

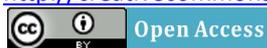
What's in a Bot? L2 Lexical Development Mediated through ICALL

Kelly Arispe

Boise State University, Boise, USA
Email: kellyarispe@boisestate.edu

Received 10 November 2013; revised 13 December 2013; accepted 20 December 2013

Copyright © 2014 by author and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY).
<http://creativecommons.org/licenses/by/4.0/>



Abstract

In recent years, the field of Second Language Acquisition (SLA) has made great strides to refocus its attention on the essential role that vocabulary plays in becoming a proficient L2 learner (Nation, 2001). Moreover, Computer Assisted Language Learning (CALL) has made advances in providing interactive online tools that help L2 learners strategically engage and work through their vocabulary development. This present study reports on how an Intelligent CALL tool (ICALL), *Langbot*, helps learners at the beginner and intermediate levels with their lexical acquisition. Modeled after instant messaging systems, which create a synchronous communicative environment, *Langbot* acts like a pedagogical scaffold or online *buddy* that caters to the vocabulary needs of each individual learner. It provides 1) translation requests with examples in context, 2) a frequency-based “word of the day” and 3) quizzes based on recent inquiries and a specific frequency range according to the learner’s level. The results from breadth and depth tests (N = 142), suggest that learners at all levels that have access to *Langbot* significantly improve their vocabulary breadth, while only intermediate-high learners with access to *Langbot* improve their vocabulary depth. Furthermore, survey data demonstrate *Langbot*’s effectiveness and accessibility based on learner perceptions.

Keywords

CALL; ICALL; SLA; Vocabulary Learning; Explicit Learning

1. Introduction

One of the major obstacles in learning a second language (L2) is the challenge of acquiring a completely new set of words that have a whole host of socially and culturally constructed meanings. L2 learners are met with what Nation (2001) calls the “learning burden” in that each individual approaches new words with different cultural as well as educational backgrounds and has to negotiate their meaning according to his/her learning style.

Simply put, the learning burden is the amount of effort it will take for the L2 learner to acquire the word's meanings and uses. In addition, "the sheer number of words and their patterns variously explains why language learning takes so long..." (Ellis, 2003: p. 69). Fortunately, research has demonstrated just how crucial lexical knowledge is for progress in the L2 (Cobb, 2008; Groot, 2000; Kelly, 1991; Koda, 1989; Laufer, 1992; Laufer & Hulstijn, 2001; Laufer & Nation, 1995; Linnarud, 1986; Nation & Webb, 2011; Ranalli, 2009; Sagarra & Alba, 2006; Segler, Pain, & Sorace, 2001; Singleton, 1999; Zapata & Sagarra, 2007). Moreover, research within the field of Computer Assisted Language Learning (CALL) has provided educators and learners alike with online vocabulary tools that provide scaffolded tutorials and practice in and out of the classroom.

As technology becomes more ubiquitous, it is not surprising that the field of SLA is rapidly incorporating CALL as a key component in the L2 learning environment (Blake, 2008; Jamieson, Chapelle, & Preiss, 2005; Kern & Warschauer, 2000). Nor is it surprising that CALL materials have become popular mediators used for vocabulary development (Chun & Payne, 2004; De la Fuente, 2003; Groot, 2000; Tozcu & Coady, 2004; Zapata & Sagarra, 2007). These tools help enhance language instruction and offer additional practice that doesn't deter from limited classroom instruction. Furthermore, in recent years, ICALL (Intelligent Computer Assisted Language Learning) has contributed to the field of SLA because their intelligent applications offer language learners a personalized experience. ICALL is a sub-field of CALL, which uses computational linguistics and natural language processing in order to augment the language learning process. While in recent years some research has been dedicated to vocabulary learning and ICALL, specifically vocabulary learner strategies (Segler, Pain, & Sorace, 2002), research has yet to answer the fundamental question, do vocabulary ICALL tutors help learners learn more words?

This paper, therefore, looks in-depth at the ways an ICALL tool, *Langbot*, acts as the vehicle whereby Spanish L2 beginner and intermediate learners can strengthen their vocabulary breadth and depth¹ through greater access to word definitions and context with the purpose of improving vocabulary development. Modeled after instant messaging systems, which create a synchronous communicative environment, *Langbot* acts like a pedagogical scaffold or online *buddy* that caters to the vocabulary needs of each individual learner. It provides 1) translation requests with examples in context, 2) a frequency-based "word of the day" and 3) quizzes based on recent inquiries and a specific frequency range according to the learner's level. Additionally, this tool forces L2 learners to be aware of words they don't know by holding them accountable to remember recent inquiries (through a quiz function) over a longer period of time.

1.1. Theoretical Underpinnings

In large part, the value placed on lexical knowledge in language development in recent years stems from a connectionist approach to how language is processed in the brain. Connectionist models show that associative-learning mechanisms account for complex language representations. Most importantly, connectionists believe that the parsing of chunks are what account for the strengthening of these associations (Ellis, 2003). As learners encounter a greater frequency and range of contexts in which they encounter a word, they make stronger associations. Thus, connectionism underscores the importance of the lexicon and its effect on L2 development².

Furthermore, learners, as well as researchers, recognize that vocabulary is essential for their linguistic progression. For example, Gass & Selinker (2001) demonstrated that lexical errors are the culprit for most communicative breakdown. In fact, they found that both native speakers and L2 learners perceive lexical errors as the fundamental reason why communication was disrupted. These perceptions alone challenge researchers to continue uncovering the way L2 learners process and remember words.

While there has been much debate about whether or not vocabulary learning is implicit, incidental, explicit or intentional³, it seems that research has established that lexical acquisition is *both* an implicit and explicit process (Ellis, 1994). In fact, Skehan (1992) suggests that adults will only be partially successful when learning an L2 incidentally. Moreover, most psychological models for memory posit that attention is actually necessary for "encoding a stimulus into long-term memory" (Logan, 1988, cited in Schmidt, 1993: p. 209). When it comes to

¹There are many ways to measure vocabulary acquisition. This topic has been widely debated and researched and is beyond the scope of this current study. To see a summarized discussion on this topic in great detail, consult Nation (2001).

²For more in-depth discussion on the main tenets of connectionism, consult Ellis Gasser, M., (1990) and Houghton, G. (2005).

³See Chun & Plass, 1996; Hulstijn, 1992; Knight, 1994; Laufer & Hulstijn, 2001; and Lupescu & Day, 1993. These studies look at the value of intertextual glosses, dictionaries, words embedded in context, etc. as evidence for and against implicit and explicit (and/or incidental vs intentional) learning.

lexical acquisition, Ellis (1994b) argues that explicit learning is needed to deal with “higher” facets of L2 vocabulary learning.

1.2. The Role of CALL in Vocabulary Learning

In many ways, teachers are often limited when it comes to effectively introducing and rehearsing new vocabulary. Furthermore, strategic, cognitive vocabulary training is something most language teachers either don't know how to teach or don't have time for in class. Although Sagarra and Alba (2006) powerfully demonstrated the success of the mnemonic key-word method, these types of strategic vocabulary techniques take time and are best suited for concrete types of words. Perhaps what language instructors need most in terms of vocabulary instruction are tools and strategies that don't take away from classroom instruction, but help augment instruction in, but mostly outside of class. This is precisely why CALL has had such a popular and visible effect, specifically in the realm of vocabulary learning.

Recently, Prince (2012) elucidated seven principles of vocabulary learning that can help guide future CALL vocabulary learning programs and tutorials. Namely, he established the value of 1) giving learners access and rehearsal to the most frequent words, 2) utilizing a combination of implicit and explicit methods, 3) taking advantage of explicit methods that lead to deeper processing (as has been previously discussed), 4) allowing learners to see and understand the word according to its authentic context, 5) encountering the word at least 6 - 10 times (underscoring the value of repetition), 6) providing translations as a complementary approach to seeing the word in context, and 7) allowing learners to see the words via different modalities (for example, auditory as well as visual). It should be noted that Prince's principle regarding the value of repetition corroborates findings from Webb's (2007) robust study involving 121 Japanese learners of English that demonstrated significant gains for both receptive and productive vocabulary that had been incidentally presented in context ten times.

CALL programs have also underscored the value of frequency in tandem with explicit vocabulary learning. In a fairly recent study, Tozcu and Coady (2004) revealed that explicit attention to vocabulary has the potential to greatly improve not only vocabulary gains, but to also augment reading comprehension and decrease word recognition reaction time when reading. Their study was grounded on Coady, Carrell and Nation's (1985) theory of learning which values the emphasis of high frequency words in language instruction because it leads to automaticity and greater overall L2 development gains.

Tozcu and Coady tested 56 intermediate level students learning English as a second language. 28 participants in the experimental group had access to a CALL Tutorial program, *New Lexis*, that targets the 6400 most frequent words in English via three main components: study, practice, and review. The other 28 students in the control group completed reading and comprehension task activities that simulated normal classroom exposure to vocabulary without explicit attention or instruction. Pre and posttests were administered to both groups over a period of two months. Although both groups improved their scores in all three areas (vocabulary, reading comprehension and reaction time), the experimental group showed a significant difference in all three at the $p < 0.001$ levels. They suggested that their findings “indicate that individualized vocabulary learning on the computer will almost certainly facilitate vocabulary acquisition” (p. 491). They ultimately concluded that learners should focus on learning the 2000 - 3000 most frequent words as quickly as possible through direct vocabulary instruction (p. 492). This study is important not only because it highlights the benefit of explicit attention to vocabulary, but it also recognizes the value of frequency and CALL as key components in L2 language development. Moreover, the authors explain:

Individualized instruction on the computer appears to be a highly efficient method for learning the most frequent words since second language learners tend to know different subsets of the highly frequent vocabulary, limiting the likelihood of effective group instruction (Coady et al., 1993). Furthermore, individualized instruction on the computer is beneficial since it is done outside of class without utilizing precious class time (p. 492).

Indeed, the authors underscore the advantage of having access to an individualized tutor. In fact, their CALL Tutorial had many of the features that have characterized the latest advances in online language learning tools, namely, intelligent CALL (ICALL).

1.3. Personalized Vocabulary Learning via Intelligent CALL

One of the greatest gains in recent technological advances is that computer programs have evolved from online

mechanical exercises to dynamic language tutors. Heift & Schulze (2007) have helped elucidate the historical development of ICALL tools and the challenges that await computational linguists aiming to provide more intelligent feedback to learners. Undoubtedly, the quality of ICALL that is most attractive for language development is that it provides individualized feedback that stimulates language growth specific to the level of the learner, thereby drawing attention to the explicit, strategic process of learning a language. Essentially, ICALL programs provide personalized language practice, while training learners to become more aware of how they are accomplishing their linguistic goals⁴.

ICALL systems inherently provide more learner control than traditional CALL programs due to their sophisticated answer processing mechanisms. Unlike the more traditional drill and practice programs, ICALL software employs Natural Language Processing (NLP) which overcomes the rigidity of the response requirements of traditional CALL (Heift, 2002: p. 296).

One of the major advantages of using ICALL for vocabulary development is that it interacts with the learner on an individual basis. Heift (2002) acknowledged that ICALL has been shown to be effective specifically for form-based instruction. In her study that looked at error correction and an ICALL application (a German Intelligent Language Tutoring System), she found that the program interacted with each individual according to his/her language skill level and situated learners in a learner-controlled environment. For this reason, ICALL has the ability to scaffold language learning because it intelligently interacts with each individual learner at the appropriate level. Furthermore, ICALL has the potential to raise learner awareness about the learning process. Bull (1997) states that intelligent CALL programs are “designed to raise awareness of the variety of strategies available and to allow students to make informed choices about the approaches most useful to them” (p. 12, cited in Ranalli, 2009).

1.4. What’s in a Bot

This present study, looks specifically at an ICALL vocabulary tutor called Langbot. Langbot is an intelligent agent that serves as an interactive online vocabulary “buddy” that is both a language reference and a vocabulary-training assistant that offers a whole host of vocabulary tools, targeting the most frequent 3000 words in Spanish (Payne, 2009: p. 6). By design, a Bot is a small computer application that works alongside other computer applications or programs and can be used instantaneously for vocabulary development in the L2.

Langbot offers opportunities for both implicit and explicit vocabulary development and provides the following features: 1) translations for words, phrases or sentences, 2) examples of words or phrases used in a wide variety of contexts with hyperlinks to additional online resources, 3) individualized vocabulary quizzes, and 4) a “word of the day” specific to a frequency range as determined by the individual’s level. Furthermore the bot encourages the learner to communicate in the target language to receive feedback (although English is allowed).

Langbot has the ability to guide each individual user in the next step of his/her development because it is an adaptive, intelligent tool that can cater to the individual progression of the learner. “It can automatically record its interactions with users, store these data in a database, and perform analyses behind the scenes to inform future interactions with individual users and access their vocabulary development” (Payne, 2009: p. 5). For example, vocabulary quizzes contain words from recent inquiries as well as the most frequent words just beyond the learner’s level so as to scaffold him/her to a higher level of lexical proficiency. Thus, Langbot is dynamic because it has the power to individually assess each learner’s progress so as to correctly guide the learner in his/her L2 vocabulary development.

1.5. The Study

This study explored L2 vocabulary development as mediated by Langbot. It was cross-sectional in nature and measured vocabulary growth (breadth and depth) at three main levels: beginner, intermediate-low and intermediate-high. A quantitative analysis measured overall vocabulary improvement based on the outcome gains of a pre and posttest by comparing results between the experimental (access to Langbot) and control (no access to Langbot) groups. In addition, a Language History Questionnaire was administered to give better insight to vocabulary learning from the student’s perspective and a *Langbot Evaluation Survey* (for experimental groups only)

⁴Learner autonomy is beyond the confines of this article. To read more about the connection between CALL and learner autonomy, see Colentine, 2011; Hafner & Miller, 2011; Lee, 2011; and Reinders & White, 2011.

measured learners' overall satisfaction with the tool. The study probed the following questions:

- 1) Does Langbot help learners acquire more words at a faster rate (breadth) than those who do not use this ICALL tool?
- 2) Does Langbot use lead to greater depth of knowledge for the words acquired at the intermediate level?
- 3) Does Langbot help make learners more aware of their L2 vocabulary development?

2. Method

2.1. Research Design

In order to measure L2 vocabulary growth, a cross-sectional study was conducted over the duration of ten weeks (one quarter) and administered to six classes (N = 142) at three different proficiency levels: beginner (N = 57), intermediate-low (N = 42) and intermediate-high (N = 43). At each level there were at least two classes, one with access to Langbot (the experimental group) and one without (the control group). Results were triangulated based on vocabulary assessments, survey data and qualitative analysis⁵.

2.2. Participants

The participants in this study were divided into three main groups according to predicted proficiency in Spanish as a second language: beginner (first year), intermediate-low (second year) and intermediate-high (third year). Due to the nature of the treatment (Langbot), the control and experimental groups were divided according to course level. This gave the researcher the ability to collaborate with the instructor on how to implement Langbot into the course structure for the classes that pertained to the experimental group. In order to get a robust sample size, two to four classes at each level (year) were selected to participate in the study. In addition, only the intermediate-low and intermediate-high groups (both control and experimental) received the depth test since beginner learners would not be at an appropriate level to participate in this type of meaning based vocabulary test. An outline for the participant groups according to their treatments is listed in [Table 1](#).

2.3. Exclusions

There were three main exclusions that were taken into consideration for the participants in this study: 1) participant's first language, 2) number of years of high school or college Spanish, and 3) less than half of quizzes completed with Langbot. This information was ascertained first, through the Language History Questionnaire and second through access to the Langbot Manager (to be explained in the next section).

First, participants that were bilingual Spanish speakers, or who spoke Spanish with at least one parent, were excluded from the study. Provided that this research was interested in vocabulary growth for traditional beginner and intermediate L2 learners⁶, it was determined that these participants were false beginner/intermediate learners. Second, participants that spoke a first language other than English and spoke that language at home with every family member (i.e. parents and siblings) were also excluded from the study. Given the fact that both vocabulary measures (breadth and depth tests) were administered in English and required participants to acknowledge the

Table 1. Distribution of subjects according to level and treatments.

Level	# of participants	Access to Langbot	Breadth test	Depth test	Langbot survey
Beginner	25 (2 classes)	Yes	Yes	No	Yes
Beginner	32 (2 classes)	No	Yes	No	No
Intermediate-low	20 (1 class)	Yes	Yes	Yes	Yes
Intermediate-low	22 (2 classes)	No	Yes	Yes	No
Intermediate-high	25 (2 classes)	Yes	Yes	Yes	Yes
Intermediate-high	18 (1 class)	No	Yes	Yes	No

⁵Due to limitations in length, the qualitative analysis was not included in this article.

⁶*Traditional* refers to learners that have no background experience with the Spanish language. Heritage language learners may be classified as beginner or intermediate for a multitude of reasons, however the nature of this study was to determine lexical gains for learners that would not have had any outside exposure to the beginner/intermediate vocabulary frequency range.

target word's equivalent in English, it would not be a fair measure to include learners that were not dominant English speakers. Third, participants that had taken 4 years of high school Spanish were automatically excluded from the beginner level for both control and experimental groups. Lastly, participants in the experimental group were excluded if they completed less than half of the number of Langbot quizzes required. Ultimately, the re-researcher wanted to ensure that participants had used Langbot consistently over the ten-week period in order to ensure that the experimental group was indeed using the treatment⁷.

2.4. The Experimental Group

The experimental groups (N = 70) were randomly selected at each level and every class member was given access to Langbot. Within the first week of the academic quarter, the researcher gave each class a training session on how to use Langbot and obtained necessary permission to conduct the study according to IRB standards. Participants were taken to the language laboratory on the university campus and the researcher instructed learners on how to install and successfully use Langbot⁸. It is important to mention that Langbot conducted a preliminary language history “interview” with the learner before he/she could use it for individual use. This enabled Langbot to profile each learner and supply him/her with words appropriate to his/her level for the word-of-the-day and the weekly quizzes (1 - 2000 most frequent words for beginners, and 2000 - 3000 most frequent words for intermediates). **Figure 1** is an example of the language history interview:

During the training session, the researcher also explained the requirements for using Langbot as an integral component of students' course evaluation as a percentage of their homework grade. Specifically, participants were asked to sign onto Langbot daily to accompany their homework, for example, working on an essay or reading a short-story and answering discussion questions. **Figure 2** provides an example for how a participant might ask Langbot for translation help, in this specific case to translate the word *adquisición*⁹.

Additionally, each day Langbot supplied participants with a word-of-the-day that strategically came from one of the top 1 - 3000 most frequent words depending on the learner's level (Davis, 2002). The word-of-the-day would pop-up spontaneously while the participant used the tool. It supplied one frequent word in Spanish, along with its' definition and an example for how the word is used in context, as well as a hyperlink that could

Langbot Introductory Interview:

You will now begin a short proficiency assessment. Please answer the questions as accurately as possible.

How many languages do you speak?

How many years have you been learning Spanish (use .5 for one semester)?

Have you ever lived in a country where Spanish is spoken?

For how long? (Enter the number of years)

When you read a Spanish text, how much do you understand without using a dictionary (on a 1-10 scale: 1 = 'very little' to 10 = 'almost every word')?

When you listen to Spanish on a video or the radio, how much do you understand (on a 1-10 scale: 1 = 'very little' to 10 = 'almost everything')?

When you write in Spanish about topics your teacher assigns to you, how easy is this task for you (on a 1-10 scale: 1 = 'very difficult' to 10 = 'easy')?

How well can you discuss topics like sports or politics in Spanish (on a 1-10 scale: 1 = 'very difficult' to 10 = 'very easy')?

Thank you. Your answers will help Langbot provide better assistance.

Figure 1. Langbot's language history interview.

⁷It should be noted, however, that all participants were included in the evaluation survey regardless of quizzes completed.

⁸Langbot could be accessed from a normal laptop as well as tablets and mobile devices that had access to the Internet.

⁹Note: “tr” indicates “translate” and “y” means yes; these quick codes were created for students to access the tool quickly.

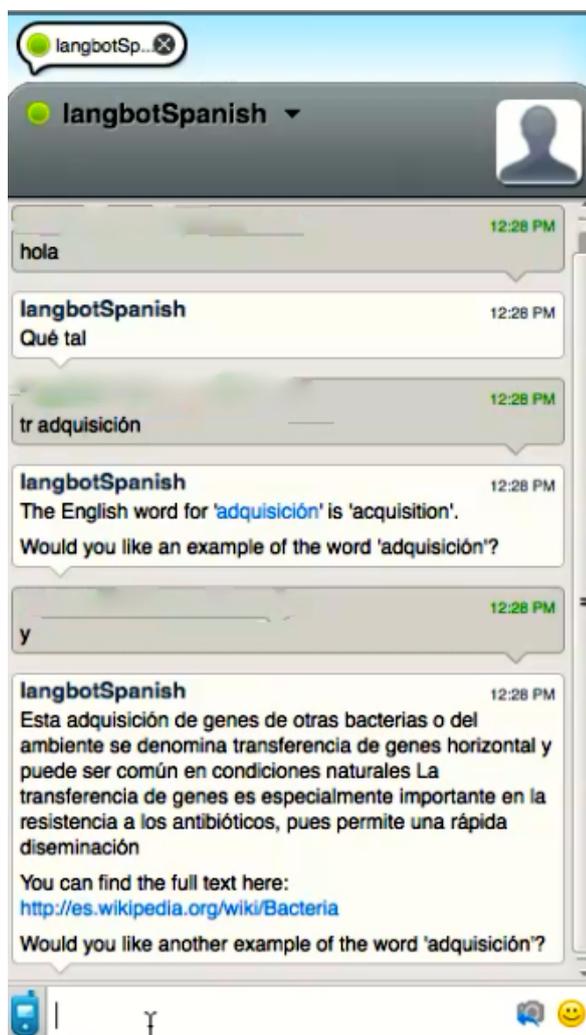


Figure 2. An example of a translation request using Langbot; the username has been blocked to protect anonymity.

instantaneously take the learner to an external site, Spanish Dict.com¹⁰, where the learner could see a list of synonyms, conjugation charts (if appropriate), and in many cases a video with a native speaker who uses the word in context, thereby providing the pronunciation of the word as well.

Lastly, participants were asked to complete a 10-word vocabulary quiz weekly. Subjects were provided with a word in Spanish and had to provide its English equivalent. In the event that the student got the definition wrong, Langbot provided the correct definition. This quiz function within Langbot could be accessed multiple times per day (by simply typing “quiz”), but for the purposes of the study, participants had to complete one quiz each week to receive credit (for a total of 10 quizzes). In many circumstances, participants completed more than one quiz, however they were not rewarded with extra points for doing so. The researcher worked with each instructor to help ensure that students completed the weekly quiz and to combat any technical difficulties. Additionally, the researcher updated the online confidential Langbot score (based on quizzes completed) every week. Figure 3 shows an example of a typical online quiz (NB: only the first three items are present, but each quiz contained 10 questions):

Participants were required, therefore, to use Langbot when working on assignments and to pay attention to the word-of-the-day for each day of the instructional week (Monday-Friday), as well as to complete the weekly vo-

¹⁰It should be noted that this site was chosen in particular because the top 3000 most frequent words in Spanish have a video explanation for the word in question. There is no way to record if and how frequently participants accessed these hyperlinks. The only data that points to this is from the evaluation survey that points to students’ perceptions about the hyperlinks.

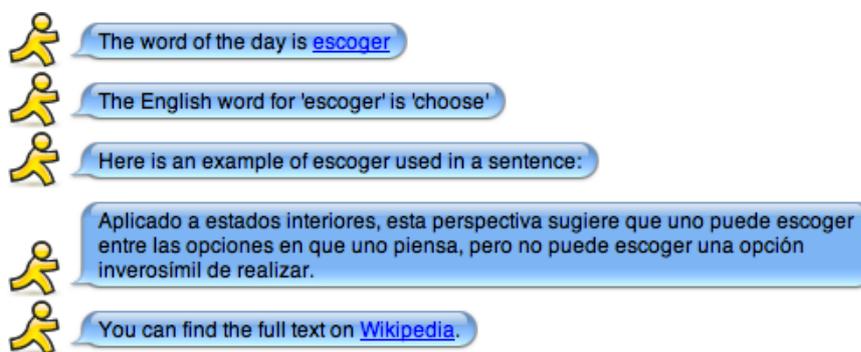


Figure 3. An example of the word-of-the-day in Langbot.

cabulary quiz. The researcher had access to each participant’s log for analyzing the data, as well as keeping track and making sure that participants were successfully completing the weekly vocabulary quizzes. If the researcher noticed attrition for a specific participant, the researcher contacted the participant to help troubleshoot any problems.

2.5. The Control Group

The control groups ($N = 72$) were randomly selected at each level. Participants in the control group did not have access to Langbot. For each proficiency level, the syllabus for the class was identical (for experimental and control groups) except for this aspect of the course evaluation (i.e. the homework grade did not include Langbot since these participants did not have access to the tool). Although at each level (beginner, intermediate-low, intermediate-high) vocabulary learning is treated as an inherent component of the course, it is not explicitly evaluated as part of the course grade (except for what is tested on exams). In other words, there is no specific “vocabulary” component to the course that makes this aspect of the L2 learning process explicit. Furthermore, instructors generally did not emphasize how to learn vocabulary strategically.

2.6. Instructors

There were a total of 10 instructors in this study—five in the experimental group and five in the control group. All instructors were graduate teaching assistants (TAs). This allowed for relative continuity in methodology and class structure across the different proficiency levels. All instructors had an overarching teaching supervisor that was responsible for designing and helping implement a standard course syllabus. Additionally, all TAs received thorough teacher training to help prepare and scaffold their teaching approach in the classroom. Nevertheless, differences in instruction have been accounted for as part of the experimental design and will be discussed in the results section. It should also be noted that the researcher was the instructor for the experimental group at the intermediate-low level.

2.7. Data Collection

The vocabulary tests were a word frequency breadth test and a word association depth test (Wesche & Paribakht, 1996). These tests were administered twice—once at the beginning (week 1) and once at the end of the ten weeks (week 10). One highly fluent and one native speaker both reviewed the tests for accuracy. Participants were administered the test online and test items were randomized. One questionnaire (administered at week 1) and one survey (administered at week 10) were also administered; the first was a Language History Questionnaire had helped the researcher determine the linguistic background for each participant, the second gathered student perceptions based on the role of vocabulary in L2 learning.

2.8. Evaluation Measures

Word Frequency Vocabulary Tests

A random selection of 30 words from the 1000 most frequent words was selected from Davis’ word frequency list for the beginner level. A random selection of 30 words from the top 3000 most frequent words was selected

for the intermediate-low and intermediate-high groups. The researcher excluded obvious Spanish-English cognates and grammatical terms (i.e. prepositions like *de*, *en*, *por*, etc.). Participants at the intermediate levels were also asked to complete a 10 word association depth test (explained below) and, for this reason, it was decided that these participants only complete 20 words from the definition breadth test so that there wouldn't be the risk of hindering test reliability due to attrition.

Although there are several measures that can adequately determine whether or not a learner has improved his/her vocabulary breadth, Nation (2001) argues that one of the most reliable, efficient, and effective breadth tests is a simple translation recognition task between the L1 and L2. Therefore, all participants in each level were asked to match the target word with its definition in English. The overall gain over the ten week period was then compared between the experimental and control groups. This measure helped answer the question, "Does Langbot help learners acquire more words at a faster rate (breadth) than those who do not use this ICALL tool?"

As a second evaluation measure, word association depth tests, were also administered twice (at weeks 1 and 10) to participants at the intermediate levels (both low and high). Hulstijn & Tangelder (1991 & 1993) found that word meaning tests (for this study, word association depth tests) were not appropriate for beginner learners who were still struggling to formally recognize the phonological and/or orthographic characteristics of a word. However, at the intermediate/advanced levels, the ability to know a word's meaning as it relates to other words (collocations, synonyms, etc.) conveys a greater depth of knowledge. In a previous study, Read's Word Associate Format, "WAF" (1993), asked participants to match the target word with four others (out of 8 total) that shared some sort of relationship: paradigmatic (synonyms), syntagmatic (collocates) and analytic (representing part of the meaning of the word) (Read, 1997: p. 317). Read found that this type of test was a successful measure and that "the format is an economical means of assessing the learner's range of knowledge of high-frequency content words" (317). The WAF test has been widely used for pedagogy as well as research because it has demonstrated both testing efficiency and reliability¹¹.

The word association depth test in this study was based largely on Read's WAF, however for obvious reasons the target word and word associates (possible responses) were in Spanish. As previously explained, the target words were selected according to the frequency range that was appropriate for intermediate learners, in this case up to 3000. The researcher wanted to avoid the problem of overt guessing, and for this reason participants were told to select all answers that applied (rather than 4, as is the case for the WAF). Interestingly, this supports one of the suggestions that Schmitt, Ng, & Garras, (2011) highlighted in their study when they reviewed Read's WAF. Ultimately, participants had to choose the correct responses out of a total of 5 possible options. Additionally, participants were asked to give the basic definition of the target word. This was done for two reasons. First, the researcher wanted to award participants for partial knowledge (i.e. the simple definition). Second, and most importantly, the researcher wanted to ensure that participants weren't simply guessing. There were ten questions for the word associates test, with six possible points awarded for each question (up to five if they correctly did or did not select the word associate, and one point for adequately defining the word). Like the breadth tests, a random selection of 10 words were chosen for the depth test. Obvious cognates and grammatical terms were also excluded from this measure. Ultimately, this assessment helped answer the research question, "Does Langbot use lead to greater depth of knowledge for the words acquired at the intermediate level?"

3. Results

3.1. Research Question 1: Vocabulary Breadth

The statistical significance of experimental observations was assessed using the mixed model procedure lmer () in R (Bates & Maechler, 2009). The analysis of breadth was conducted in a model that included the fixed effects access/no access and level (beginner, intermediate-low, and intermediate-high) and the random effect of class (there were ultimately 10 classes). Ultimately the random effect was removed from the model because it was determined to be insignificant. Contrasts were setup to determine mean comparisons against the main treatment, access/no access to Langbot and level (the fixed effects). The Bonferroni correction was used to control the type 2 error rate. The asterisks in Table 2 denote significance at the 0.05 level, indicating that all levels (beginner, intermediate-low, and intermediate-high) that had access to Langbot showed significant improvement when comparing their post test score with their pre test score for vocabulary breadth. None of the levels that did not have access to Langbot showed significant improvement.

¹¹Schoonen and Verhallen (2008) found that Read's WAF had a reliability of 0.75 - 0.83.

Table 2. Analysis of variance for access to LangBot across levels according to vocabulary breadth.

Vocabulary breadth	Treatment		df
	Access $\mu \pm SE$	No access $\mu \pm SE$	
Beginner	1.36 \pm 0.74* N = 25	0.23 \pm 0.49 N = 32	1
Inter-low	3.40 \pm 0.44* N = 20	2.27 \pm 0.57 N = 22	1
Inter-high	2.48 \pm 0.39* N = 25	0.29 \pm 0.49 N = 18	1

* $p < 0.05$ within rows.

3.2. Research Question 2: Vocabulary Depth

The statistical significance of experimental observations for vocabulary depth was assessed using the mixed model procedure `lmer()` in R (Bates & Maechler, 2009). The analysis of depth was conducted in a model that included the fixed effects access/no access and level (beginner, intermediate-low, and intermediate-high) and the random effect of class. Just as was the case for breadth, the random effect was removed from the model because it was again determined to be insignificant. Contrasts were setup to determine mean comparisons against the main treatment, access/no access to Langbot and level (the fixed effects). The Bonferroni correction was used to control the type 2 error rate. The asterisks in Table 3 denote significance at the 0.05 level, indicating that the intermediate-high group that had access to Langbot showed significant improvement when comparing their post test score with their pre test score for vocabulary depth.

Langbot Usage

Lastly, it was important to take overall usage into consideration to provide a broader analysis of vocabulary development as it pertains to both breadth and depth. Langbot kept track of all messages and transcripts so that the researcher and/or instructor could monitor each individual's progress. Therefore, the researcher quantified the total number of messages that each individual participant had with Langbot and then analyzed overall usage (number of messages) with the total gain as determined by the difference in posttest and pre test scores for both breadth and depth. The statistical significance for this correlation was assessed using PASW Statistics 18.0 (IBM SPSS Inc.). Pearson Correlations were used in order to measure the strength of dependence for the two continuous variables (number of messages and improvement score). The results in Table 4 indicate that breadth and overall usage positively correlate at the 0.05 level.

3.3. Research Question 3: Lexical Awareness

The *Langbot Evaluation Survey* was administered to all participants in the experimental group in order to understand, from the users' perspective, if and how Langbot was benefiting their lexical acquisition. Ultimately, this measure helped answer the third research question, "Does Langbot make learners more aware of their L2 vocabulary development?". The survey consisted of twelve main questions that can be reviewed in Appendix A. Even though these participants did indeed improve their vocabulary breadth and the intermediate-high group their vocabulary depth, the researcher believed it was important to give voice to their overall impressions and intuitions. After all, learners are also consumers and their overall satisfaction with the tool is crucial in long-term learning.

Table 5 summarizes the frequencies for each response on the survey. For the reader's convenience, the question has been included in an abbreviated form to the left of the table. There are 4 main columns: the first is the average frequency for all participants (from groups beginner, intermediate-low, and intermediate-high). The columns to the right depict the frequency for each group according to level.

First and foremost, it was important to ascertain whether or not participants felt that their vocabulary improved as a direct result of using Langbot. They were not aware of their vocabulary pre and posttest scores, and thus had no way to measure, aside from their own intuitions, that Langbot had indeed improved their vocabulary development. Their intuition is valuable because it helps predict if learners will continue to use the tool in the future.

Table 5 demonstrates that indeed, most participants believe that their vocabulary did improve as a result of using Langbot (74.0%). Furthermore, their satisfaction increases (79.9%) when the intermediate learners are

Table 3. Analysis of variance for access to LangBot across intermediate levels according to vocabulary depth.

Vocabulary depth	Treatment		df
	Access $\mu \pm SE$	No access $\mu \pm SE$	
Inter-low	2.45 \pm 1 N = 20	4.3 \pm 1 N = 22	1
Inter-high	4.62 \pm 0.92* N = 25	1.47 \pm 0.79 N = 18	1

* $p < 0.05$ within rows.**Table 4.** Pearson correlations for LangBot messages and breadth and depth scores.

	Messages
Breadth	0.245*
Depth	-0.003

* $p < 0.05$.**Table 5.** Frequencies for Langbot evaluation survey (N = 70).

Question	Total	Beg.	Int. low	Int. high
Do you feel that your vocabulary improved from using Langbot?				
Yes	74	52.9	86.4	73.3
No	20.5	38.2	13.6	16.7
Somewhat	5.5	8.8	0	10
Did you become more aware of your need to improve your vocabulary?				
Yes	91.8	88.2	86.4	100
No	8.2	11.8	13.6	0
Did Langbot encourage you to consult additional resources online?				
Yes	56.2	52.9	72.7	53.3
No	43.8	47.1	27.3	46.7
Did you find Langbot to be more interactive than other online tools?				
Yes	45.2	41.2	22.7	60
Somewhat	43.8	38.2	68.2	30
No	11	20.6	9.1	10
†When did you consult Langbot most often?				
Writing	41.1	47.1	31.8	30
Preparing for a test	4.1	11.8	4.5	0
Reading	39.7	20.6	50	60
Quizzes	13.7	20.6	13.6	10
Do you agree that the quiz function benefits Spanish language learners?				
Strongly agree	27.4	17.6	27.3	30
Agree	56.2	55.9	63.6	60
Disagree	15.1	23.5	9.1	10
Strongly disagree	1.3	2.9	0	0

Continued

Were the examples it gave you helpful?				
Very helpful	8.2	17.6	9.1	6.7
Somewhat helpful	64.4	44.1	68.2	73.3
Not helpful	27.4	38.2	22.7	20
Will you use Langbot more if you have access via facebook?				
Yes	37	38.2	18.2	53.3
Indifferent	42.5	35.3	40.9	46.7
No	20.5	26.5	40.9	0
Did you go to the Wikipedia site when you were given an example?				
Always	0	0	0	0
Sometimes	32.9	26.5	36.4	33.3
Rarely	39.7	23.5	45.5	43.3
Never	27.4	50	18.2	23.3
†Do you want examples to continue to be provided?				
Yes	58.9	44.1	59.1	76.7
No	39.7	52.9	40.9	23.3
Will you use the tool in the future?				
Yes	63.1	58.8	54.5	66.7
Maybe	30.1	32.4	27.3	30
No	6.8	8.8	18.2	3.3

†Indicates that there was missing data (less than 2%) for this question.

solely considered.

Also very crucial to language learning in general, is a learner's awareness in terms of one's development and progress. Because Langbot is a personalized ICALL tool, it can give each individual learner specific feedback by probing quiz questions that are specific to the learner's recent inquiries. Not only does this help with frequency (i.e. the learner sees the word more than once), but it also holds the learner accountable and reveals whether or not one has actually mastered the word's meaning. Results suggest that an overwhelmingly high percentage (91.8%) of participants believe that Langbot made them more aware of their need to improve their vocabulary. Furthermore, every single member of the intermediate-high group (100%) found this to be true.

4. Discussion

Segler, Pain, and Sorace (2002) called for more research to understand the relationship between ICALL tools and vocabulary development. Consequently, this study contributes to a growing field within SLA that is just beginning to understand the implications for ICALL tools and L2 vocabulary learning. ICALL tutors are beneficial because they bring to light an area of the SLA process that is often overlooked in traditional L2 classrooms and it is an area that L2 learners themselves struggle to practice independently, as well as effectively.

The quantitative analysis suggests that learners who use Langbot improve their vocabulary breadth. In other words, learners with access to Langbot regardless of the level—beginner, intermediate-low, and/or intermediate-high—show significant gains in vocabulary breadth from pre to posttest ($p < 0.05$). Conversely, learners that did not have access to Langbot—at all levels—showed no significant improvement from pre to posttest. These findings were further supported by the correlational data that found breadth to be a significant factor when looking at the relationship between overall Langbot usage (measured by number of messages with Langbot) and

gain in pre to posttest score. The positive correlation (0.245) suggests that the more learners use Langbot, the greater the probability of improving their vocabulary score.

The issue of vocabulary depth has become very important as the field of SLA gives greater attention to vocabulary learning in general. Both researchers and educators recognize that the value of lexical proficiency is measured not only by knowledge of a vast number of basic words, but also by the implied syntagmatic and paradigmatic relationships and social-cultural characteristics associated with any given word. More often than not, learners are never explicitly taught to make sense of these complex relationships that embody words; instead, they are left to implicitly or independently make sense of this depth of meaning through reading or independent study. This study sought to determine if ICALL can also help learners deepen their understanding of any given word's associations.

The results from the vocabulary depth test suggest that learners at the intermediate-low level—regardless of having access to Langbot—do not show significant improvement in terms of depth of meaning. On the other hand, only intermediate-high learners with access to Langbot significantly improved their vocabulary depth. The correlational data ultimately conclude, however, that the number of messages does not indicate a higher depth score (unlike the significant correlation for breadth). There are several possible interpretations that, given further study and analysis, could help explain why 1) intermediate-low learners do not improve their vocabulary depth scores, and 2) there is no correlation between usage and depth—despite the significant improvement for depth with intermediate-high learners that did have access to Langbot.

First, it could very well be that learners at the intermediate-low level, like beginners, are still grappling with the basic translation-knowledge of words and are not capable of recognizing or dealing with lexical depth. Moreover, there is a fundamental difference in course content that is presented at the intermediate-low versus intermediate-high levels (and was the principal reason why the researcher was interested in measuring the difference between the two). The significant gain for intermediate-high learners that had access to Langbot may indeed reflect the combination of two main factors—first, the sheer exposure to words in context through a significantly higher demand in reading, and second, the assistance of Langbot to manage this reality. Thus, the nature of this relationship (i.e. demands of reading and Langbot assistance) may be the reason why the correlation is ultimately non-conclusive.

Finally, this study was interested in examining the relationship between the benefit of ICALL tools and learner awareness—a quality of language learning that has been highly valued in achieving and working towards autonomous learning¹². Results from the *Langbot Evaluation Survey* show that 91.8% of learners agree that Langbot made them more aware of their need to learn vocabulary. What is interesting is that beginner learners respond favorably 88.2% of the time, as well as intermediate-low learners (86.4%). Every single participant at the intermediate-high level, however, agrees that Langbot successfully raises awareness in relation to vocabulary learning (100%).

Furthermore, there seemed to be a greater need for certain aspects of the tool as the level increased. Specifically, when participants were asked to respond whether or not they wanted to continue to receive examples for how the word is used in context, their positive responses increase with each level (beginners, 44.1%; intermediate-low, 59.1%; intermediate-high, 76.7%). This response alone suggests that there is a pattern for breadth and depth when considering the difference between beginner and intermediate learners. It seems that learners at more advanced levels value having more information about the words they do not know (i.e. they want depth of meaning). It might be suggested that because they have more lexical associations they are therefore wanting to strengthen those associations via bigger semantic and grammatical networks. Conversely, beginner learners may simply want, and perhaps need, the basic definition without the “extras”.

On the one hand, [Tozcu and Coady \(2004\)](#) stated that intermediate learners should concentrate on the 2000 - 3000 most frequent words as quickly as possible through direct vocabulary instruction (i.e. explicit learning). However, L2 learners are inundated with a vast array of unknown words that are academic, as well as colloquial in nature. Lexical demands become more diverse and vastly idiosyncratic as the learner advances and as the course content shifts from a language learning focus to a literature or content focus; there is little if no class time to address these individual lexical needs. ICALL tools have the possibility and opportunity to fill this pedagogical void and provide both implicit and explicit attention to lexical development and word association building

¹²Learner Autonomy (LA) was not discussed in this present study, however the researcher acknowledges the importance of LA in addition to vocabulary breadth and depth as a result of access to ICALL tools. For more discussion on the value of LA in CALL, see Little, 2004 & 2007; Porto, 2007; and Schwienerhorst, 2008.

(among many things). Nevertheless, the value of Langbot was also supported at the beginner level where participants need exposure and practice to build a lexical foundation in the L2. This is indeed the point; ICALL tools like Langbot are adaptive and learners need this flexibility throughout their learning process.

5. Conclusion

5.1. General Implications of the Study

If the need for explicit attention to vocabulary becomes more complex as the level increases (i.e. breadth as well as depth), and the classroom environment is no longer conducive to dealing with these issues, ICALL tools can fill a very important niche in L2 classrooms and SLA in general. Moreover, the participants in this study valued the accessibility of Langbot in addition to their overall impression that the tool was effectively improving their vocabulary development. Ultimately, this study corroborates previous research on the importance ICALL in L2 learning. The capacity for Langbot to cater to individual needs in L2 lexical development supports findings that ICALL tools help learners become more aware of their learning (Bull, 1997; Heift, 2002).

5.2. Limitations and Directions for Future Research

The main findings in this study reflect the benefit an ICALL tool, Langbot, can exert on vocabulary breadth and depth, however, there were several limitations that should be addressed in future research. First, this study was cross-sectional and measured receptive vocabulary gains in a relatively short-term, 10-week span. There is no question that future research must also demonstrate Langbot's capacity to help learners retain vocabulary breadth and depth longitudinally. It would be beneficial to know, for example, the optimal number of repetitions in quizzes to transfer word recognition from short to long-term memory.

Second, this study was interested in comparing the typical Spanish L2 classroom (at beginner and intermediate levels) that was enhanced by ICALL assistance (i.e. access to Langbot). It could be argued, however, that *any* explicit attention to vocabulary learning is beneficial. Future research, therefore, should probe which qualities of ICALL vocabulary tutors are best, as well as to compare ICALL vocabulary tutors vs. explicit attention in a classroom setting. For example, future studies may wish to compare four different types of possible L2 learning environments: 1) the traditional classroom, 2) the traditional classroom + ICALL vocabulary tutor, 3) explicit vocabulary classroom instruction, vs. 4) explicit vocabulary classroom instruction + ICALL vocabulary tutor. Ultimately, it would be helpful to know which learning environment provides an optimal environment for learning L2 vocabulary and the degree to which, if any, exposure to an ICALL vocabulary tutor supersedes other types of explicit vocabulary attention.

Third, this study was interested in measuring lexical gain according to the typical L2 learner. The researcher acknowledges, however, that the learning environment is not homogenous. Specifically, the L2 classroom is made up of L2 and heritage language learners (HLL). The HLLs were excluded in this because they would have skewed the data. Nevertheless, HLLs might greatly benefit from access and accountability to knowing academic terms, for example. Future research should find a way to creatively measure lexical gains for HLLs since lexical proficiency is of great need and importance to these types of learners as well.

Finally, future research needs to consider broader ways to measure the relationship between Learner Autonomy (LA) and ICALL tools. Just as Reinders & White (2011) recognize that the definition of autonomy can be ambiguous and challenging to pin point, there is potential for ICALL tools to meaningfully draw personalized attention to areas where learners need to reinforce, review and enhance their learning.

References

- Bates, D., & Maechler, M. (2009). lme4: Linear Mixed-Effects Models Using Eigen and S4. *R package version 0.999375-32*.
- Blake, R. J. (2008). *Brave New Digital Classrooms*. Georgetown, Washington DC: Georgetown University Press.
- Bull, S. (1997). Promoting Effective Learning Strategy Use in CALL. *Computer Assisted Language Learning, 10*, 3-39. <http://dx.doi.org/10.1080/0958822970100102>
- Chun, D., & Payne, S. (2004). What Makes Students Click: Working Memory and Look-Up Behavior. *System, 32*, 481-503. <http://dx.doi.org/10.1016/j.system.2004.09.008>
- Coady, J. M., Carrell, P., & Nation, P. (1985). *The Teaching of Vocabulary in ESL from the Perspective of Schema Theory*. Milwaukee: Midwest TESOL.

- Coady, J., Magoto, J., Hubbard, P., Graney, J., & Mokhtari, K. (1993). High Frequency Vocabulary and Reading Proficiency in ESL Readers. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second Language Reading and Vocabulary Learning* (pp. 217-228). Norwood, NJ: Ablex Publishing Corporation.
- Cobb, T. (2008). Commentary: Response to McQuillan and Krashen. *Language Learning & Technology*, 12, 109-114.
- Davis, M. (2002). Corpus del Español. M. Davis.
- De la Fuente, M. (2003). Is SLA Interactionist Theory Relevant to CALL? A Study on the Effects of Computer-Mediated Interaction in L2 Vocabulary Acquisition. *Computer Assisted Language Learning*, 16, 47-81. <http://dx.doi.org/10.1076/call.16.1.47.15526>
- Ellis, N. & Beaton. (1993). Psycholinguistic Determinants of Foreign Language Vocabulary Learning. *Language Learning*, 43, 559-617. <http://dx.doi.org/10.1111/j.1467-1770.1993.tb00627.x>
- Ellis, N. (1994a). Consciousness in Second Language Learning: Psychological Perspectives on the Role of Conscious Processes in Vocabulary Acquisition. *AILA Review*, 11, 37-56.
- Ellis, N. (1994b). Vocabulary Acquisition: The Implicit Ins and Outs of Explicit Cognitive Mediation. In N. C. Ellis (Ed.), *Implicit and Explicit Learning of Languages* (pp. 211-282). London: Adademic Press.
- Ellis, N. (2003). Constructions, Chunking, and Connectionism: The Emergence of Second Language Structure. In C. J. Doughty, & M. Long (Eds.), *The Handbook of Second Language Acquisition* (pp. 63-103). Malden, MA: Blackwell. <http://dx.doi.org/10.1002/9780470756492.ch4>
- Gass, S., & Selinker, L. (2001). *Second Language Acquisition: An Introductory Course* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.
- Groot, P. J. M. (2000). Computer Assisted Second Language Vocabulary Acquisition. *Language Learning and Technology*, 4, 60-81. <http://ilt.msu.edu/vol4num1/groot/default.html>
- Heift, T. (2002). Learner Control and Error Correction in ICALL: Browsers, Peekers and Adamants. *CALICO*, 19, 295-313.
- Hulstijn, J. H. (2001). Intentional and Incidental Second-Language Vocabulary Learning: A Reappraisal of Elaboration, Rehearsal and Automaticity. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 258-286). Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139524780.011>
- Hulstijn, J., & Tangelder, C. (1991). *Intralingual Interference: Lexical Errors as a Result of Formal and Semantic Similarities among English Word-Pairs*. Paper Presented at the First Annual Meeting of the European Second Language Association (EUROSLA), Salzburg.
- Hulstijn, J., & Tangelder, C. (1993). Semantic and Phonological Interference in the Mental Lexicon of Learners of English as a Foreign Language and Native Speakers of English. In J. CHapelle, & M. T. Claes (Eds.), *Actes: 1er Congrès International: Mémoire et Mémorisation dans l'Acquisition et l'Apprentissage des Langues/ Proceedings: 1st International Congress: Memory and Memorization in Acquiring and Learning Languages*, Louvain-la-Neuve: CLL.
- IBM SPSS Inc. (2010). PASW STATISTICS 18.0 Command Syntax Reference. SPSS Inc., Chicago.
- Jamieson, J., Chapelle, C., & Preiss, S. (2005). CALL Evaluation by Developers, a Teacher, and Students. *CALICO Journal*, 23, 93-138.
- Kelly, P. (1991). Lexical Ignorance: The Main Obstacle to Listening Comprehension with Advanced Foreign Language Learners. *IRAL*, 29, 135-149.
- Kern, R., & Warschauer, M. (2000). *Network-Based Language Teaching: Concepts and Practice*. Cambridge: Cambridge University Press.
- Koda, K. (1989). The Effects of Transferred Vocabulary Knowledge on the Development of L2 Reading Proficiency. *Foreign Language Annals*, 22, 529-540. <http://dx.doi.org/10.1111/j.1944-9720.1989.tb02780.x>
- Laufer, B. (1992). How Much Lexis Is Necessary for Reading Comprehension. In Arnaud & Béjoint (Eds.), *Vocabulary and Applied Linguistics*, Basingstoke: Macmillan.
- Laufer, B., & Hulstijn, J. (2001). Incidental Vocabulary Acquisition in a Second Language: The Construct of Task-Induced Involvement. *Applied Linguistics*, 22, 1-26. <http://dx.doi.org/10.1093/applin/22.1.1>
- Laufer, B., & Nation, P. (1995). Vocabulary Size and Use: Lexical Richness in L2 Written Production. *Applied Linguistics*, 16, 307-322. <http://dx.doi.org/10.1093/applin/16.3.307>
- Linnarud, M. (1986). *Lexis in Composition*. Lund: Lund Studies in English.
- Logan, G. D. (1988). Toward an Instance Theory of Automatization. *Psychological Review*, 95, 492-527. <http://dx.doi.org/10.1037/0033-295X.95.4.492>
- Nation, I. S. P., & Webb, S. A. (2011). *Researching and Analyzing Vocabulary*. Heinle, Cengage Learning.
- Nation, P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139524759>

- Payne, S. (2009). A Proposal for the Langbot Project. Unpublished.
- Prince, P. (2012). Towards an Instructional Programme for L2 Vocabulary: Can a Story Help? *Language Learning & Technology*, 16, 103-120.
- Ranalli, J. (2009). Prospects for Developing L2 Students' Effective Use of Vocabulary Learning Strategies via Web-Based Training. *CALICO Journal*, 27, 161-186.
- Read, J. (1993). The Development of a New Measure of L2 Vocabulary Knowledge. *Language Testing*, 10, 355-371. <http://dx.doi.org/10.1177/026553229301000308>
- Read, J. (1997). Vocabulary and Testing. In P. Bogaards, & B. Laufer (Eds.), *Vocabulary in Second Language: Selection, Acquisition, and Testing*, 10, 302-320.
- Read, J. (1998). Validating a Test to Measure Depth of Vocabulary Knowledge. In A. Kunnan (Ed.), *Validation in Language Assessment* (pp. 41-60). Mahwah, NJ: Erlbaum.
- Sagarra, N. & Alba, M. (2006). The Key Is in the Keyword: L2 Vocabulary Learning Methods with Beginning Learners of Spanish. *Modern Language Journal*, 90, 228-243. <http://dx.doi.org/10.1111/j.1540-4781.2006.00394.x>
- Schmidt, R. (1993). Awareness and Second Language Acquisition. *Annual Review of Applied Linguistics*, 13, 206-226. <http://dx.doi.org/10.1017/S0267190500002476>
- Schmitt, N., Ng, J. W. C., & Garras, J. (2011). The Word Associates Format: Validation Evidence. *Language Testing*, 28, 105-126. <http://dx.doi.org/10.1177/0265532210373605>
- Schoonen, R., & Verhallen, M. (2008). The Assessment of Deep Word Knowledge in Young First and Second Language Learners. *Language Testing*, 25, 211-236. <http://dx.doi.org/10.1177/0265532207086782>
- Segler, T. M., Pain, H., & Sorace, A. (2002). Second Language Vocabulary Acquisition and Learning Strategies in ICALL Environment. *Computer Assisted Language Learning*, 15, 409-422. <http://dx.doi.org/10.1076/call.15.4.409.8272>
- Singleton, D. (1999). *Exploring the Second Language Mental Lexicon*. Cambridge Applied Linguistics. <http://dx.doi.org/10.1017/CBO9781139524636>
- Tozcu, A., & Coady, J. (2004). Successful Learning of Frequent Vocabulary through CALL Also Benefits Reading Comprehension and Speed. *Computer Assisted Language Learning*, 17, 473-495. <http://dx.doi.org/10.1080/0958822042000319674>
- Webb, S. (2007). The Effects of Repetition on Vocabulary Knowledge. *Applied Linguistics*, 28, 46-65. <http://dx.doi.org/10.1093/applin/aml048>
- Wesche, M., & Paribakht, T. S. (1996). Assessing Second Language Vocabulary Knowledge: Depth versus Breadth. *Canadian Modern Language Review*, 53, 13-40.
- Zapata, G. C., & Sagarra, N. (2007). CALL on Hold: The Delayed Benefits of an Online Workbook on L2 Vocabulary Learning. *Computer Assisted Language Learning*, 20, 153-171. <http://dx.doi.org/10.1080/09588220701331352>