

【原著】

Using Spaced Repetition Software with a TOEIC Wordlist: Preliminary Results

Arthur Rutson-Griffiths and Jack Bower

間隔反復ソフトを用いた TOEIC 単語の学習に関する予備調査について

Arthur Rutson-Griffiths and Jack Bower

Abstract

This paper presents the preliminary results from the second phase of a study in which participants are studying a TOEIC wordlist using spaced repetition software (SRS). The overall aim of this study is to investigate whether increased usage of the SRS leads to statistically significant gains in TOEIC overall and reading scores. The aim of the second phase of the study is to replicate the first stage of the study, which used the SRS Anki and yielded disappointing results, with the more feature-rich and user-friendly SRS Cooori, and to compare participant attitudes to them. Participants were given a 640 item TOEIC wordlist, and access to this list in Cooori, and asked to study it over one academic year. Participants are tested on their knowledge of the items via 20 item multiple-choice quizzes every two weeks in order to encourage regular study. TOEIC scores at the start of the project will be compared to those at the end, and attitudes are obtained through administration of a survey. This working paper presents preliminary attitudinal data, which suggests participants favor Cooori over Anki, and preliminary usage data, which shows greatly increased usage in the second phase of the study compared to the first. As final TOEIC scores are not yet available, SRS usage is correlated with quiz scores using a non-parametric test. The results suggest that TOEIC scores predict quiz scores, and that while SRS usage does not yet strongly predict quiz scores, usage of the 'quiz' feature of Cooori does predict scores. Preliminary results show increased SRS usage and more positive attitudes when compared to the first phase, and encouraging results from the correlations between Cooori usage and quiz scores.

Context

The Global Communication Department at Hiroshima Bunkyo Women's University

Hiroshima Bunkyo Women's University (HBWU) is a small, private women's university in the outskirts of Hiroshima, Japan. It currently has approximately 1400 students and five departments. In 2010 a new Global Communication Department (GCD) was established at the university. The GCD aims to produce graduates with the skills necessary to compete in an increasingly globalized economy in the 21st century. Perhaps the most important of these GCD graduate skills is seen to be English language proficiency, as the English language remains the dominant language of

international business, trade and tourism (Crystal, 2003). In an effort to obtain an objective measure of GCD student progress in English language proficiency for the business sphere, GCD uses the Test of English for International Communication (TOEIC). A brief description of the TOEIC and reasons for choosing it as the primary measure of GCD student English proficiency are given in the following sections. This is followed by a rationale for this study supported by a brief literature review.

The Test of English For International Communication (TOEIC)

The TOEIC is a test of business English, which is well known in Japan. It has a 45-minute listening section with 100 items, and a 75-minute reading section with 100 items. Each correct response is worth five points, and candidates can achieve a maximum score of 495 on each section. The TOEIC was first administered in Japan in 1979 and has risen to become one of the most widely used and accepted tests of business English proficiency, both in Japan and abroad (McCrostie, 2009). The TOEIC is taken by more than five million people each year (Educational Testing Service, 2007), and was taken annually by more than 1.7 million people in Japan in the late 2000s (Kamijo, 2010). In Japan minimum TOEIC scores are often stated in job advertisements, and TOEIC scores are sometimes set as requirements for promotion within companies or for postings overseas (Institute of International Business Communication, 2013). Also, there is an increasing trend for Japanese high schools and universities to set a target TOEIC score for their graduates (Takahashi, 2012), with several universities also using the TOEIC for English course placement purposes (Yazawa, Takeuchi & Mochizuki, 2009).

An Imperative to Raise GC Department TOEIC Scores

In an attempt to set a high standard for English language proficiency for graduates from the new GCD, HBWU set an ambitious minimum requirement for graduates of TOEIC 600. However, when the first cohort of GCD students graduated in early 2014, unfortunately very few students had achieved this standard. Indeed only three students out of 18 graduates or 17% of the cohort had achieved this target.

We believe that it is important that the target of a TOEIC score of over 600 for GCD graduates be met for three reasons. Firstly, for students it will enhance their chances of finding a good job in which they can utilize their English skills upon graduation. Secondly, for teachers and administrators in the GCD, and in the Bunkyo English Communication Center (BECC), which handles many English language classes for GCD students, it will provide accountability for the efficacy of the English language program. Finally, for the university higher TOEIC scores will be useful for public relations and advertising purposes.

Washback from the TOEIC 600 Target on the GC Department Curricula

'Washback' refers to the influence of testing on classroom instruction, and assumes that stakeholders, including teachers, learners and administrators "do things they would not necessarily otherwise do because of the test" (Alderson & Wall, 1993, p. 5), both in and out of the classroom. The setting of a TOEIC score of 600 target for GC graduates has caused washback on

the GCD curriculum. Three of these washback effects are apparent to the authors of this paper. The first is the creation of GCD courses that teach TOEIC test taking strategies and give test practice, which were set up by Japanese GCD faculty. The second is the creation of cooperative learning groups by foreign teachers of GCD students in the BECC, in which first and third year students collaborate to study for the TOEIC. The third is having GCD students study a TOEIC vocabulary list using Spaced Repetition Software (SRS). This paper describes this third project and gives some preliminary results of its second phase.

Literature Review

Genesis of the SRS TOEIC Word List Projects

In an effort to boost GCD students' TOEIC scores, the authors of this paper had the idea of having students study a TOEIC word list using SRS. The rationale for this is given in the following sections, each supported by a brief literature review. Firstly the reasons for introducing a TOEIC word list are given, followed by a brief description of the word list chosen. Next an outline of the advantages of SRS over regular flashcard memorization is given, along with a short review of the evidence for the importance of vocabulary as a component of reading and listening comprehension. Finally, previous studies on SRS for language learning are briefly reviewed.

Word Lists

Word lists for language learning are based on computer corpus analysis, which scans large amounts of text to identify the most frequent words used. The resulting lists allow language learners to focus their study on the most useful words for communication and comprehension without wasting time learning low frequency words, which they are unlikely to encounter often (Maera, 1995; Nation & Waring, 1997). We believed that by introducing a word list targeted at the TOEIC we could help learners study more effectively for the test by prioritizing the study of those words that are most likely to appear in the test.

A TOEIC Word List

The TOEIC vocabulary list chosen for this study was one constructed by Chujo and Genung (2005) based on a corpus of retired TOEIC tests, practice TOEIC tests and common Japanese high school textbooks. They analyzed this corpus to find words which commonly appear in the TOEIC but are unlikely to be known by the typical Japanese high school graduate, thus creating a list of 640 words that would boost the word coverage of Japanese university students to the 95% knowledge of words, or 'threshold' (Nation, 2001; Read, 2000) needed to comprehend a typical TOEIC reading section text. Chujo and Genung's TOEIC word list was thought to be the most suitable for GCD students, because it is targeted specifically at Japanese university students, and it also seemed to have a sound methodology behind its construction. In addition, no previous studies have investigated the effects of systematic study of this list on TOEIC scores.

The Relationship Between L2 Vocabulary and L2 Reading and Listening Comprehension

A number of studies have shown a significant relationship between second language (L2)

vocabulary knowledge and L2 reading comprehension (Laufer, 1992; Maera & Jones, 1989; Brisbois, 1995). Other studies have also found a clear relationship between L2 vocabulary and L2 listening comprehension (Chiang, 2007; Mehrpour & Rahimi, 2010). Given that the relationship between L2 vocabulary and L2 listening and reading comprehension is clearly supported by research, the authors considered it worthwhile to investigate if raising student knowledge of words that are likely to appear on the TOEIC could have the effect of increasing student gains on the TOEIC test. However, the question of the best way to teach this vocabulary list to students remained.

Choosing a Vocabulary Learning Strategy

There is evidence that deliberate focused study of words leads to faster retention of new words than is acquired through 'incidental learning' or learning through simply reading and listening to lots of level-appropriate L2 material (see Nation, 1982 for an early review of such studies). Nation and Waring (1997) recommended two strategies for helping learners to memorize the meaning of important L2 words. The first strategy is mnemonics and the second is flashcards. In the mnemonics strategy learners memorize an L2 word by linking it to a mental image, sound or phrase. For the flashcard strategy learners memorize a word by writing the L2 word or phrase on one side of a flashcard, and writing the first language (L1) equivalent on the other side. Learners then test themselves by attempting to recall the L2 meaning of the L1 word on the other side of the card, and vice versa.

For this study the mnemonic technique was thought to be impractical, as either the researchers or the learners would have to come up with mnemonics for the 640 words on the TOEIC word list. In addition there is some evidence that the mnemonic strategy does not lead to long-term retention of words for Japanese learners of English (Iodine-Shirai, 2007). The researchers decided to utilize a version of the flashcard strategy for this study, which also employs spaced repetition. The advantages of spaced repetition and particularly Spaced Repetition Software SRS are briefly outlined in the next section.

The Advantages of Spaced Repetition Software

The spacing effect refers to the phenomenon whereby "for a given amount of study time, spaced presentations yield substantially better learning than do massed presentations", and has been understood for more than a century (Dempster, 1988, p. 627). Spacing allows for more efficient learning as long-term memory retention is improved, and hence less time is required to relearn forgotten material. Its effectiveness has been demonstrated with vocabulary acquisition (Bahrlick et al., 1993), and moreover the spacing effect has been shown to also increase the efficiency of reviewing previously learned material by allowing an increasingly large amount of time between subsequent reviews of retained knowledge (Cepeda et al., 2006). Thus the spacing effect can be utilized in two ways: to make both the initial learning and subsequent review of material more efficient.

Advances in computing power have allowed software developers to create computer programs that

utilize the spacing effect to make learning and review more efficient. One of the earliest examples of this was the first spaced repetition software system Supermemo (Wozniak, n.d.), which demonstrated the rich potential of utilizing the spacing effect in computer software. Supermemo paved the way for, and was eventually superseded by, a number of feature-rich and free to use SRSs, such as Anki (Elmes, n.d.) and Mnemosyne (2014).

SRSs utilize the spacing effect in spacing the presentation and review of electronic flashcards. The user either creates their own 'decks' of cards or downloads a pre-made set, which contain a problem or question on one side and the answer on the other. The user views the card and attempts to recall the answer to the question. She then presses a button to reveal the answer on the reverse side of the card and rates herself on how well she was able to recall the knowledge from a choice of three or four levels. Options typically include an incorrect option and a few differing levels of how easily the user was able to produce the answer, with the next review time ascending with the ease of recall. Review times increase with subsequent successful recall, from a few minutes up to many years. Insufficient recall schedules the card to be presented again more frequently.

Previous Studies Using SRS for L2 Vocabulary Learning and Justification for the Present Studies

Researchers advocate the use of SRS as a superior way to learn L2 vocabulary lists (Goodwin-Jones, 2010), and some previous research has shown the advantages of SRS over more traditional forms of language learning. For example Miles and Kwon (2008) showed that SRS had superior learning outcomes to traditional classroom vocabulary learning, and Hirschel and Fritz (2013) showed that using an SRS resulted in slightly better long-term vocabulary gains than using a vocabulary notebook. Ko and Gorenson (2014) also found that use of an SRS resulted in statistically significant improvements in receptive and productive vocabulary for Korean students after ten weeks of study; however, a methodological weakness was that their study lacked a control group.

Agawa, Black and Heriman (2011) examined the effect of having 76 second year Japanese university students study a TOEIC word list using the SRS *Word Engine* for around 30 minutes per week for eight weeks. Effects of the SRS study were examined by correlating SRS study time and TOEIC scores, and by comparing participants' TOEIC gains to previous years' cohorts over the same period. This study found no statistically significant relationship between amount of SRS study and TOEIC scores, nor did it demonstrate a statistically significant TOEIC score gain over previous cohorts. However, shortcomings of this study were the short length of the intervention, and the small amount of overall study time using the SRS.

In spite of the growing research body into the use of SRS for L2 vocabulary learning, we know of no research into the effects of using an SRS to practice a test-specific vocabulary list on improving scores for a standardized test over a longer term period of one academic year. Nor do we know of any research into learner attitudes to such a longer-term pedagogical intervention. The two studies reported in this paper aim to address these gaps in the research.

A Two Phase Research Project

This research project consists of two major phases. Phase one took place in 2012 and used the SRS *Anki*. Based on shortcomings of the Anki SRS revealed in the first phase, a second phase of the project is underway in 2014 utilizing the SRS *Cooori*. In the next two sections the methodology and results of the Anki project are briefly summarized, follow by a description of the phase two project utilizing *Cooori*.

The First Study Using the SRS Anki

In 2012 all GCD first and second year students were required to study the TOEIC word list with the SRS *Anki* for one academic year. There were 24 first year participants with a starting mean TOEIC score of 314 and a standard deviation of 69, and 28 second year participants with a starting mean TOEIC score of 389 and a standard deviation of 82. Participants were given a paper copy of the TOEIC word list and also were given an online Anki account with a deck of the TOEIC word list installed. During semesters 1 and 2 participants were tested on an incrementally increasing range of the word list with vocabulary quizzes every few weeks. Participants were expected to study an additional 50 words from the list every two weeks during semesters 1 and 2.

To assess the motivational effects of studying the TOEIC word list with Anki a survey was given to the participants at the end of the study. To assess the effects of studying a TOEIC word list with Anki on students' TOEIC scores the difference between pre and a post TOEIC scores were compared to the previous year's cohort's annual TOEIC gain. In addition, statistical tests were run to check for correlations between the amount of Anki study of the TOEIC word list and annual TOEIC gains.

Results of the motivational survey showed that students did not understand the benefits of SRS over other vocabulary learning strategies, with many students preferring to study the paper copy of the word list. Students also requested a pronunciation feature, and example sentences for the SRS flash cards. Furthermore, no statistically significant differences were found between the 2011 first and second year GCD cohorts' annual TOEIC gains, and the gains of the 2012 first and second year GCD cohorts who studied for a year with Anki. Also, no correlation was found between amount of Anki study and annual TOEIC score gains. The main reason for the lack of an apparent effect on TOEIC scores was judged to be insufficient amount of study time, and lack of regular study on the SRS (see Bower & Rutson Griffiths, 2014 for a complete description of phase 1 of the project).

Methodology

The Second Study Using the SRS Cooori

The second phase of the study was designed in order to replicate the first phase and improve upon its problematic features. The participants are 39 first year students with an average starting TOEIC score of 317.7 and a standard deviation of 91.3, and 33 second year students with an average TOEIC starting score of 415.9 and a standard deviation of 96.47. All students were

required to study the same TOEIC word list, which was provided in an SRS for students to use. Additionally, students in the second phase all have their own iPad minis, given to them upon entering the university, that they can use to study in class or at home (see Runnels & Rutson-Griffiths, 2013). Students were also given the full list of words, translations and example sentences in the form of a PDF downloaded to their iPad minis. As in the first phase of the study, participants are quizzed every few weeks on their knowledge of increasing numbers of the word list. Students were also surveyed for their attitudes to the TOEIC list and using an SRS; the survey was delivered after 30 weeks in the first phase and after 6 weeks in the present study, and will be delivered again after 30 weeks. The surveys were delivered online, and contained twelve common items, ten of which were statements rated on a 6-point likert type scale, and two of which were open-ended responses.

Improvements

Based upon the first phase study results, and feedback obtained from the phase one motivational survey, the following areas were identified as possible areas of improvement: adding example sentences and audio for each word, improving the user interface, better explaining to students the benefits of using an SRS, and encouraging them to study more. As some of these issues, such as user interface, could not be solved by continuing to use Anki, the present study was carried out using the online SRS Coori (2014) in collaboration with its creators. Using Coori allowed the shortcomings of the exploratory study to be improved in the following ways.

Example Sentences and Audio

For this study, at least one original example sentence was written for each item. A recording of each sentence, read by a native English speaker, and a Japanese translation was made for each sentence. Audio was also recorded for all individual vocabulary items. In contrast to a card in Anki in the first phase of the project, which presented students with only an item and its translation, Coori presents students with an item and its translation, an example sentence and its translation, and audio for both the item and the example sentence (see Figure 1). These sentences and audio are also taken advantage of in Coori by providing students with listening questions and fill-in-the-gap questions in addition to flashcards.

Improved User Interface

As well as containing more information for students when studying, the interface for Coori is more user-friendly (see Figure 1). When reviewing cards, the number of buttons available for students to grade how well they remembered an item are reduced from four in Anki (very easy, easy, good, again) to a more readily understandable two in Coori (I knew it, I didn't know it). As all students have their own iPad minis for use in class, students are able to open Coori on their devices and start studying immediately, both in and out of class, which is hoped to encourage more usage (see below).

Better Explanations for Students

Many students in the exploratory study reported that they did not understand the benefits of

using an SRS (see Figure 2). In the current study, representatives from Cooori visited the university to give all students an induction in using Cooori on April 16th 2014. Cooori also has a number of explanatory features on its website, such as FAQs and videos, that students can access at any time.

Encouraging Students to Study More

As student usage on the SRS was disappointing in the first study (see Table 1), efforts to increase student usage were increased. In addition to classroom time being given to students to study using Cooori, students receive weekly emails from Cooori encouraging and reminding them to study and were provided with the incentive of small prizes to use Cooori during the summer vacation. Cooori itself also contains a number of features that may encourage students to study more. Students are able to view their progress on a number of measures on the 'progress' page. Students are also able to use the 'quiz' function to test themselves on sets of 50 words. This feature, which is separate from the spaced-repetition 'learn' mode, is expected to be useful for students wanting to revise before taking an in-class test (see Table 4 in the results section) and may encourage familiarity and willingness to use Cooori.

Figure 1: The User Interface in Cooori's SRS Mode



Research Questions

The research questions for this study are:

1. Do cohorts using Cooori to study a TOEIC word list for one year have a statistically significant greater gain in TOEIC scores than previous cohorts using Anki?
2. Do Cooori repetitions correlate with TOEIC reading and listening score gains?
3. Do cohorts using Cooori report more positive attitudes to spaced repetition software than cohorts using Anki?
4. Do cohorts using Cooori study more regularly, and in greater amounts than cohorts using Anki?

Results and Discussion

Usage - Anki and Coori

Total usage comparing the first study and the present study are shown in Table 1. Usage for the present study covers the period for which Coori was used in the first semester of 2014 (approximately 16 weeks). Final usage data (30 weeks) for the first study is presented, as only final usage data for Anki is available. The results show that students have been doing an average of 30 extra reviews per week in Coori compared to Anki. This is a considerable number of extra reviews, but it is difficult to compare number of reviews directly due to the differing amounts of data presented on the flashcards in each SRS. Students in the first phase Anki study were presented with less information on each card, and were able to move onto the next card more quickly. There was also evidence in the first phase data that some students were not paying attention to the cards on screen, but were instead 'button mashing' and going through as many cards as quickly as they could to try and finish. For example, one student had an average answer time per card of only two seconds. As Coori contains listening and fill-in-the-gap questions, it is less likely that students will button mash and get through cards quickly without paying attention.

More significant is the average number of learning sessions per week. Students in the first phase were on average only using the SRS once a week. As Anki, and all similar spaced-repetition software, is designed to be used daily (Elmes, 2014), once a week usage is not sufficient to benefit from the spacing effect. The average weekly sessions in the present study was more encouraging, with an average of 4.08 study sessions per week being closer to the ideal daily usage.

Table 1: Summary of Usage for the First and Second Stages

SRS	Cohort	Total reviews	Average weekly reviews/student	Total review sessions	Average weekly learning sessions/student
Anki (30 weeks)	1st year (n=23)	24,826	35.98	490	0.71
	2nd year (n=27)	87,264	107.73	693	0.85
	Total (n=50)	112,090	74.73	1,183	0.79
Coori (16 weeks)	1st year (n=39)	62,269	99.79	2,383	3.82
	2nd year (n=33)	58,266	110.35	2,314	4.38
	Total (n=72)	120,535	104.63	4,697	4.08

Note: Final Anki usage data was not available for all students.

The usage data from students who used the quiz feature in Coori, not available in Anki, is presented in Table 2. Quiz answers are not a part of Coori's SRS function, and so are not included in the data presented in Table 1, but it can be seen that students are using this feature a significant amount.

Table 2: Summary of Usage for Coori's Quiz Feature

Cohort	Total quiz answers (16 weeks)	Average weekly quiz answers/student	Total quiz sessions	Average weekly quiz sessions/student
1st year (n=24)	16,313	42.48	606	1.58
2nd year (n=28)	23,642	52.77	816	1.82
Total (n=52)	39,955	48.02	1,422	1.71

Attitudes - Anki and Coori

Attitudinal data collected from the common items in the surveys distributed to students are shown in Figures 2 and 3. As data were collected at different stages of the two studies, attitudinal data for the present study is considered preliminary, and some differences in results may disappear by the time the final survey is administered after 30 weeks.

Figure 2: Comparative Survey Responses to Likert Type Items About SRS

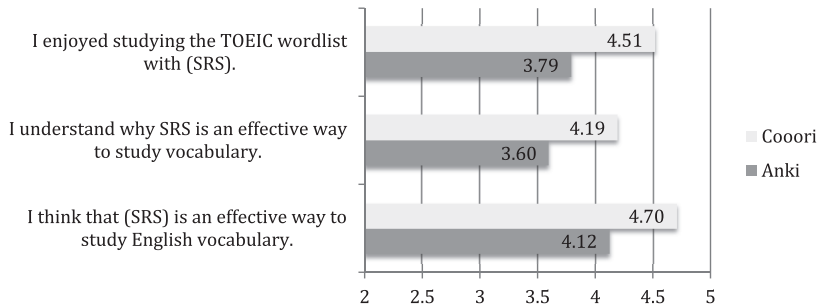
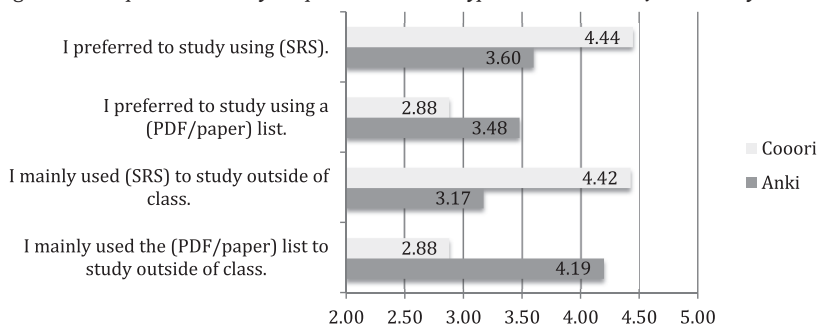


Figure 2 shows student attitudes to using SRS and their level of understanding of why it is effective. Students in the current study reported higher levels of enjoyment and understanding SRS when compared to the first phase study. Figure 3 shows preferences and self-reported behavior in regard to SRS and the paper or PDF list. Whereas students in the exploratory study showed little overall difference in preference for the SRS or paper list, students in the current study report preferring to use the SRS, Coori. Students using Coori also reported that they study outside class using the SRS rather than the PDF list, in contrast to the users of Anki, who

Figure 3: Comparative Survey Responses to Likert Type Items About Preferred Study Method



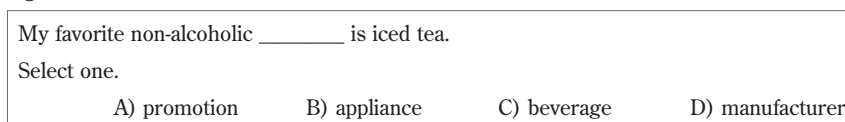
as a group reported using the paper list more than the SRS.

The Relationship Between TOEIC Word List Quiz Scores and Coori Repetitions

The following section examines the relationships between the amount of study on Coori flash cards and Coori practice quiz questions with scores on TOEIC word list quizzes administered in class. The results give a very preliminary indication of possible effects on TOEIC scores from studying a TOEIC word list with Coori.

In semester 1, 2014 all first and second year GCD students took a twenty-item quiz on the TOEIC word list approximately every two weeks. The dates of the quizzes are shown in the appendix. As this paper represents just a preliminary analysis, only the results of the final quiz of the semester (Quiz 6) are analyzed. The final TOEIC quiz of the semester consisted of ten items testing words from the list range 251-300, and ten items testing words from the list range 1-250. All of the quiz questions had the same multiple-choice format, which had quiz takers choose the best option from four word choices to complete a sentence (see Figure 4). The key and all distractors for the quiz items were taken from words 1-300 of the TOEIC word list. The quiz had a Cronbach's alpha or coefficient of internal consistency of 0.8, which is considered to be 'good' reliability and within the acceptable range of 0.7-0.9 (Tavakol & Dennick, 2007).

Figure 4: TOEIC Word Quiz 6 Item



Several of the variables to be analyzed did not meet the assumptions of normality based on Shapiro Wilk tests of normality in SPSS. Therefore to test for the relationship between TOEIC word list Quiz 6 scores and students' beginning TOEIC scores, Spearman's Rho, a non-parametric test for correlation, was run. The results of the test are shown in Table 3. Only those students who had taken the TOEIC after entering the GC department, and who also took the final TOEIC word list quiz were able to be included in this analysis.

Table 3: TOEIC Word List Quiz 6 Score Correlations with Beginning TOEIC Scores Using Spearman's Rho

	TOEIC Total	TOEIC Listening	TOEIC Reading
First Year (n=34)	r=.698** p=.000004	r=.709** p=.000003	r=.462** p=.006
Second Year (n=28)	r=.542* p=.003	r=.312 p=.106	r=.607* p=.001

*Correlation is significant at the 0.05 level (two tailed)

**Correlation is significant at the 0.01 level (two tailed)

TOEIC word list Quiz 6 scores correlated significantly with TOEIC reading and overall TOEIC scores both for first and second year students. TOEIC listening scores also correlated significantly with the TOEIC word list Quiz 6 scores for first years, but not for second years. Scores on the TOEIC Quiz 6 account for around 29% of the variance in TOEIC total scores for second year students, and nearly 49% of the variance for first year total TOEIC scores. The difference between the amount of variance accounted for is likely due to the different dates of TOEIC administration: the TOEIC was administered three months earlier for first years and six months earlier for second years. Given the clear relationship between TOEIC scores and scores on TOEIC word list Quiz 6 we would argue that if Coori study of the TOEIC word list seems to have an effect on TOEIC word list Quiz 6 scores it would bode well for showing a later effect on TOEIC reading scores after one year of study with Coori.

Next, total repetitions of Coori flashcards, and of Coori quizzes both for the whole of semester 1 before Quiz 6 and for the two weeks before Quiz 6 are given in the following table. Due to the staggering of dates for class tests for different classes, the total repetitions were calculated separately for each class and then combined. The non-parametric test Kendall's tau was used for these calculations, because it is better than Spearmans rho for data sets in which many scores have the same rank (Field, 2000). Many scores were ranked together at the bottom this data set because a majority of students did not use the quiz practice feature on Coori giving them a zero score.

Table 4: TOEIC Word List Quiz 6 Score Correlations with Coori Flash Card and Practice Quiz Repetitions using Kendall's Tau

	Preceding two week flash card reviews	Semester 1 flash card reviews	Preceding two week quiz practice reviews	Total semester 1 quiz practice reviews
First Year (n=35)	r=.242* p=.047	r=.172 p=.154	r=.292* p=.034	r=.370** p=.003
Second Year (n=32)	r=.111 p=.400	r=.152 p=.239	r=.510** p=.000199	r=.443** p=.001

*Correlation is significant at the 0.05 level (two tailed)

**Correlation is significant at the 0.01 level (two tailed)

From the above table it can be seen that the amount of Coori flash card reviews for the two weeks preceding the test did not correlate significantly with TOEIC word list Quiz 6 scores for second-year students. First-year student flash card reviews did, however correlate significantly at the .05 level for first year students and TOEIC word list Quiz 6 scores; however, the effect size was small, accounting for only 6% of variance in Quiz 6 scores. Total semester 1 flash card reviews did not correlate with TOEIC word list Quiz 6 scores for either student cohort.

However, for both first and second year groups the amount of practice on Coori practice quizzes had a clear relationship with TOEIC word list Quiz 6 scores. The relationship between Coori quiz practice and TOEIC word list Quiz 6 was statistically significant for both the two weeks preceding

the test, and for the total quiz practice for the whole semester preceding TOEIC word list Quiz 6 for both first years and second years. The preceding two weeks of Coori quiz study accounted for 8.5% of variance in Quiz 6 scores for first years and 26% of variance in Quiz 6 scores for second years. Total semester 1 Coori quiz practice reviews accounted for around 14% of variance in Quiz 6 scores for first years and around 20% of variance in Quiz 6 scores for second years.

Conclusion

This study will attempt to answer four research questions. These research questions are:

1. Do cohorts using Coori to study a TOEIC word list for one year have a statistically significant greater gain in TOEIC scores than cohorts using Anki?
2. Do Coori repetitions correlate with TOEIC reading and listening score gains?
3. Do cohorts using Coori report more positive attitudes to spaced repetition software than cohorts using Anki?
4. Do cohorts using Coori study more regularly, and in greater amounts than cohorts using Anki?

We will address the research questions in order. At this stage, we are unable to answer questions 1 and 2, as students will not take the TOEIC until the end of the study. However, the preliminary results reported for the second phase of this study are encouraging. A problem in the first phase of the study was that students did not appear to use Anki sufficiently to benefit from the spacing effect. In the second phase usage of Coori is so far much higher than Anki. Given the moderate association between Quiz 6 scores, and TOEIC overall scores and reading section scores, the significant correlation between Coori quiz usage and Quiz 6 scores seems to be a positive preliminary indicator for the end of study TOEIC results for regular users of Coori.

For question 3, although students in the first and second phases of the study were surveyed at different points (30 weeks and 6 weeks respectively), students in the second phase demonstrated more positive attitudes to the SRS than first phase students, as well as better understanding of the benefits of using SRS and higher usage of SRS compared to the paper or PDF list. Presumably if students understand the benefits of using SRS and think it is an effective way to study vocabulary, their usage will continue to be higher and there is subsequently a higher chance that their regular SRS usage will lead to TOEIC gains.

Regarding research question 4, initial results after one semester clearly show more regular study with Coori than was the case with Anki, and also higher usage overall. This may be due to a number of factors including: a more user-friendly interface than Anki, extra features offered by Coori over Anki, better explanation of the benefits of SRS in the second project phase, incentives offered to participants to study over the summer holiday period, and the ease with which participants could access Coori through iPad minis, which were provided to all participants. The results presented in this working paper are preliminary. Research questions 1 and 2 will be

answered at the end of this second phase when all participants take the TOEIC. Usage data after 30 weeks will be correlated with TOEIC total scores and reading scores. Another attitudinal survey will also be administered in order to answer question 3, and usage data examined to answer question 4.

References

- Agawa, G., Black, G., & Herriman, M. (2011). Effects of web-based vocabulary training for TOEIC. In A. Stewart (Ed.), *JALT2010 Conference Proceedings*. JALT: Tokyo.
- Alderson, J. C., & Wall, D. (1993). Does washback exist?. *Applied linguistics*, 14(2), 115–129.
- Bahrack, H. P., Bahrack, L. E., Bahrack, A. S., & Bahrack, P. E. (1993). Maintenance of foreign language vocabulary and the spacing effect. *Psychological Science*, 4(5), 316–321.
- Bower, J., & Rutson-Griffiths, A. (2014). *Effects of studying a TOEIC word list with spaced-repetition flashcard software*. Manuscript in preparation.
- Brisbois, J. E. (1995). Connections between first- and second-language reading. *Journal of Reading Behaviour*, 27(4), 565–584.
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132(3), 354.
- Chiang, A. C. (2007). The impact of vocabulary preparation on L2 listening comprehension, confidence and strategy use. *System*, 35, 534–550.
- Chujo, K., & Genung, M. (2005). Utilizing the British National Corpus to analyze TOEIC tests: The quantification of vocabulary-usage levels and the extraction of characteristically used words. *TOEIC Research Report 3*. Retrieved from http://www.toEIC.or.jp/library/toEIC_data/toEIC_en/pdf/newsletter/KiyomiChujo_E.pdf
- Coori (2014). *About Coori*. Retrieved from http://www.coori.com/en/about_coori/
- Crystal, D. (2003). *English as a global language*. Cambridge: Cambridge University Press.
- Dempster, F. N. (1988). The spacing effect. *American Psychologist*, 43, 627–634.
- Educational Testing Service. (2007). *TOEIC user guide: Listening and reading*. [Brochure]. Retrieved from: http://www.ets.org/Media/Tests/Test_of_English_for_International_Communication/TOEIC_User_Gd.pdf
- Elmes, D. (n.d.). Anki [Software]. Available from <http://ankisrs.net>
- Elmes, D. (2014). *Anki 2.0 user manual*. Retrieved from <http://ankisrs.net/docs/manual.html>
- Field, A. P. (2000). *Discovering statistics using SPSS for Windows: Advanced techniques for the beginner*. London: Sage.
- Goodwin-Jones, R. (2010). From memory palaces to spacing algorithms: Approaches to second language vocabulary learning. *Language Learning and Technology*, 14(2), 4–11. Retrieved from <http://lsa-cmsf5test.lsa.umich.edu/german/hmr/531/llt/emerging.pdf>
- Hirschel, R., & Fritz, E. (2013). Learning vocabulary: CALL program versus vocabulary notebook. *System*, 41(3), 639–653.
- Institute of International Business Communication (2013). *TOEIC newsletter*, 116. Tokyo: IIBC. Retrieved from: http://www.toEIC.or.jp/toEIC_en/publications/newsletter.html
- Iodine-Shirai, D. (2007). Using mnemonics in vocabulary tests. *The Internet TESL Journal*, 8(2). Retrieved from <http://iteslj.org/Articles/Iodine-Mnemonics.html>
- Kamijo, T. (2010). TOEIC reading comprehension: Analysis of the texts and questions, practical application, and its implications. *立命館高等教育研究*, 10, 173–185.
- Ko, M. H., & Gorenson, J. (2014). Technology-assisted vocabulary learning and student learning outcomes: A case study. *Multimedia-Assisted Language Learning*, 17(1), 11–33. Retrieved from http://journal.kamall.or.kr/wp-content/uploads/2014/04/Ko+Goranson_17_1_01.pdf
- Laufer, B. (1992). How much lexis is necessary for reading comprehension? In H. Bejoint & P. Arnaud (Eds.), *Vocabulary and applied linguistics* (pp. 136–132). London: Macmillan.
- Maera, P. (1995). The importance of an early emphasis on L2 vocabulary. *The Language Teacher*, 19(2), 8–10.

Using Spaced Repetition Software with a TOEIC Wordlist: Preliminary Results

- Maera, P., & Jones, G. (1989). *Eurocentres Vocabulary Test 10 KA*. Zurich: Eurocentres.
- McCrostie, J. (2009). TOEIC no turkey at 30: Once-struggling test flying high in Japan as corporate partners take on larger role. *The Japan Times*. Retrieved from http://www.japantimes.co.jp/community/2009/08/11/issues/toEIC-no-turkey-at-30/#.UyuKNIUb_D9
- Mehrpour, S., & Rahimi, M. (2010). The impact of general and specific vocabulary knowledge on reading and listening comprehension: A case of Iranian EFL learners. *System*, 38, 292–300.
- Miles, S., & Kwon, C. (2008). Benefits of using CALL vocabulary programs to provide systematic word recycling. *English Teaching*, 63(1). Retrieved from <http://praxised.com/research/miles-kwon.pdf>
- Mnemosyne. (2014). Mnemosyne (Version 2.2.1) [Software]. Available from <http://mnemosyne-proj.org/download-mnemosyne.php>
- Nation, I. S. P. (1982). Beginning to learn foreign vocabulary: A review of the research. *RELC Journal*, 13(1), 14–36.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press
- Nation, I. S. P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In Schmitt, N. & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6–19). Cambridge: Cambridge University Press
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press
- Runnels, J., & Rutson-Griffiths, A. (2013). Tablet PCs in a paperless classroom: Student and teacher perceptions on screen size. *JALT CALL Journal*, 9(3), 275–285. Retrieved from http://journal.jaltcall.org/articles/9_3_Runnels.pdf
- Takahashi, J. (2012). An overview of the issues on incorporating the TOEIC test into the university English curricula in Japan. *Tama University School of Global Studies Bulletin*, 4, 127–138.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55. Retrieved from <http://www.ijme.net/archive/2/cronbachs-alpha.pdf>
- Wozniak, P. (ND). *History of SuperMemo*. Retrieved from <http://www.supermemo.com/english/history.htm>
- Yazawa, H., Takeuchi, Y., & Mochizuki, H. (2009). Universities add TOEIC to the curriculum. *Education in Japan Community Blog*. Retrieved from <http://educationinjapan.wordpress.com/college-entrance-angst/universities-add-toiec-to-the-curriculum/>

—平成26年10月15日 受理—

Appendix

TOEIC Word List Vocabulary Quiz Dates for GCD Reading Classes Semester 1 2014

	Quiz 1 (words 1-50)	Quiz 2 (words 1-100)	Quiz 3 (words 1-150)	Quiz 4 (words 1-200)	Quiz 5 (words 1-250)	Quiz 6 (words 1-300)
RS1 Tuesday	May 9	May 27	June 10	June 24	July 8	July 22
RS1 Thursday	May 9	May 22	June 5	June 19	July 3	July 17
RS3 Wednesday	May 7	May 21	June 4	June 18	July 2	July 16
RS3 Friday	May 2	May 23	June 6	June 20	July 4	July 18