



Robust learning of vocabulary: investigating the relationship between learner behaviour and the acquisition of vocabulary

Alan Juffs, Lois Wilson, Maxine Eskenazi, Jamie Callan
Jonathan Brown, Kevyn Collins-Thompson, Michael Heilman, Timothy Pelletreau, James Sanders
University of Pittsburgh & Carnegie Mellon University

Introduction

- Need for long-term studies of L2 learning in Intensive English programs
- Need for clarification of success of incidental and explicit vocabulary learning
- Need for a wider range of vocabulary activities in the ELI reading curriculum
- Need for tracking of individual learner's moves at the Microlevel
- Availability of REAP technology: 10 million texts searchable by grade level, length and vocabulary item

Experiment Design

- Target vocabulary selected from Coxhead word list
- Pretest to select individual student's set of target words
- Based on target word list, student receives texts at grade 6 – 8 level with 2 to 4 target words
- If students get the question on the target word correct, the word is seen again later.
- If the student gets the question wrong, target word shown again almost immediately
- "Mastery" was defined as 2 successful correct answers for a given word
- Subjects could click on the word's definition in English.

Discussion

The experiment showed that making hints available on line to students did provide support that they were interested in receiving. Students clicked on 71% of hints.

However, use of hints did not improve mastery over the course of the term. It is possible that because all the students used the hints, the effects were not visible.

Clearly, some students were able to read more texts than others. More rapid reading, more text read, led to more words mastered. This is not surprising, but it confirms a view of some reading teachers that slowly poring over individual words is not efficient. However, it is hard to persuade students of this fact.

The REAP software showed the ability to supply texts that correspond to individual needs, but will be extended so that it also finds documents that are 'interesting' to the students.

A follow up on testing long-term retention is planned.

Incidental and explicit instruction

N. Ellis and Sinclair (1996) and others (e.g. R. Ellis and He 1999; Hulstijn et al., 1996; DeKeyser, 1995) have shown in laboratory conditions that focus on form and interaction with input and output can enhance learning. Hulstijn et al. (1996) investigated Dutch-speaking learners of French vocabulary. They divided the learners into three groups: a group that was given a marginal gloss for the new word, a group that was given a dictionary but not required to use it, and a control group. The target items were 16 words, 8 of which appeared only once in the readings and 8 of which appeared 3 times. The results were that in correct recalls, the marginal gloss group outperformed both the dictionary and the control groups, which did not differ. (The dictionary group did not use the dictionary very much – only about 12-15% of words, but when they did use the dictionary, they tended to recall the word's meaning.) There was a main effect for frequency, and a significant interaction between group and frequency for exact recall, but not when the exact and partial meaning were results were counted together. Ellis and He (1999) showed that interactive output was also beneficial. In the review of previous studies, Huckin and Coady (1999) mention only OMC computer assisted study (Chun and Plass, 1996). Chun and Plass (1996) used one 762 word text with a target set of 82 words which could be 'looked up' by clicking on a link to a definition or in some cases a picture. Results suggested that a quarter of the new words were learned, but no long-term study was completed. In a literature search of Linguistics Abstracts, only one other CALL-mediated study on vocabulary was found in the major journals: Jones and Plass (2002) used a (short) 2 minute 20 second listening activity in French as a second language with various multimedia annotations. They report that a combination of pictures and translations helped recall of vocabulary. However, no long-term follow-up is reported.

1. Does interaction with a definition of an unknown word on-line provide better learning over time than simply 'noticing' (in context) by a student? Words were selected on the basis of not being in the Reading 4 textbook, and being in one of the important academic word families identified by Coxhead (2000).

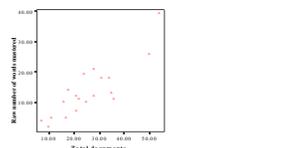
2. How do individual students use the tool once it is provided to them. Does simple availability of the tool ensure they will use it and benefit from it?

Basic correlations (n=19):

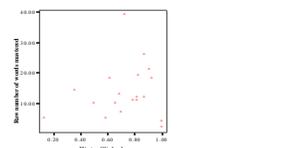
1. Number of texts read and total words mastered: $r = .86, p < .0001$
2. Hints clicked and total words mastered: $r = .16, ns$
3. MTELP score and total words mastered: $r = -.11, ns$
4. Accuracy on questions and total words mastered: $r = .10, ns$

Results

Correlation between number of texts read and total words mastered.



Correlation between number of hints clicked and total words mastered.



The Stimuli

Texts from the world-wide web via REAP prefiltered by reading level, length, and content of the target words
A set of important academic word families identified by Coxhead (2000). Each student had a different set of words, according to their individual knowledge.

Background to REAP: <http://orieans.lti.cs.cmu.edu/Reap/#?project>

REAP includes
i) a search engine that finds text passages satisfying very specific lexical constraints,
ii) Selection of materials from an open-corpus (the Web), thus satisfying a wide range of student interests and classroom needs,
iii) the ability to model an individual's degree of acquisition and fluency for each word in a constantly-expanding lexicon so as to provide student-specific practice and remediation.

This combination enables research on a wide range of reading comprehension topics that were formerly difficult to investigate.

The project includes a series of studies that test the central instructional implication of this project: Texts that closely reflect the reader's vocabulary and comprehension level can be used to support reading comprehension and vocabulary growth. The goal of these studies is to test the feasibility of text-to-reader matching on vocabulary while simultaneously testing hypotheses about the relationship between vocabulary-stretch and comprehension as well as the use of definitions. Our high-level hypothesis is that moderate stretch can be tolerated and can even produce increased learning of word meanings, whereas large stretches, because they tax the reader's ability to understand, may not. Reader characteristics are also important. Individuals vary in their tolerance of text difficulty and their tolerance likely depends on their incentive to read a given text, as well as reading ability and vocabulary knowledge.

Information about Participants

The students in the ELI are from a variety of countries, including China, Korea Taiwan, the Middle East and Latin America.

The students are from Level 4 reading classes, i.e. Intermediate English.

Many students intend to pursue academic studies in the US or Canada, but many are also studying English for general purposes.

Interpretation of Results

- Students seem to use the hints a great deal (on average 71% of hints are used), but this use of the tool does not predict "mastery"
- Number of documents read is the strongest predictor of words "mastered".
- "Mastery" is NOT predicted by general proficiency scores, which means that the tool is providing information not available in the students' vocabulary.
- Focus on accuracy in answering target questions may slow students down, and result in less learning since accuracy on target questions does not predict "mastery" of new words.

Future Questions

Future research will consider whether providing more input through REAP on known words will increase knowledge of:

- a) collocations that are idiosyncratic and
- b) verb meaning – syntax correspondences that are not idiosyncratic.

Summary

An experiment using REAP technology to automatically select texts that matched individual students' vocabulary learning needs showed that the number of documents that a student read was the best predictor of mastery.

Most students used hints provided on-line. This contrasts with an earlier study by Hulstijn et al (1996), which showed that students did not use dictionary resources when they were available offline.

We would like to thank the Robert H. Anderson Language Media Center.
This research has been funded by NSF grant SBE-0354420. Any opinions, findings and conclusions or recommendations expressed in this material do not necessarily reflect the views of the National Science Foundation.