Virtual Vocabulary: Research and Learning in Lexical Processing

ULF SCHUETZE
GERLINDE WEIMER-STUCKMANN
University of Victoria

ABSTRACT
This article presents the concept development, research programming, and learning design of a lexical processing web application, Virtual Vocabulary, which was developed using theories in both cognitive psychology and second language acquisition (SLA). It is being tested with first-year students of German at the University of Victoria in Canada, specifically looking at intervals and time-sets in a large study on spaced retrieval. The findings are being used to refine language-learning software for second language lexical acquisition.

KEYWORDS
Second Language Acquisition (SLA), Cognitive Psychology, Working Memory, Spaced Retrieval, Spaced Rehearsal, Vocabulary Learning, Intentional Learning, Repetition

INTRODUCTION
Virtual Vocabulary (ViVo) was developed for college-level German and is being tested with first-year students. However, ViVo can also serve as a platform for other languages as well as second-, third- and fourth-year students. This use would require some cosmetic changes, but the principles of programming and design would remain the same. ViVo presents lexical items to learners using images, sound files, lexicogrammatical information, target language sample sentences, and intercultural information. Our approach is influenced by research on multimedia (Jones & Plass, 2002; Kim & Gilman, 2008; Rimrott, 2009). For research purposes, insights from intelligent computer-assisted language learning (ICALL) were used in the programming of intervals according to the principles of spaced retrieval developed in cognitive psychology (Carpenter & DeLosh, 2005; Landauer & Bjork, 1978). ViVo analyzes language at the level of the lexical item. Currently, learners practice lexical items with ViVo according to interval schedules. The results of this research will be used to program optimal rehearsal frequencies that will allow learners to study lexical items efficiently according to their individual rehearsal schedules. The program is currently being tested as both a research tool and as language-learning software at the University of Victoria in Canada.

RATIONALE
Neurology has for many years informed research issues in cognitive psychology and linguistics. Yet there still seems to be a gap between the two disciplines. Jacobs and Schumann (1992) suggested that language acquisition researchers must begin to incorporate a degree of neurobiological reality into their perception of the language acquisition process. Fifteen years
later, Hulstijn (2007) still cautions against the divide in the research in neurophysiology and linguistics and recommends that researchers of both disciplines work together. Interestingly, although for language learners the acquisition of large amounts of new vocabulary is a major task, textbooks, curriculum designers and researchers often neglect this aspect of language learning (Schmitt, 2008).

**Online Vocabulary Program: Concept**

Initially, we developed ViVo as a vocabulary trainer. Later, we modified it for use as a research tool. For the design of its interface we consulted theories in cognitive psychology on memory, working memory, and spaced retrieval as applied to second language processing.

**Memory**

Conway, Jarrold, Kane, Miyake, and Towse (2007) describe memory as “the ability to mentally maintain information in an active and readily accessible state, while concurrently and selectively processing new information” (p. 3). Research into lexical processing is linked to research into memory. All aspects of lexical acquisition such as input, processing, storage, retrieval, and output are inseparable from cognitive processes of our brain (Aitchison, 2003; Friederici, Mueller, & Rüschemeyer, 2006). In milliseconds (Friederici, 2002; Kandel, 2007), we can access memories dating from years ago or experience an immediate recall of something encountered mere seconds ago.

Models of memory at work are abundant. We can distinguish two strongly debated theoretical schools of research: the structuralist approach (memory is comprised of different components, i.e., the model of working memory by Baddeley & Hitch, 1974) versus the functional approach (memory is comprised of one store, retention differences are viewed as differences of depth of processing, i.e., Craik & Lockhart, 1972).

**Working Memory**

The "three-store model" introduced by Baddeley and Hitch (1974) is described in more detail below because its cognitive psychology findings are often cited when addressing second language acquisition (SLA) conditions. It consists of a sensory buffer, a working memory, and a long-term memory. These registers each serve a different purpose and differ in capacity, length of retention, and encoding processes. The sensory buffer is modality based (i.e., phonological, visual, and olfactory). It receives sensory data input but does not forward it for further processing unless this input receives more attention. Only when data are attended to are they remembered. The working memory is described as a construct with three components: the central executive with its substructures of the visuo-spatial sketchpad and the phonological loop. Later, Baddeley added the episodic buffer as a fourth component. He describes it as the interface and a binding mechanism between the three working subsystems and the long-term memory (Baddeley, 2007). The central executive controls incoming sensory data and manages their processing, the visuo-spatial sketchpad processes nonverbal information, and the phonological loop stores phonological information and rehearses it on a subvocal level. This aspect is particularly interesting for lexical processing: lexical items need to be rehearsed to receive the amount of attention necessary for forwarding. The long-term memory then comprises a human being’s world knowledge, emotions, and thoughts. Its workings are believed to trigger the formation of a complex system of neurophysiological structures.
**Lexical Processing**

Lexical processing in language learning consists of two procedures: recognition as part of comprehension and retrieval as part of production. Even today the more commonly propagated computer metaphor assumes an organization in which lexical units are stored and accessible for retrieval following algorithms likened to computer organization. It is referred to as an interface (Raupach, 1994) that allows us to process cognitive concepts as input or output that are then transposed into encoded representations in our minds. How these processes manifest themselves is still under debate. Two models have been predominant: Marslen-Wilson and Welsh’s (1978) cohort model and the interactive activation model as described by Aitchison (2003). In both models, several lexical items are activated by a sound. A matching process unfolds until the correct item is chosen. The matching process is influenced by the lexical environment, and, whereas for Marslen-Wilson and Welsh this process is hierarchical, Aitchison (2003) and Singleton (1999) view it as parallel. In general, the item that receives the most attention is activated. Hulstijn and Laufer (2001) call this the “involvement load hypothesis:” they distinguish among the need to use a word in a sentence, the search to find the word form expressing a concept, and the evaluation by the learner to compare that word with other words and recognize differences. In each of these three phases, the lexical item in question receives attention. Once the lexical item has been found and checked, it is activated. The more the phonological loop rehearses the lexical item subvocally, the higher the activation will be.

**ONLINE VOCABULARY PROGRAM: PROGRAMMING**

For the sake of consistency, ViVo presents to every student every lexical item the same number of times (frequency) over the same period of time (time-set). Recently, Schmitt (2008) compiled a survey of all studies carried out on second language lexical processing. Frequency of exposure to a lexical item in these studies ranges from three to 20 or more encounters over a time period of a few days to 2 weeks. Typically this amounts to about five encounters over 10 days. Accordingly, the project described here set up five encounters: all students were introduced to the lexical item (first encounter) and then had four review sessions over a time-set of 7 days (four encounters) followed by an online and a print quiz. This was also done for practical reasons because the acquisition of lexical items in a first-year German class at the university level often covers a time frame of 2 weeks per chapter and the learning needs to take place within that time frame.

For this project ViVo investigated the ‘spacing effect’ described above (Baddeley, 2007; Hulstijn & Laufer, 2001). It reflected two spaced learning intervals: one with uniform intervals (see Table 1) and the other with graduated intervals (see Table 2).

| Table 1 |
| Uniform Interval Schedule |
| Distribution in class | day | interval |
| 1. Encounter: practice and immediate review | day 1 | 0 |
| 2. Encounter: review | day 3 | 2-day |
| 3. Encounter: review | day 5 | 2-day |
| 4. Encounter: review | day 7 | 2-day |
| 5. Encounter: review | day 9 | 2-day |
| Online quiz | day 11 | 2-day |
| Print quiz | day 15 | 4-day |
Table 2
Graduated Interval Schedule

<table>
<thead>
<tr>
<th>distribution in class</th>
<th>day</th>
<th>interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Encounter: practice and immediate review</td>
<td>day 1</td>
<td>0 interval</td>
</tr>
<tr>
<td>2. Encounter: review</td>
<td>day 1</td>
<td>1/2-day interval</td>
</tr>
<tr>
<td>3. Encounter: review</td>
<td>day 2</td>
<td>1-day interval</td>
</tr>
<tr>
<td>4. Encounter: review</td>
<td>day 4</td>
<td>2-day interval</td>
</tr>
<tr>
<td>5. Encounter: review</td>
<td>day 7</td>
<td>3-day interval</td>
</tr>
<tr>
<td>Online quiz</td>
<td>day 11</td>
<td>4-day interval</td>
</tr>
<tr>
<td>Print quiz</td>
<td>day 15</td>
<td>4-day interval</td>
</tr>
</tbody>
</table>

The uniform and graduated schedules were designed according to research on spaced retrieval (Carpenter & DeLosh, 2005; Landauer & Bjork, 1978), the “forgetting curve” (Ebbinghaus, 1913), and Oxford’s (1990) suggestions for second language vocabulary acquisition.

**Spaced Retrieval**

The basic idea of spaced learning is to return to a previous learning task after some time has passed by simply repeating a lexical item, choosing different modes of presentation (i.e., audio, visual), or expanding the information on the item at each rehearsal. It is therefore different from massed learning in which a lexical item is simply repeated several times without a time delay or a distractor in between. Recent research has shown that spaced learning leads to higher retention than massed learning (Balota, Duchek, & Logan, 2007; Cull, 2000). However, research into interval length of spaced learning has been inconclusive. For example, Landauer and Bjork (1978) carried out an experiment in which students had to remember in their first language (English) 16 name/face cards that were part of a card deck of 50 cards. After each of the 50 cards was presented once for 9 seconds each, students had to go through the card deck three more times. Each card was again rehearsed for 9 seconds; 12 cards appeared four times each within the deck, and the other 4 cards were interleaved as distracters. The distracters were presented in either uniform (i.e., 4-4-4) or expanded (i.e., 1-4-10) patterns. Students were assigned randomly to one of the patterns. After the students went through the card decks three times, they listened to a 30-minute lecture to distract them and then took a retention test on the 12 cards. Results showed that the expanded pattern led to much higher retention rates than the uniform pattern. Carpenter and DeLosh (2005) carried out a similar experiment. However, name/face cards were presented on a computer, the number of items was increased to 30, and, after each item was presented for 6 seconds, the three rehearsals had no time limit for students. When students finished, they were asked to name as many US States as possible in 5 minutes. Immediately afterwards, the 30 items plus the 18 distracters were tested. The distracters were spaced 3-3-3 (uniform) and 1-3-5 (graduated). Results showed no differences. A second experiment expanded the graduated pattern to 3-5-7, but there were still no differences.

These studies were carried out in the context of cognitive psychology using controlled learning environments in which seconds were measured. The question is, how would this type of spaced retrieval unfold if the methodology were used in the environment of language learning?
Second Language Memorization Systems

In addition to the question of interval length, it is not clear how often and in what time frame the acquisition of a lexical item should occur to be successfully retrieved. Interestingly, some of the studies on spaced learning as well as some of the studies on word memorization cited Ebbinghaus’s research of the late nineteenth century in which he conducted experiments on himself concerning the forgetting of words.

After having created test corpora of short nonsense syllables, Ebbinghaus memorized sets of eight, twelve, or sixteen of the nonsense syllables and measured the amount of time he needed for these rehearsals. It is important to note that Ebbinghaus set out to study forgetting, not learning. He therefore observed this process of forgetting and measured the seconds he saved when relearning a previously accomplished task. He defined the learning task as completed as soon as he had achieved the objective of two errorless reproductions of the learned item. In various test sessions he addressed the following question (Ebbinghaus, 1885): How much time and learning effort can be saved after repetition with spaced intervals of 20 minutes, 1 hour, 9 hours, 1 day, 2 days, 6 days, and 31 days? He compared the learning time with the relearning time and measured the seconds he had saved when relearning the material. Many years later, in 1913, he published a conversion of this data into a table, commonly known as the “forgetting curve.”

Many researchers share the view of Hulstijn and Laufer (2001) that an explicit memorization stage following other strategies (i.e., inferring, verifying) will greatly improve retention. Oxford (1990) promotes a staggered processing of learning material in her renowned SLA textbook. She suggests seven encounters with optimal intervals of 15 minutes, 1 hour, 2 hours, 1 day, 4 days, 1 week, and 2 weeks. However, this has not been tested. Nation (2001) promotes “direct learning of vocabulary” with word cards and states that this method of direct learning should be part of an overall vocabulary learning agenda.

A prominent learning device using the concept of structured cyclical repetition is Leitner’s (1972) “hand computer.” The hand computer, or as Leitner calls it in German, die Lernkartei, is a memorization device consisting of flashcards and a box with five to six sections of progressively larger size (see Figure 1).

Figure 1
Leitner’s Hand Computer, die Lernkartei

The term to be learned is written on one side of the flashcard, and the prompt on the other side. Ideally, the daily input, determined by the size of a daily vocabulary learning routine, is limited to 12 to 15 new lexical items. The flashcards then begin their learning cycle in the first compartment. Users take them out and memorize them. As the first compartment is filled with flashcards, these are reviewed. Cards for correct responses move up to the next
compartment. Incorrect ones drop back to, or remain in, the first compartment. The size of these compartments is progressively larger. Because it takes longer to fill the compartments and new input is limited to twelve and to fifteen items, the review periods are spaced.

However, the question is, are they really spaced? The flash card boxes have been in use in the German school system for many years. Their usefulness is promoted in teacher training programs (Schröder & Roedig, 2007). English as a Second Language textbooks publish sets of flashcards to match their corpora or create multimedia vocabulary learning environments based on cyclical learning. Apart from concerns raised about the impracticability of flash cards and issues of how to edit and proofread content (Lüders, 2005; Mondria & Mondria-De Vries, 1994; Schmitt, 2000), there has been no published research on how these flash cards are processed and handled. Students determine when and how long they will practice (and memorize). Even the practicing of the cards in the first compartment may therefore vary from seconds to days.

Based on the theory of working memory as well as what has been learned so far from research into spaced learning, we designed and programmed web-based application software to not only present and practice the lexical corpora, but also, for the purpose of research, track and document the process of learning of every single lexical item for every single user.

**ONLINE VOCABULARY PROGRAM: DESIGN**

ViVo applies the concept of explicit vocabulary instruction as a learning process that requires attention (Schmidt, 2001). It displays advantages that the nonelectronic, handheld *Lernkartei* did not provide. Mondria and Mondria-De Vries (1994) recommended the design of an electronic version favoring multimodal presentation options.

The concept of ViVo can be compared to a flash card system with target items in L1 and their L2 representation. However, as a web application, ViVo allows for additional presentation features and interactive tasks, for the design of which we looked to research on multi-media learning (e.g., Jones & Plass, 2002; Kim & Gilman, 2008) and on different learning style preferences (Cohen, 2003; Oxford, 2003). At Simon Fraser University, Rimrott (2009) has explored different presentation modes (visual, audio, definition, and International Phonetic Alphabet) in her recent research with first-year German students. Her findings suggest that a combination of all four options was overall better for most students.

Furthermore, the contextualization and expanded intercultural and lexicogrammatical information correspond to a concept of depth of vocabulary knowledge (Nation, 1990). The acquisition of a target item is a process and encompasses the sum of ‘subknowledges’ (i.e., register collocation, pronunciation, grammatical features, and morphological information). However, as a program for beginners, ViVo represents target items in their basic, most frequently used meanings. In addition to L1 and L2 target item presentations, ViVo uses image files, sound files, lexicogrammatical information, target language sample sentences, intercultural information, and practice writing fields with spell check.
Figure 2 illustrates the Practice mode of the program.

Figure 2
Practice Mode in ViVo

In Review mode, the interface provides students with the option of producing L2 target items (see Figure 3). Correct spelling is indicated by the background color, incorrect spelling is corrected in a prompt. This self-testing procedure is graded from the mere copying of the lexical items to production induced by an L1 prompt.

Figure 3
Review Mode in ViVo
If necessary, users can return to the Practice interface of the one item they are studying. However, users cannot go back to the beginning of the training session, and, once their schedule’s session of 40 items (see below) has been completed, they receive the message “You are done for today.” This ensures that while students have unlimited time to study and complete their session, they only do so according to their schedule of practice intervals.

In sum, ViVo provides controlled enriched multimodal training tasks for beginners.

**ViVo IN USE**

ViVo is currently being used and tested at the University of Victoria. In the fall of 2008, it was used with four sections of first-semester German and in the spring of 2009 with three sections of second-semester German with approximately 30 to 35 students in each section.

**Corpora**

The textbook used at the University of Victoria is *Deutsch Na Klar* (Di Donato, Clyde, & Vansant, 2007), and our vocabulary is based on that text. In German 100A/B, students acquire an active/passive learner corpus of about 100-120 words per chapter. Three selection decisions underlay the ViVo corpora. The first was to select 40 out of the 100-120 words per chapter as the active lexical corpora. Students cannot generally rely on short-term memory for as many as 40 words (Baddeley, 2007; Cowan 2001, 2005), and any attempt to learn all lexical items the day before the quiz is forestalled. The second decision was to use Jones and Tschirner’s (2006) *A Frequency Dictionary of German* as a filter to select the more frequent words. This dictionary is based on 4.2 million words of contemporary German that had been compiled from various registers (spoken, literary, journalistic, academic, and instructional language). The third was to achieve a balanced mix of content words (nouns, verbs, adjectives), function words (prepositions, conjunctions, pronouns, adverbs), and cognates (defined as L1 and L2 lexical items that are identical or nearly identical, i.e., *die Lampe* ‘the lamp’; Carroll, 1992). Aitchison (2003) refers to content words as the “bricks” and function words as the “mortar” although the metaphor is only approximate. The ratio of content to function words was set at three to one.

**Participants**

One hundred seventeen students were enrolled in first-semester German and 90 students in second-semester German. The students in first-semester German had had no prior German language instruction or knowledge of German as identified by means of a questionnaire. For research purposes, another questionnaire was administered at the end of each semester in order to eliminate students from the data analysis who learned vocabulary from sources other than ViVo.

**ViVo as a Research Tool**

In the Lernkartei, practice intervals were in random distribution; in ViVo, they were controlled. Furthermore, ViVo functioned as a tracking tool and documented the participants’ entries.
The students enrolled in the four sections of first-semester German were divided into two groups according to two test conditions. One group used ViVo to learn vocabulary items following a graduated spaced learning schedule (GG). The second group used ViVo to learn vocabulary items following a uniform spaced learning schedule (UG). Most students continued their study of the language in the second-semester course. If they were part of the GG group, they stayed in this group; if they were part of the UG group, they stayed in that group. A few of the new students in second-semester German had some prior knowledge of German. They completed the questionnaire and, if selected, were assigned to one of the two testing conditions.

All user entries were tracked and documented in the database of ViVo. User logs were used to establish how often and when the GG and UG students accessed and completed the rehearsal tasks.

The test corpora consisted of one online quiz and one print quiz per chapter. All 40 items practiced per chapter were tested online; 20 items per chapter were tested in print. The retrieval of these lexical items based on the test scores of all quizzes for both groups (GG and UG) is currently being analyzed. The main research question is: Will students who learned vocabulary on a practice schedule with graduated intervals demonstrate higher results on their tests than students who practiced on a schedule with uniform intervals?

**ViVo as Language-Learning Software**

Once the research project has been completed, ViVo will be used in the curriculum of first- and second-year German at the University of Victoria. It will be programmed according to the research findings so learners work with an optimal interval length in a defined time-set. This will allow learners to study lexical items efficiently according to their individual rehearsal schedules.

**CONCLUSION**

The focus of this project was on the concept, design, and programming of ViVo, both as a research tool and as language-learning software, exploring the interrelatedness of intervals and time-sets. Our research agenda on lexical processing in SLA is still being pursued.

The next question is that of long-term retention. Both Read (2004) and Nation (2001) describe vocabulary acquisition with all its aspects as a procedural continuum. The tracking function of ViVo allows individual analysis of each student progressing through first- and second-year German thereby blending the quantitative research of the current study with qualitative research of a future study. An expansion of the program to other languages such as French or Spanish is also possible and desirable.
REFERENCES


AUTHORS’ BIODATA

Ulf Schuetze is Assistant Professor of Second Language Acquisition in the Department of Germanic and Slavic Studies at the University of Victoria. His research interest is information technology and second language acquisition. His research project on lexical processing brings together linguistics, second language acquisition, cognitive psychology, and intelligent computer-assisted language learning (ICALL).

Gerlinde Weimer-Stuckmann recently completed her MA thesis on lexical processing. With over 20 years of language teaching experience in Germany as well as in Canada and as a trained speech pathologist, she is now preparing for her Ph.D. in second language pedagogy.

AUTHORS’ ADDRESSES

Ulf Schuetze
Department of Germanic and Russian Studies
P.O. Box 3045
University of Victoria
Victoria, BC, V8W 3P4
Canada
Phone: 250 721 7323
Fax: 250 721 7319
Email: ulfs@uvic.ca

Gerlinde Weimer-Stuckmann
Department of Germanic and Russian Studies
P.O. Box 3045
University of Victoria
Victoria, BC, V8W 3P4
Canada
Phone: 250 721 7503
Fax: 250 721 7319
Email: gweimer@uvic.ca