

# Investigating an intelligent system for vocabulary learning through reading

**Glenn Stockwell**

Waseda University, JAPAN

gstock@waseda.jp

While learners can acquire vocabulary through extensive reading (Pigada & Schmitt, 2006), research suggests that acquisition can be more effective when supplemented with targeted vocabulary activities (e.g., Paribakht & Wesche, 1997). Problems arise, however, in determining what vocabulary learners have acquired, and what items should be focused on in these vocabulary activities. The main purpose of the current study is to describe the development and trial implementation of an intelligent system which created individualized vocabulary activities for each learner based on hyperlinked words that were clicked on during reading written passages in the target language. In this preliminary study, 43 Japanese learners of English were required to read several mid-length passages on computer during class time. Content words were hyperlinked to a separate window which provided meanings, pronunciation, and example sentences. Whenever a word was clicked it was added to a database and activities were automatically created for each learner that could be completed outside of class time, on either a computer or a mobile phone. Learners were given a pre-test at the beginning and a post-test on completion of the semester. Logs were kept regarding clicking patterns during reading. There were two specific objectives in the study. Firstly, the ways in which learners clicked on words on the passage were investigated to see if there were any differences in how words deemed as “known” and “unknown” were looked up, and whether this linked to acquisition of unknown words. Secondly, lists of the words looked up by the learners were analysed to determine if it was possible to create individual profiles of learners’ vocabulary knowledge. Data collected in the current study included the correlation

*between the words considered as unknown according to the pre-test and the words clicked on during reading, and the length of time spent looking at word descriptions, and the results of the vocabulary post-tests.*

## Introduction

There is a significant body of research that argues for the existence of a clear relationship between vocabulary and reading (Grabe, 2009; Waring, 2009), and that reading extensively can play a role in facilitating incidental learning of vocabulary (e.g., Coady, 1997; Nation, 2001). Reading has the potential to help learners acquire not only the meaning of the words they encounter, but also improve spelling and grammatical usage (Pigada & Schmitt, 2006). In saying this, however, learners need to be exposed to an extremely large amount of input in order for gains to be made and retained (Nation, 2008) and acquisition can be both unpredictable and time-consuming (Zimmerman, 1997).

In order to deal with this problem, Paribakht and Wesche (2007) argue that vocabulary acquisition can be significantly enhanced if reading is supplemented with activities which target specific vocabulary items. While this may seem an ideal way to improve learners' vocabulary knowledge, it can be very difficult to determine what vocabulary items to target. One possibility is controlling the content of what learners read and then to create activities which focus on what the teacher assumes that the learners need to learn. Given that teachers often do have a reasonable idea of what their learners do and do not know, this can work if the learner group is relatively homogeneous, but if there is large variation in the knowledge of the learners within a class, then this has the potential to be frustrating for those with a larger (or significantly smaller) vocabulary. Another possibility is to collect data about the words learners look up the meaning of during the reading process, and then to base activities on the words that they look up. This has the potential to provide a clearer picture of the vocabulary that each individual learner is unsure of the meaning of, but of course there are problems with this approach as well. Firstly, this can only give an indication of the words that learners actually look up, which excludes words that learners guess from context. This is a problem that would indeed be difficult to overcome, but at the very least, it would be possible to provide learners with activities that target words that they may be less confident with. Another conceivably more serious problem is that keeping records of items that learners look up would be very difficult to manage. Learners could keep manual logs themselves, but this has the potential to interrupt the reading process. Similarly, it would be exceptionally difficult, if not impossible, for teachers to manually keep records of words looked up, particularly if there are a large number of learners in a single class.

Given the labour-intensive nature of manually keeping records of learners' activities, technology seems to be a logical alternative, where record-keeping can be done in an automated and non-intrusive manner. Using technology for learning vocabulary is far from new, and there has been a range of studies that have appeared in the **CALL** literature. Indeed, teaching and learning vocabulary has consistently attracted attention from teachers and researchers since the early days of **CALL** (see Healy, 1999). Making up around one-third of all empirical research in four major **CALL** journals from 2001 to 2005, vocabulary still remains one of the most commonly researched areas in the field, and empirical research has been carried out using various types of courseware, online activities, online and electronic dictionaries, and corpora and concordancing (Stockwell, 2007). Studies that look specifically

focus, tending to concentrate mostly on the effect of looking up words and the information provided to learners when these words are looked up, as described in the following section.

## Learning vocabulary through reading in CALL

Encouraging learners to look up words that they do not know during reading might be considered as an important step towards acquiring them. It can be both time-consuming and frustrating for learners if they are required to infer the meaning purely from the context, and vocabulary learning is enhanced when learners consult dictionaries while reading compared with when they do not (Knight, 1994). If the process of looking up new words requires effort – as is often the case when using paper-based dictionaries – there is the danger that learners will simply skip over unfamiliar words rather than taking the time to look them up. Evidence for this is provided in a study by Koyama and Takeuchi (2007), who showed that learners engaged in reading comprehension activities were far more likely to look up words using handheld electronic dictionaries than they were to use paper-based dictionaries, although they did not find any relationship between the frequency of looking up of words and acquisition of these words. This suggests, then, that the information gained from looking up unknown words on a single occasion has a similar effect as looking up words repeatedly. Thus, the challenge then becomes not to make learners look up a word a number of times, but to make it as easy as possible for learners to look up words, and then to provide sufficient information the first time around to enhance their knowledge of it.

A means that has been used to encourage learners to look up unknown words during over the past several years has been the use of annotations, of either a textual (e.g., De Ridder, 2002) or non-textual nature such as pictures (e.g., Yoshii & Flaitz, 2002) or pictures and sound (Yeh & Wang, 2003), either with or without textual annotations. The use of annotations makes it easier for learners to click on the words that they want to find out the meaning of without interrupting the reading process any more than is necessary, which would result in learners being more likely to look up words that they are unfamiliar with rather than simply skipping them. Simply looking up words does not necessarily provide the best conditions for learning unknown vocabulary, and there is evidence suggesting that looking up words during reading for comprehension leads to improved retention of vocabulary. Peters (1997) showed that Dutch learners of German were more likely to acquire new words when reading passages for meaning than when informed that the vocabulary appearing in reading passages would appear in an upcoming vocabulary test. In this study, Peters required two groups of learners to read a passage written in German in order to complete reading comprehension activities, using an online dictionary if they encountered any unknown words while. In the first group, learners were not informed of a vocabulary test that took place directly after the reading was completed, while the learners in the second group were told about the test. She found that the learners in the second group looked up more words than the learners in the first group, but there was no significant difference between the groups in the number of times that words were looked up. In contrast, learners in the first group scored significantly better on the post-reading vocabulary test, particularly if the words that they looked up were required in order to complete the reading comprehension activities.

A recurring argument in the literature has been the need for care to be taken in the design of reading comprehension tasks and activities, as differential results have been achieved in retention of vocabulary acquired through reading depending on the type of task **261**

which has been used. De Ridder (2002) points out that if the activities are too demanding, then they have the potential to distract learner attention away from the new vocabulary that they encounter, which could decrease retention. Similarly, as Peters (2007) suggests, reading comprehension tasks need to be designed in such a way that the meanings of vocabulary items the teacher wishes to focus on must be understood in order to complete the tasks. Both of these are important issues to consider in designing reading comprehension tasks and activities which have a specific focus on learning vocabulary. The tasks and activities must be simple enough so that the learners' attention is not diverted entirely away from the process of learning vocabulary, but at the same time, these tasks and activities must require sufficient comprehension of the particular items to complete, meaning that they must be adequately challenging so that they cannot be completed without clearly understanding the meaning of key vocabulary that appear.

While this research sheds valuable light on the process of learning vocabulary through reading, one of the problems is that it still largely depends on learners acquiring words incidentally during the reading process. Even if learners are provided with various cues regarding the meanings of the words as a result of annotations, the primary purpose remains, for the learner, at the very least, comprehension of the reading comprehension task or activity at hand. In line with Paribakht and Wesche's (2007) arguments, acquisition is enhanced if supplemented with activities which target specific vocabulary, but, as pointed out above, this is difficult without using technology to assist in keeping accurate records of words that learners look up during reading. Thus, the purpose of the current study was to investigate the development and trial implementation of a system which not only provided learners with annotations they could access during reading comprehension activities, but also kept records of the items that were looked up, and generated activities targeting the words that they looked up during the reading process. In addition, based on the clicking patterns of the learners, the system also aimed to construct a profile of the learners' vocabulary knowledge which could be used to give teachers better insight into what vocabulary items learners actually need assistance with and help choose or design materials more suited to the learners. The design of the system is described in the following section.

## System description

An intelligent system was developed which consisted of two main components: a reading activities component and a vocabulary activities component. Written using **PHP** and **MySQL** and integrated into Moodle, the system was intelligent in that provided vocabulary activities that were generated specifically for each learner depending on the words that they clicked on during the reading and on their performance during the vocabulary activities themselves. The architecture of the system was designed based on the modules outlined by Kang and Maciejewski (2000), these being the expert knowledge module, the student knowledge module, the tutoring module, and the user interface module. The expert knowledge module is what provides the knowledge to be taught, and in the current study this was a dynamic module which changed based on the words that were clicked on while engaged in the reading comprehension activities. This expert knowledge module was fed by a larger overall module which contained the complete list of words in the **JACET** 8000 frequency list, plus a further 2000 words that appeared in the reading passages but not in the **JACET** list. The list was not deemed as being the ideal list given the relatively small range of vocabulary covered in it, but was used as it had already been put into digital form

along with information about the vocabulary items as part of an earlier project. The student knowledge module kept a record of the words that the learner got correct and incorrect during the vocabulary activities which were done after the reading comprehension, while the tutoring module and the user interface module controlled the interactions between the system and the learner. More information regarding the reading activities and the vocabulary activities are given below.

### *Reading activities*

The reading activities were carried out during class time, and were made up of a total of ten reading passages each of about 2500–3000 words in length, all of which needed to be completed by the end of the semester. They were based on topics that were considered to be of interest to the learners, such as alternative forms of energy, conservation, aged care, and so forth. Learners could access the reading passages in any order that they wished, and they were read during class time. Each of the content words in the reading was hyperlinked so that when it was clicked, a separate window opened which included the annotations. In this developmental stage, time constraints meant that it was not possible to provide as wide a range of non-textual information in the annotations as was initially hoped, and as a result, information was limited to L1 and L2 textual annotations (L1 translation, L2 meaning and example sentence, part of speech, and inflections) and an audio annotation (an audio recording of the word itself that learners could click on in the annotation window). The links in the reading passage were not initially salient, but changed colour when the mouse was passed over them.

The reading comprehension activities consisted of self-scoring multiple choice questions, and required the learners to have an understanding of a number of key words that appeared in the passages in order to answer them correctly. There were around 15–20 questions per reading passage, although there were some questions to which learners were required to select more than one answer out of a group of 8–9 choices. Records were kept of all of the words that were clicked on during reading and each word was entered into the expert knowledge module along with the information about the word from the larger database. The amount of time the window was kept open was also recorded, and kept a record of whether or not they clicked on the audio file of the word, and if so, how many times they clicked it. Learners were given a total of 30 minutes in which to read the passage and to answer the questions, which was considered to be adequate based on the practice session in the first class and through observation of learners during the first couple of weeks of class.

### *Vocabulary activities*

In contrast to the reading activities, the vocabulary activities were carried out outside of class time, to be completed before the following class a week later. The vocabulary activities were generated from the words that were clicked on while doing the reading comprehension activities, and consisted of five types: choosing the appropriate word for an English sentence, choosing the appropriate English word for a Japanese meaning, choosing the appropriate English word for an English definition, writing a word in English for an English definition, and writing the appropriate English word for an English sentence. The vocabulary component was developed as part of a separate project, and has been described in depth in an earlier study (see Stockwell, 2010), so further information has not been provided here. **263**

The activities could be accessed either from a desktop computer or from learners' mobile phones in the same way as the earlier study. Learners had completed vocabulary activities previously using this component of the system, hence were familiar with how to do them. The system kept records of scores, access times and the platform the learners used when they completed the activities.

One problem that was foreseen in designing the vocabulary activities was that once learners realised that the words that they clicked on during the reading comprehension activities would appear in the vocabulary activities, they would avoid clicking on words to reduce the amount of work to be done outside of class time. In order to overcome this, if there was an insufficient number of words clicked on during the reading passages, a number of words were randomly selected from the words appearing in the passage that the learner read based on clicking patterns of other students in the classes, bringing the total up to a minimum number of 12 words. There was no maximum cap on the number of words that could appear in the vocabulary activities.

## Method

As this was an exploratory study, the purpose was to get some indication of what learners would do with the system so that further developments could be made to refine both the system itself and the way in which it was implemented. A pre-test was administered to get some preliminary idea of what words the learners were familiar with from the outset, and then information regarding what they clicked on and how long they spent looking at the descriptions was collected. A post-test was used in order to measure acquisition and the impact of clicking on known and unknown words. The following research questions were posed:

1. Do learners look up meanings of words that are deemed "unknown" according to a pre-test?
2. Can vocabulary profiles be constructed through the annotated reading activities?
3. Does the looking up of unknown words lead to acquisition?

While not reaching the level of hypotheses, there were some general expectations behind the research questions that were posed. The idea behind the first question was the expectation that learners would predominantly click on words that they were not familiar with, but of course it was considered feasible that they would click on other words as well, even if the pre-test deemed that they knew them. The second question was posed in order to determine whether the system had value as a means of representing learners' knowledge that could be useful for other purposes such as materials design or test construction. The final question sought to determine whether or not looking up of unknown words through the system had any impact on their acquisition. Details of how the study was carried out are provided in the following subsections.

## Participants

The participants in the study were 43 first-year law major students at Waseda University, taking a compulsory English reading and writing course held in the second semester of the academic year. The levels of the learners ranged somewhat, but the majority were of a level

(one of 21 learners and the other of 22 learners), and their ages ranged from 18 to 21, made up of 29 male and 14 female students. Classes were held once a week for a 90-minute period over a period of 15 weeks. The students had used the Moodle system in which the reading system was integrated in the first semester, and accessed the system regularly in order to check grades and to submit essays that were required to be written during the semester.

### *Data collection and analysis*

There were two main ways through which data were collected in the study, vocabulary pre- and post-tests and system logs of learner access. The pre- and post-tests were given online and consisted of 120 items which were compiled from 10–14 words from each of the reading passages. Each consisted of two parts. In the first part, learners were asked to indicate how well they knew a word according to a four-point scale: (a) I know the meaning and I can use it in a sentence, (b) I know the meaning but I can't use it in a sentence, (c) I've seen it before but I don't know the meaning, and (d) I've never seen it before. In the second part, multiple choice questions were provided for each of the words that appeared but in a different order from the first part. If the learners indicated either (a) or (b) and got the word correct in the multiple choice word, it was considered as being "known." The reason why (b) was included here was that learners were required to know the meaning of the word in order to complete the reading comprehension activities, but there was no production making it impossible to distinguish between the two levels. The results of the pre-test were kept with details of the learner. An independent code was used to identify each learner in the online system which was completely separate from both their student number and the ID assigned to them in Moodle.

The system kept a log of every word that each learner clicked on, which could then be checked against the words in the pre- and post-test. In addition to this, the system indicated the amount of time that the annotation window was open, although, as described earlier, it was not possible from the system as it stood to determine exactly what the learners did with the information provided in the annotation window. At the end of the semester, the words that were clicked on during the reading passages were then collated with the pre-test results to determine the relationship between words that were deemed as "known" according to the pre-test and learner clicking patterns. These were later correlated with the post-test to determine whether clicking on the words had any relationship with acquisition.

### *Procedure*

In the first week of the semester, after the initial orientation for the class was completed, learners were provided with the pre-test. On completion of the pre-test, learners were given an explanation of the reading activities using a sample passage and reading comprehension activities. The learners were told that the reading activities were to be completed during class time, and that they should click on any words that they didn't know, which would result in the annotation window opening. They were told that they should close the window after they were finished with it so that they could continue doing the activities, but in case they forgot and left the window open in the background, a script was included to make sure that the window closed after 60 seconds, although learners were not told this. There were very few instances of the window closing automatically, but these were omitted from the study as it was considered unlikely that the learner looked at the description continuously

for this period of time. Learners were observed during the reading comprehension activities to see how they used the annotations. All of the computers in the computer laboratory that were used included headphones that allowed learners to listen to the audio, and all of the learners chose to wear them while reading. The system also kept a record of when learners accessed the audio of the words.

The reading comprehension activities only made up 30 minutes of the overall class, and the remaining 60 minutes were spent on teaching academic writing. Learners were told that they would have to complete vocabulary activities outside of class time based on words that they clicked, but, as described above, even if they did not click on any words, they would still be required to complete vocabulary activities based on words in the passage they read, so they were encouraged to click on words they wanted to know without fear of increasing their workload. Because the vocabulary component of the system adapted to learner responses, if learners scored correctly in the activities, they took a far shorter amount of time to complete than if they got questions wrong. The learners were told that they must complete the activities before the following class, or they would not be able to go on to the next reading passage in class. While there was no penalty for not completing the activities before the class, for the most part learners were very diligent in making sure that they completed what was required of them. Although learners were given the option of using either the desktop computer or their mobile phones for the vocabulary activities, as was seen in the first semester, the vast majority preferred to do the activities on desktop computers. The post-test was administered in the last class of the semester, using the same site as the pre-test. The same items were included in both the pre- and post-test, but the order was randomised for each student, and was unlikely to have been the same in each test.

## Results

The data were analysed to determine what the patterns of the learners were in looking up words using the reading component of the system. The look-up patterns were compared against the pre-test data, and revealed that most of the words that were looked up were deemed as being “unknown” according to the pre-test (i.e., that they indicated either that they didn’t know it or got it wrong in the multiple choice question). Figure 1 shows that of the words that both appeared in the pre-test and were looked up by learners, over 80% of them were deemed to be “unknown” while the remainder were words that were considered as being “known.”

There were also a number of words that appeared in the pre-test but were not looked up by the learners during the reading comprehension activities. As is shown in Figure 2, over two-thirds of the words that were not looked up while completing the reading comprehension activities were classed as being “known” with the remainder being “unknown.”

In addition to the words that were marked as “known” and “unknown,” information regarding the amount of time that was spent using the annotation window was also recorded, as can be seen in Table 1. While there is not a great difference, the table shows that learners were likely to spend slightly longer looking at the annotation window of words that were regarded as being “unknown” compared to those that were considered to be “known.” The fact that the mean time was relatively long (over ten seconds in both cases) provides some indication that the learners did spend some time looking at the information written in the annotation window. There was, however, a very large standard deviation in



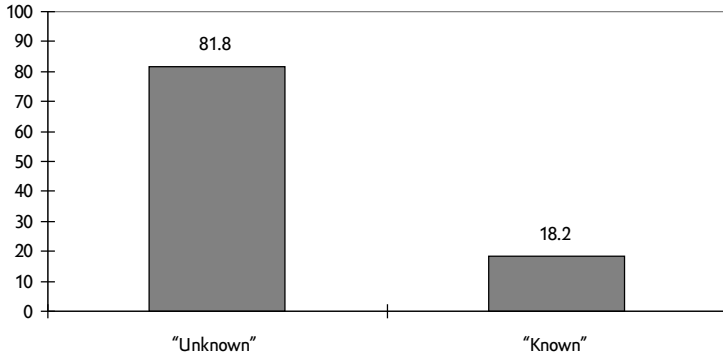


Figure 1. Percentage of words looked up that appeared in the pre-test

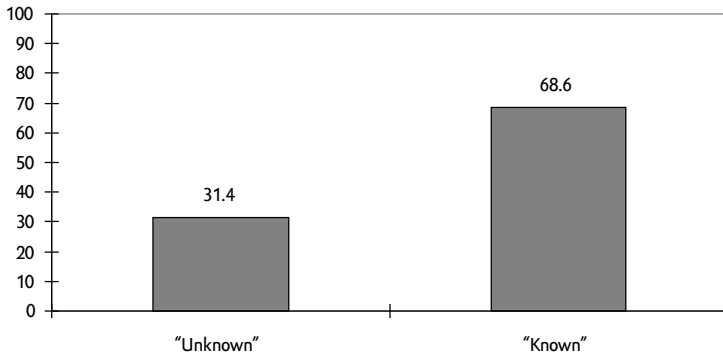


Figure 2. Percentage of words that appeared in the pre-test but were not looked up

both cases, which shows that there was a lot of variation in the amount of time spent looking at the descriptions, depending on both the word and the learner.

Table 1: Time spent by learners on reading annotation window

Category	M (Seconds)	SD
"Unknown"	12.4	5.89
"Known"	10.1	5.48

Data were also collected in order to determine the feasibility of the system as a means of constructing a profile of learner knowledge. Table 2, based on all data collected throughout the project, presents an overview of the "known" vocabulary of all the learners ( $N=43$ ) mapped against the JACET 8000 list. The "1000" here refers to the most frequently occurring **267**

one thousand words, “2000” to the second thousand, all the way through to the final 8000 (the least frequently occurring). The word “Other” was refers to those items which were included in the reading passages but were not in the JACET 8000 list. While there are again quite large standard deviations for each category, the table shows that in general terms, the number of “known” vocabulary in categories of higher frequency are significantly higher than those in the lower frequency categories.

Table 2: Overview of “known” vocabulary according to clicking patterns (N=43)

JACET level	M	SD
1000	82.1	9.54
2000	75.6	9.65
3000	63.4	10.98
4000	66.7	9.76
5000	48.9	11.91
6000	43.5	12.18
7000	45.1	16.13
8000	28.2	16.52
Other	17.9	16.28

In addition to the overall data, information pertaining to a single student was also mapped to get a picture for a single student, as can be seen in Table 3 and again graphically in Figure 3. Student 21 was selected as the profile was in a very neat pattern according to the frequency categories. While there were other students who exhibited similar patterns, there were some who showed quite a large variation compared to what might be expected if the frequency categories are to be used as a guide.

Table 3: Sample of “known” vocabulary according to clicking patterns for Student 21

JACET level	M
1000	83.3
2000	75.1
3000	56.8
4000	60.3
5000	53.7
6000	41.6
7000	44.4
8000	31.5
Other	17.8

As is evident in Figure 3, this learner shows a higher number of words classed as “known” in the higher frequency words (i.e., the first 1000 to 2000 words) compared to those words that are much lower in frequency (8000 or Other). This does provide some indication that frequency lists can give some indication of what words learners may be expected to know, but the large variation shows that caution is needed in making generalisations.

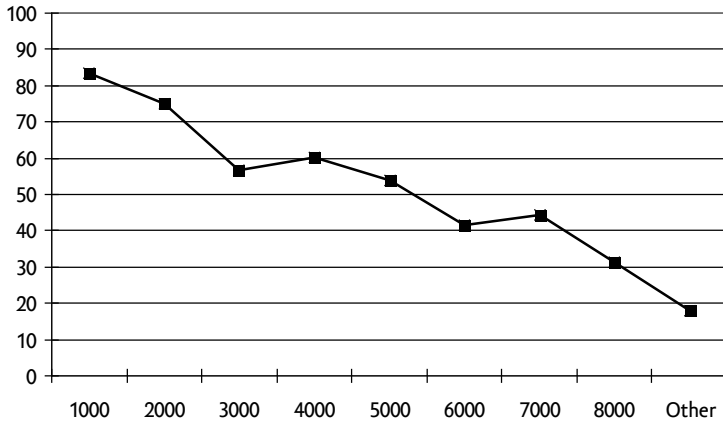


Figure 3. Sample graphical representation of "known" vocabulary for Student 21

Table 4: Results of post-test correlated with "Known" and "Unknown" words

Category	Clicked (M/SD)		Not Clicked (M/SD)	
	M	SD	M	SD
"Unknown"	65.2%	6.53%	14.2%	11.96%
"Known"	93.4%	7.39%	89.6%	5.48%

Finally, the results of the post-tests were correlated with the clicking patterns and the pre-test. The words that were deemed as "unknown" in the pre-test (i.e., those that were incorrect or claimed by the student to be unknown) were correlated against the post-test depending on whether they were clicked or not, as has been outlined in Table 4. The table shows that words that were clicked on were far more likely to be acquired than those that were not, with nearly two-thirds of the words being looked up being correct in the post-test compared with 14.2% for those words that were not looked up. There was also a difference in the words that were considered as "known" in the pre-test, with 93.4% of words scored correctly in the post-test when they were looked up, compared to 89.6% for when they were not. It is of interest that the majority of learners did complete the online activities based on the words looked up outside of class, and this is likely to have contributed to the high score for the post-test for the "unknown" words.

## Discussion

The purpose of the current study was to investigate whether a system that kept logs of words that learners clicked on while engaged in reading activities would offer a feasible means through which vocabulary activities that were tailored to learners' individual needs could be created. Learners were provided with reading comprehension activities which 269

included hyperlinks to annotations that gave L1 and L2 textual information and an audio recording of the word. Based on the words that were looked up, the learners were given activities which targeted the words that they looked up while completing the reading comprehension activities. The factors that may have contributed to the ways in which the learners looked up new words, along with some of the main issues surrounding the construction of the learner profiles are outlined in the following subsections.

### *Patterns of looking up words*

The study showed that there were many words that the pre-test indicated that learners did not know but were not looked up during the reading activities. There are several possibilities for this. Firstly, it is conceivable that learners were able to guess the meanings of the words from the context in which they appeared, and as such there was no need to look the items up in order to complete the activities. Secondly, it is possible that learners had learned the word elsewhere since the pre-test, perhaps as a result of seeing it in the pre-test, but also possibly through other sources, since the learners were enrolled in other English language classes going on at the same time. Another possibility is simply that the pre-test was not an accurate reflection of learners' vocabulary knowledge. If learners were not confident with their knowledge of a word that appeared in the pre-test, they may have marked it as (c) "I've seen it before but I don't know the meaning," when in fact they did have a good idea of its meaning. In the current study, a word was considered as being "Unknown" if it was marked as (c) or (d) regardless of whether they got it right in the multiple choice question, as it was thought that learners may simply guess it correctly, so their own evaluation of their knowledge was given preference over the score (unlike when a word was marked as either (a) or (b) but incorrect on the multiple choice). Thus, in this way, the pre-test simply may have been too insensitive to learners' vocabulary knowledge, resulting in discrepancies.

The results also suggested that although learners clicked predominantly on the words that were considered as "unknown" according to a pre-test, there were also several instances where words that were deemed as "known" were also looked up. While of course the same potential problems regarding the pre-test may have applied, there are also other possibilities that might be considered. One of the most likely is that the learners wanted to confirm the meaning of a word, despite the fact that they knew it. The word may have appeared in a context which they were unfamiliar with which could have prompted learners to check to make sure that they did understand it. Related to this is the fact that as there were audio recordings provided of each of the words, it is possible that learners wanted to hear how a word was pronounced, either to remind them of the meaning, or simply because they were curious about how it was pronounced. In either case, the annotation may have been a source of information that the learners found useful, even for words that they already knew.

One possible advantage of learning through the type of system described in the current study is that it has potential to reduce the immediate load on learners to try to learn or keep lists of unknown vocabulary during the reading process. As De Ridder (2002) argues, if time pressures are too high, learners are more likely to focus only on the meaning of the passage and direct less attention to vocabulary, which could have a detrimental effect on acquisition. In the current study, the fact that activities were given to learners based on the words that they looked up during reading regardless of how long they spent reading the annotations could be a means of reducing pressure on learners to try to remember the word at the time. Because learners know that they will have opportunities to try to learn the word

in their own time, it means that learners may choose to click on the word to quickly get an idea of the meaning in order to complete the reading comprehension activities without the added pressure of giving too much attention to remembering the words. An implication of this is that the annotation itself perhaps does not need to contain as much information as has been included in previous studies where the annotation was the primary source through which learners got information about the words they were looking up (see Yeh & Wang, 2003). However, given the fact that multimodal information about new words can enhance acquisition as predicted by the Dual Coding theory (see Yoshii, 2006, for a discussion), ensuring that the vocabulary activities provide varied information using different modes may play a role in promoting acquisition.

The survey logs indicated that, on average, the learners spent over ten seconds looking at the vocabulary annotations. This was somewhat longer than was first expected, and does provide some indication that learners took some care in going through the information that appeared in the annotations, which, as described above, included an L1 meaning, a definition and example sentence in the L2, and an audio recording of the word. Access to the audio was surprisingly small, with the vast majority of the students appearing to access only the Japanese translation of the words, an outcome which bears similarities to the results provided by Davis and Lyman-Hager (1997), who found that learners typically limited their consultation to word definitions, even when other information was available.

### *Recycling vocabulary*

Through providing reading activities for learners, it is important to ensure that there is a sufficiently wide range of vocabulary covered. Included in this is the need to allow learners to be exposed to lexical items a number of times in varying contexts. One way of doing this could be to introduce graded readers, which have attracted attention from researchers for many years (e.g., Tudor & Hafiz, 1989). Graded readers are generally short stories or novels which are limited in their vocabulary and syntax depending on level, and if used systematically can provide learners with vocabulary of increasingly difficult levels. Using these different levels can make it possible to recycle vocabulary items so that learners can get multiple exposures to them. Waring and Takaki (2003) showed that less than 20% of words encountered in a single graded reader were retained “well” but that learners did retain some memory of as many as 60% of them. However, these vocabulary gains were gradually lost when the learners were not further exposed to the words in other contexts. If learners are exposed to words in multiple reading passages, such as graded readers or other types of materials, then there is a greater chance that learners will retain these words for longer.

In the current study, there was no systematic coordinating of the ten different reading passages with regard to vocabulary (or syntax). When selecting materials, learners may have a better chance of learning new vocabulary they encounter if they are exposed to the word in more than one of the reading passages. Of course, as learners are able to do practice activities which would be expected to help them to learn the words they clicked on during reading, they were not provided with opportunities to encounter the word again after they had engaged in vocabulary activities about it. It is possible that learners may be motivated if they encounter words after having spent some time learning it, as they may feel that they are able to apply the knowledge that they learnt to a practical situation, this being reading for meaning.

## *Construction of learner profiles*

The study suggested that it is possible to create some kind of indicative profile of learners' vocabulary knowledge. It should be pointed out, however, that this profile has a number of limitations. The biggest problem with the profile is that it was constructed entirely based on the words that the learner clicked on during the reading comprehension activities. That is to say, it is based on the words that the learners actually click, and fails to take into consideration words that they clicked that they already knew, or words that they didn't know and chose not to click on. In addition, a profile that is based on a frequency list can only be as accurate as the list on which it is based, so if there are problems with the list itself, the profile also loses its value. While the **JACET 8000** list has received some support in Japan (e.g., Mizumoto, 2004), there are other lists have been more widely used such as West's General Service List and the University Word List (see Ghadirian, 2002, for a discussion), and may provide a more accurate picture. Another potential shortcoming of the learner profile is that it is in essence only valid for a single semester at best. Learners' vocabulary knowledge is naturally going to be highly changeable and dynamic, and as such a profile can only provide a snapshot of the words that the learner had clicked on during the reading over a specified period.

This does have potential advantages as well, however. The fact that the profiles for the most part fit the frequency categories of the **JACET 8000** list is encouraging, although obviously the list itself is not highly comprehensive and the number of words that learners clicked on was relatively small. Despite this, however, if such profiles are constructed on a semester-by-semester basis, they do have the potential to motivate students where they can see their own improvement in graphical form. If teachers are able to locate reading materials and/or vocabulary activities that fit the individual levels of the lists that the profiles are measured against, it may make it easier to target learners' individual needs. Learners themselves may not have a real idea of where they stand with regard to their own vocabulary level or what materials are appropriate for them, so providing this type of guide can eliminate much of the guesswork.

## *Limitations*

There were a number of limitations associated with the current study, some of which were expected at the outset of the study, and others that became apparent as the study progressed. Of the expected problems, one of the biggest was, as stated above, that the profiles were constructed based on words that were deemed as "known" or "unknown" by whether or not the learners clicked on them. While it was not anticipated that learners would click on all of the words that they did not know, it was still hoped that the system would provide at the very least a rough indication of learners' vocabulary knowledge. Another expected limitation was that due to the exploratory nature of the study, the reading passages were quite short and the range of vocabulary was very small. There would be value in expanding the study to include passages of a more substantial nature, thereby increasing learner exposure to vocabulary, both in terms of the number of words, and the frequency of encountering the words.

## Conclusions and further research

Learning vocabulary is a critical but time-consuming aspect of learning a second language. While reading can provide a means through which learners can acquire new words with a fuller comprehension of the range of meanings of the words and how they are used, reading only is a time-consuming process which can be enhanced if supplementary activities targeting specific vocabulary items are provided. One of the main problems with this approach is, however, that without specifically looking at what the learners actually know, there is a need for guesswork in choosing the content of these supplementary activities. The current study was intended to be exploratory in nature, and showed that it is possible to provide vocabulary activities tailored to learners based on the words that they click on during reading passages, but it has raised a number of other issues which require further investigation. The study did indicate that looking up items that were not known by the learners was likely to lead to their acquisition, and this high figure is likely due to the fact that many learners carried out activities based on these unknown words. There would be value in determining whether there would be gains even if the learners did not complete the activities, and if so, the effect of the number of times the words was clicked along with the types of annotations used.

In conclusion, learning vocabulary through reading in **CALL** contexts is an area with a great deal of promise, but one where research is still very much needed. Technology has the potential to give us insights into what learners do during reading that could be used for their benefit in acquiring vocabulary. It is up to teachers and **CALL** practitioners, then, to think through the ways in which the technologies can look into the learners' world, and to bring them what they need.

## References

- Coady, J. (1997). L2 vocabulary acquisition through extensive reading. In J. Coady & T. Huckin (Eds), *Second language vocabulary acquisition* (pp. 225-237). Cambridge: Cambridge University Press.
- Davis, J. N., & Lyman-Hager, M. A. (1997). Computers and L2 reading: Student performance, student attitudes. *Foreign Language Annals*, 30 (1), 58-72.
- Ghadirian, S. (2002). Providing controlled exposure to target vocabulary through the screening and arranging of texts. *Language Learning & Technology*, 6 (1), 147-164.
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge: Cambridge University Press.
- Healy, D. (1999). Classroom practice: Communicative skill-building tasks in **CALL** environments. In Egbert, J. and Hanson-Smith, E. (Eds), *CALL environments: Research, practice and critical issues* (pp. 116-136). Alexandria, VI: **TESOL**.
- Kang, Y.-S., & Maciejewski, A. (2000). A student model of technical Japanese reading proficiency for an intelligent tutoring system. *CALICO Journal*, 18 (1), 9-40.
- Knight, S. (1994). Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities. *The Modern Language Journal*, 78, 285-299.
- Koyama, T., & Takeuchi, O. (2007). Does look-up frequency help reading comprehension of **EFL** learners? Two empirical studies of electronic dictionaries. *CALICO Journal*, 25 (1), 110-125.

- Mizumoto, A. (2000). **JACET8000-to TOEIC-no sōkansei: TOEIC mini kōpasu-o tsukatte-no sūryōteki kentō** [A correlation between **JACET 8000** and **TOEIC**: A quantitative analysis using a **TOEIC** mini-corpus.] In Basic Language Revision Commission (Ed.), **JACET8000 katsuyō jireishū** [A collection of case studies using the **JACET 8000** list] (pp. 60–61). Tokyo: **JACET**.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2008). *Teaching vocabulary: Strategies and techniques*. Boston, MA: Heinle, Cengage Learning.
- Paribakht, T. S., & Wesche, M. (1997) Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady & T. Huckin (Eds), *Second language vocabulary acquisition* (pp. 174–200). Cambridge: Cambridge University Press.
- Peters, E. (2007). Manipulating L2 learners' online dictionary use and its effect on L2 word retention. *Language Learning & Technology*, 11 (2), 36–58.
- Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, 18 (1), 1–28.
- Stockwell, G. (2007). A review of technology choice for teaching language skills in the **CALL** literature. *ReCALL*, 19 (2), 105–120.
- Stockwell, G. (2010). Using mobile phones for vocabulary activities: Examining the effect of the platform. *Language Learning & Technology*, 14 (2), 95–110.
- Tudor, I., & Hafiz, F. (1989). Extensive reading as a means of input to L2 learning. *Journal of Research in Reading*, 12, 164–78.
- Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from a graded reader? *Reading in a Foreign Language*, 15 (2), 130–163.
- Waring, R. (2009). The inescapable case for extensive reading. In A. Cirocki (Ed.), *Extensive reading in English language teaching* (pp. 93–111). Munich: Lincom.
- Yeh, Y., & Wang, C.-W. (2003). Effects of multimedia vocabulary annotations and learning styles on vocabulary learning. *CALICO Journal*, 21 (1), 131–144.
- Yoshii, M. (2006). L1 and L2 glosses: Their effects on incidental vocabulary learning. *Language Learning & Technology*, 10 (3), 85–101.
- Yoshii, M., & Flaitz, J. (2002). Second language incidental vocabulary retention: The effect of picture and annotation types. *CALICO Journal*, 20 (1), 33–58.
- Zimmerman, C. B. (1997) Historical trends in second language vocabulary instruction. In J. Coady & T. Huckin (Eds), *Second language vocabulary acquisition* (pp. 5–19). Cambridge: Cambridge University Press.

## Author biodata

Glenn Stockwell (PhD, University of Queensland) is Professor in Applied Linguistics at Waseda University, Tokyo, Japan. His research interests include investigating the impact of the context on technology uptake and usage, **CALL** theory, and the role of training on both teachers and learners in the language learning process.