>
<td>Strongly agree</td>
<td>6</td>
</tr>
<tr>
<td>Agree</td>
<td>13</td>
</tr>
<tr>
<td>Neither agree or disagree</td>
<td>9</td>
</tr>
<tr>
<td>Disagree</td>
<td>0</td>
</tr>
<tr>
<td>Strongly disagree</td>
<td>0</td>
</tr>
<tr>
<td>4. Studying words with the computer program is more useful for</td>
<td></td>
</tr>
<tr>
<td>remembering vocabulary than studying with the picture cards.</td>
<td></td>
</tr>
<tr>
<td>Strongly agree</td>
<td>6</td>
</tr>
<tr>
<td>Agree</td>
<td>15</td>
</tr>
<tr>
<td>Neither agree or disagree</td>
<td>3</td>
</tr>
<tr>
<td>Disagree</td>
<td>3</td>
</tr>
<tr>
<td>Strongly disagree</td>
<td>1</td>
</tr>
<tr>
<td>5. Studying words with the picture cards is more useful for</td>
<td></td>
</tr>
<tr>
<td>remembering vocabulary than studying with the computer program.</td>
<td></td>
</tr>
<tr>
<td>Strongly agree</td>
<td>5</td>
</tr>
<tr>
<td>Agree</td>
<td>7</td>
</tr>
<tr>
<td>Neither agree or disagree</td>
<td>4</td>
</tr>
<tr>
<td>Disagree</td>
<td>11</td>
</tr>
<tr>
<td>Strongly disagree</td>
<td>1</td>
</tr>
</tbody>
</table>
AUTHOR’S BIODATA
Andrew Oberg is Assistant Professor at Surugadai University. His academic interests include vocabulary learning, memory and neural change, and the application of CALL methods to the language classroom.

AUTHOR’S ADDRESS
Andrew Oberg
Sha Maison Pinolo 203
Mama 408
Kochi-shi, Kochi-ken
Japan 780-0973
Phone: +81 (90) 8970-7146
Email: obergandrew@gmail.com