

Efficiency in Second Language Vocabulary Learning

Ulf Schuetze

University of Victoria, Canada

A question some researchers have been trying to answer for decades is how to optimize vocabulary learning in second language acquisition. In this context, learning refers to the ability to actively produce a word as Cook and Singleton (2014) have pointed out. There are two methods. One has been labeled as incidental learning and the other as rote learning. They have to be seen in relation to the level of proficiency. Incidental learning assumes that students learn words by reading authentic texts because these texts provide a rich source of a language's vocabulary (Hulstijn, 1992). This method is often used once learners have reached an intermediate level of proficiency because it relies on the knowledge of a base vocabulary. It also requires much reading (Hulstijn, 2001; Schmitt, 2010; van Zeeland & Schmitt, 2013). For beginning learners an effective method is rote learning, also called intentional learning (Barcroft, 2009; van Zeeland & Schmitt, 2013). Rote learning isolates words from a given context. Although it is desirable that a learner is able to use a word in a given context from a pedagogical point of view (Nation, 2001), from a psycholinguistic point of view, rote learning is efficient because repetition fosters the so-called subvocal rehearsal process (Ellis, 1995).

In the context of rote vocabulary learning, the technique of "spaced repetition" has been promoted as being particularly efficient since the sixties (Leitner, 1972; Mondria & Mondria-De Vries, 1994; Pimsleur, 1967). Cognitive psychologists conducted many experiments in this area (for an overview, see Balota, Duchek, & Logan, 2007). These experiments researched the efficiency of different intervals (massed, uniform, expanded intervals) in a single learning session with undergraduate students having to remember uncommon words in their first language, non-words, or digits and being tested five, ten, or thirty minutes later. They showed that a massed interval, that is, repeating a word several times in a row without a break in-between repetitions, leads to low recall rates. The problem with this type of interval becomes obvious when looking at studies of working memory.

The function of working memory is to prolong the time during which words can be processed. This is a temporary process. The main component is the so-called central executive that directs, divides, and switches attention to incoming words, a process that demands much energy (Baddeley & Hitch, 1974; Baddeley, 1986; Cowan, 1999; Gathercole & Alloway, 2008). If attended to, words are processed by another device called the phonological loop, which rehearses and temporarily stores the phonemes of those words (Baddeley, 2007; Gathercole, Pickering, Hall, & Peaker, 2001). This happens subvocally. The loop can code linguistic information that is auditory (listening/speaking) as well as visual (reading/writing). It does so by binding novel sounds into sequences that make up words. The phonemes must be identified by their combination of vowels and consonants, and must also be matched to a lexeme (the core of a word that carries meaning) in Wernicke's area of speech comprehension. The loop has two separate subsystems: rehearsal and temporary storage. They work in parallel. The phonemes of one word (or more) can be rehearsed at the same time as a new word (or more)

arrives. If the phonemes are not identified right away, they can be parked in the storage subsystem. If they are identified but cannot be matched to a lexeme in Wernicke's area, they can go back to the loop and be parked again. If many words are rehearsed, need to go back, or come in simultaneously, they compete for attention as they are matched. If a word does not get the attention, it falls out of the loop. The loop can hold information for a couple of minutes (Baddeley, 2007). After each rehearsal there needs to be a pause or a distraction in order to stop the process and start a new process when the word is repeated. Therefore, uniform and expanded intervals that incorporate such pauses or distractions are more efficient and lead to higher recall rates (Cull, 2000; Carpenter & DeLosh, 2005; Landauer & Bjork, 1978).

Nation (2009) and Schmitt (2010) point out that although not much research using this technique was carried out in the discipline of second language acquisition, the findings were applied to second language vocabulary learning by creating flashcards. The earliest examples of these are Pimsleur's approach (1967) and Leitner's (1972) *Lernkartei*. These flashcards assist undergraduate students taking courses such as Beginning German to face the challenge of learning and remembering many new words in a short period of time. Most of the many flashcard apps that are available these days, either as part of or separate from textbooks or learning programs, offer a fun interface incorporating audio, images, and games. Unfortunately, few of them track a learner's progress or provide assistance as to how to acquire those words efficiently (Milton, Jonsen, Hirst, & Lindenburn, 2012; Schuetze & Weimer-Stuckmann, 2010). A notable exception is the "etutor" developed by Heift (2010).

A factor these flashcard systems often do not take into account is to differentiate between short-term gains and long-term retention (Roediger & Karpicke, 2010). The question is where to draw the line. A standard practice in experiments in cognitive psychology is to have participants recall information five, ten, or thirty minutes after words have been learned. But where do short-term gains end and where does long-term retention begin? A helpful insight in regard to this question comes from studies on sleep. Raichle showed that when our brain rests, it switches to a default state where some brain areas are very active and others are not. It is hypothesized that in that state the brain digests the information it encountered (Raichle, 2011). In order to test long-term retention, then, one has to wait at least 24 hours, ideally a few days. This notion is also supported by reports of language instructors (Oxford, 1990).

The distinction between short-term gains and long-term retention is important in second language acquisition where learners acquire words to use them weeks and month later (Schuetze, 2017). Testing needs to be done a day after the last practice but also weeks after. It should also be noted that learning words in another language usually occurs over multiple learning sessions. Each time the brain encounters a word, a new process of rehearsing that word begins. The question is which of the two intervals of the "spaced repetition" technique that lead to higher recall rates – uniform (for example, repeating words every two days) or expanded intervals (for example, repeating words on day one, two, five, nine) – is the most beneficial for students. This was the premise of a series of experiments carried out at a Canadian university in the fall of 2011 and 2012. Those experiments that are presented in a short summary in the following section investigated if the uniform or expanded interval lead to higher recall rates on short term as well as in the long run (Schuetze, 2015).

Context of Previous Experiments

In the fall of 2011 and 2012, a series of experiments were carried out investigating the uniform and expanded interval in more detail (Schuetze, 2015). Subjects were undergraduate students at a North American university at the West Coast of Canada enrolled in one of four

sections of a course called “Beginning German I.” A questionnaire asked students about their previous knowledge of German in order to have homogenous groups (for details, see the methodology section below). Students of two sections used the expanded and students of the other two sections the uniform interval. In the experiments each group was shown 39 words in PowerPoint. The number 39 had been chosen as students in the Beginning German I course need to acquire about 100 words in each chapter, 60 of those in comprehension and 40 in production. Balancing content to function words, 39 words were selected for the experiments (24 content and 15 function words). The words had been selected based on the fact that they would not be part of the textbook that was used for this course, so students would not encounter them outside the experiment. Naturally, it could not be completely ruled out that students did encounter some words outside the context of the study, however, the likelihood of that scenario was the same for all groups. The words chosen for the study were common words listed in Jones and Tschirner’s frequency dictionary (2006). In the 2011 experiments, the PowerPoint was shown to students four times. Students assigned to the expanded interval saw the words on days one, two, four, and eight. Students assigned to the uniform group on days one, four, eight, and eleven. Each time, each of the 39 words was shown for eight seconds: two seconds for the English word and six seconds for the German translation with the English word remaining on the screen. Choosing to show words for six/eight seconds was based on a method established in cognitive psychology studies (for an overview, see Balota et al., 2007). Including a two-second gap between each slide, the total time for the PowerPoint was just under seven minutes. The German words included sound. Students were asked to write down the German words on a piece of paper. After each trial, all paper was collected. Three tests were carried out in which students were given the English word and had to write down the German word. The first test was the day after the last trial, the second test four weeks after the last trial, and the third test eight weeks after the last trial. A multivariate analysis of variances was carried out for each experiment, analyzing the results of the expanded group, where 43 students had participated in each trial and each test, as well as the results of the uniform group, where 33 students had participated from start to finish.

The experiments in the fall of 2011 (Schuetze, 2015) showed that the expanded group recalled more words in the first test and the uniform group in the third test. On the second tests the two groups were equal. None of the differences were statistically significant, though. Recall rates were 36.4% (uniform group) vs. 47.4% (expanded group) in test 1, dropping to 31.1% (uniform group) vs. 21.4% (expanded group) in test 3. It was surprising how many words students had forgotten by the end of the language course although these numbers were in line with what Milton (2009) had already pointed out in his book on how to measure second language vocabulary acquisition. The experiments also showed that on each test each group recalled content words more successfully than function words and those differences were statistically significant. As interesting as these results were, the question came up if recall rates could be improved by increasing the number of encounters.

Another set of experiments was devised. The experiments in the fall of 2012 used the same methodology as those in 2011 with one exception. The number of encounters was extended to five. Thereby only students from two sections of Beginning German I were able to participate due to the class schedule. The group sizes were therefore smaller with 24 undergraduates in the expanded as well as in the uniform group. The schedule for the expanded group was days one, two, four, eight, and fifteen, and for the uniform group days one, four, eight, eleven, and fifteen.

Results (Schuetze, 2015) showed that the extra encounter gave the expanded group a boost on short-term gains, outperforming the uniform group on the first test and holding on to the higher recall rates on the second test, but on the third test students using the uniform interval

showed higher recall rates as they had in the 2011 study. As in the 2011 experiments, differences were not statistically significant. All recall rates were up compared to the 2011 experiments: in test 1 44.6% (uniform group) vs. 56.1% (expanded group), in test 3 36.9% (uniform group) vs. 30.8% (expanded group). As in the 2011 experiments, students in both groups recalled more content than function words, those differences being statistically different.

The next question that came up was how these recall rates might change if fewer words were encountered. The theoretical background was based on the model of working memory as described in the introduction. As Baddeley (2007) states, the capacity of the phonological loop is limited and if a word is not rehearsed in time, it is not processed any further. Having to process 39 words in less than seven minutes is challenging. The question then was if to halve the number of words would make a difference.

Another set of experiments was designed and is described in the following. The new research question was: Is it more beneficial to increase the number of repetitions (while keeping the number of words constant) or to reduce the number of words (while keeping the number of repetitions constant) in order to improve recall rates?

The pedagogical implication was to start with one half of the words in the first week of the chapter and the other half in the second week (each chapter being two weeks long).

Methodology

The study was carried out in the fall of 2013 with students enrolled in two of four sections of a Beginning German I course at a North American university at the West Coast of Canada. Students were divided into two groups by section with one group using a uniform interval and one group using an expanded interval. Students listened to and watched words on a Power-Point presented on a screen. All in all, twelve content and nine function words were presented in random order. The students' task was to copy each word on a piece of paper.

Participants

Students were undergraduates age 17 to 24. Students taking Beginning German I are supposed to have no or very little knowledge of German. In order to ensure that the students participating in the study were indeed beginners, students had to fill out a background questionnaire identifying previous knowledge of German, previous instruction in German, German heritage, and proficiency in other languages, and provided data on faculty, major, and years of study. Of the 55 students enrolled in two sections, 49 were allowed to participate once the questionnaires had been analyzed. Of those, 40 participated in the study from start to finish, that is, they took part in every of the five practices as well as the three tests. By chance the distribution in the two sections was equal with 20 participants in the expanded group and 20 participants in the uniform group.

Material

Twenty-one words (twelve content and nine function words) were selected for the study to be acquired within two weeks. This is less than what students usually have to learn in two weeks taking a beginner's German course that is taught three times a week at our university.

However, the study was designed to investigate if a reduced word load would improve recall rates. The previous experiments in the fall of 2011 and the fall of 2012 had used 24 content and 15 function words for a total of 39 words (Schuetze, 2015). For this study the number of content words was halved from 24 to 12. The number of function words was also reduced, but in order to have a balance between content and function words it was reduced from 15 to nine. Thereby four nouns, four verbs, and four adjectives were tested as well as three conjunctions, three prepositions, and three adverbs (see Appendix A for a complete list). The words were taken from the word list used in the previous studies in order to compare results. Words selected had been checked against the textbook used for Beginning German I to avoid that students encountered those words outside of the study. They were also checked to comply with Jones and Tschirner's dictionary (2006) of the most common used words in German.

Procedure

For each word a PowerPoint slide was created using the same font and same background color for all slides. On each slide the English word was presented on the left side of the screen. After two seconds the German equivalent word was presented on the right side and appeared for six seconds while the English word remained. The German word that appeared was spoken out loud. Each slide was shown for eight seconds. The reason to include sound was to reflect the situation of how a learner often encounters a word in another language.

Participants were asked to copy down the German word they saw and heard on a piece of paper. This was done as they had to write down the German words in the tests that followed, so the activity of writing needed to be practiced. Each PowerPoint was presented in the last ten minutes of class time. Those students not participating in the study left the classroom at that point. At the end of class, the research assistant collected all paper as it was done in the previous experiments. Naturally, in a classroom study not all factors can be controlled 100%. It was possible that a couple of students would encounter, either by chance or on purpose, the words practiced somewhere else. However, that chance was equal for both groups and since it was the performance of the groups as such that was compared, the risk of this happening was determined to be acceptable.

Both groups saw and heard the words of the PowerPoint four times. The previous study (Schuetze, 2015) had shown that the recall rates improved when the number of words was constant (39 words) but the number of repetitions increased by one (from four to five). Now the question was how the recall rates would change if the number of words was decreased (21 words) but the number of repetitions kept constant (at four). In the study presented here as well as in the previous studies, each time the order of the words was different to avoid testing effects; that is, the order of words primes the next word. Participants in the uniform group practiced the words on days one, four, eight, and eleven, whereas participants in the expanded group on days one, two, four, and eight. This was the same spacing schedule that had been used in the experiments in the fall of 2011.

Tests

Three retention tests were carried out in each study, one the day after the last practice (test 1), one four weeks after the last practice (test 2), and one eight weeks after the last practice (test 3). In each test, all of the twelve content and nine function words were tested for a total of 21

words. In all tests, the English word was given and participants had to write down the German word.

Limitations

In all studies, 2011, 2012, and 2013, it was possible that students did have contact with the words tested outside the study, for example by surfing for things German on the internet. However, based on the experience of other studies carried out on vocabulary acquisition at the beginners' level, the probability of such a situation was quite low due to the busy schedule of students taking five to six courses per term. Furthermore, the probability was the same for all groups.

The acquisition of a word in these studies was defined as to be able to produce the word in German when prompted by the English equivalent. This does not do justice to all aspects of vocabulary learning. However, it is beyond the scope of this paper to go into the details of word acquisition models such as Nation's multi-level model (2001). The focus of this paper is on rote learning and how it can benefit beginning learners such as undergraduate students enrolled in a beginning German course.

Results

Inferential statistics using a multivariate analysis of variances were carried out to determine if the reduced number of words would lead to higher recall rates on short-term gains as well as on the long-term retention, if there was a difference between content and function words, and if the type of spacing interval played a role.

Table 1. *Tests 1, 2, 3 Retention Scores of Content and Function Words*

	U Mean	SD	E Mean	SD	F	Sig
Test 1	7.95	5.49	8.28	5.27	.040	.842
Test 2	7.52	5.45	7.47	5.11	.001	.777
Test 3	7.13	5.53	5.61	4.97	.114	.297

The Mean refers to the average score of all participants in each group (U = uniform; E = expanded).

The effect size for all three tests was small (test 1: Cohen's d 0.061; test 2: Cohen's d 0.009; test 3 Cohen's d 0.289). On short-term gains, the expanded group had higher recall rates than the uniform group ($F(1,38) = .040$, $p > .05$) whereas on long-term retention it was the opposite ($F(1,38) = .114$, $p > .05$). In test 2, the scores were very similar ($F(1,38) = .001$, $p > .05$). None of the differences were statistically significant although the participants of the uni-

form group started with a lower score than those of the expanding group, but eight weeks later had retained more words. It would have been interesting to see the results of another test carried out another eight weeks later, but for practical reasons it was not possible to test the students again as the term was over. In the study presented here, the difference between the highest and lowest scores was not as distinct as in the previous studies, that is, the difference between the highest and lowest scores was more pronounced in the 2011 and 2012 studies. However, none of those differences had been statistically significant.

Table 2. *Tests 1, 2, 3 Retention Scores Uniform Group Content vs. Function Words*

	Content Mean	SD	Function Mean	SD	F	Sig
Test 1	10.58	6.74	5.22	3.41	6.153	.023*
Test 2	9.75	5.95	5.15	2.87	5.776	.027*
Test 3	9.35	6.20	4.84	2.69	5.190	.034*

* $p < .05$

Tables 2 and 3 show that participants working with the uniform interval as well as participants working with the expanded interval recalled more content words than function words. Results of all of the tests of the uniform group (test 1: $F(1,19) = 6.153$, $p < .05$; test 2: $F(1,19) = 5.776$, $p < .05$; test 3: $F(1,19) = 5.190$, $p < .05$) as well as of all of the tests of the expanded group (test 1: $F(1,19) = 5.075$, $p < .05$; test 2: $F(1,19) = 4.795$, $p < .05$; test 3: $F(1,19) = 4.250$, $p < .05$) were statistically significant. The effect size was large for both groups (uniform and expanded) in tests 1 and 2 for content words (uniform test 1: Cohen's d 1.003; uniform test 2: Cohen's d 0.984; expanded test 1: Cohen's d 1.192; expanded test 2: Cohen's d 1.063) and medium for function words (uniform test 1: Cohen's d 0.448; uniform test 2: Cohen's d 0.441; expanded test 1: Cohen's d 0.512; expanded test 2: Cohen's d 0.469). In test 3, the effect size was small for the uniform group (content words: Cohen's d 0.189; function words: Cohen's d 0.094) and medium for the expanded group (content words: Cohen's d 0.494; function words: 0.445). This confirmed the results of previous studies that function words are more difficult to remember (Schuetze, 2015).

Table 3. *Tests 1, 2, 3 Retention Scores Expanded Group Content vs. Function Words*

Sig	Content Mean	SD	Function Mean	SD	F	
Test 1	11.03	5.70	5.52	3.19	5.075	.036*
Test 2	9.84	5.66	5.08	2.84	4.795	.041*
Test 3	7.76	5.94	3.47	1.39	4.250	.043*

* $p < .05$

In summary, the results show using either interval, uniform or expanded, the retention rates for content words are significantly higher than those for function words. However, the effect diminishes with time. Comparing the uniform with the expanded group for all words (content and function words combined), recall rates were 37.8% (uniform group) vs. 39.4% (expanded group) in test 1 dropping to 33.9% (uniform group) vs. 26.7% (expanded group) in test 3. These test retention rates were similar to the 2011 study despite the fact that fewer words had to be acquired compared to that study. It is an indication that reducing the number of words did not improve recall rates. Furthermore, the rates were lower compared to the 2012 study as the 2011 rates had been lower compared to the 2012 rates as well. In the 2012 study, participants had one more repetition boosting their retention rates on short-term gains as well as on long-term retention.

Discussion

In the context of rote vocabulary learning in a beginning German course, the research question presented in this study was: In order to improve retention rates, is it more beneficial to increase the number of repetitions (while keeping the number of words constant) or to reduce the number of words (while keeping the number of repetitions constant)? The experiments presented here showed that reducing the number of words was not effective. In order to prevent forgetting, that is, that students only recall a fraction of those words they learned previously a couple of months later, increasing the number of repetitions seems to be beneficial. A previous study that had increased the number of repetitions from four to five had shown that recall rates went up.

A possible explanation lies in the phonological loop itself. Baddeley (2007) proposed that the loop consists of two subsystems: a rehearsal and a temporary storage. The rehearsal is part of Broca's area in the left prefrontal cortex and the temporary storage sits in the left side of the parietal lobe of the brain. If there is not enough time to process a word right away, it can be parked in the temporary storage. That way the trace of the word that usually expires in less than a second can be prolonged. The previous experiments showed that repeating this process several times increases the chances of the word being processed, in particular if a word is encountered five times and not four. Reducing the number of words, however, seems to have not much impact, indicating that the interplay of the temporary storage with the rehearsal is of more relevance than the sheer number of words to be handled, although naturally there is a limit. Therefore, a good strategy for a language learner to improve the recall rate is to increase the number of repetitions. For second language instructors it might be helpful to know that asking students to study fewer words based on the premise that the brain cannot handle large amounts of words at once might not be the best option.

It should be noted that these experiments were carried out in the context of a university where undergraduate students with no previous knowledge of a second language studied German for the first time. Therefore, it should not be concluded that practicing words five times is a magic number for second language vocabulary learning. It proved to be a number that worked in the context described. In a different learning context with learners of a higher proficiency level or acquiring words in an environment where German (or another modern language) is spoken pre-dominantly, which is something that cannot be said for the West Coast of Canada, the number might be different. For languages with a different sound system, for example Korean, Japanese, Cantonese, or Mandarin, the process is different altogether as the main task for a learner is to identify phonemes and match them to lexemes.

References

- Baddeley, A.D. (1986). *Working Memory*. Oxford: Oxford University Press.
- Baddeley, A.D. (2007). *Working Memory, Thought, and Action*. Oxford: Oxford University Press.
- Baddeley, A.D., & Hitch, G. (1974). Working memory. In G.H. Bower (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (Vol. 8, pp. 47-89). New York: Academic Press.
- Balota, D., Duchek, J., & Logan, J. (2007). Is expanded retrieval practice a superior form of spaced retrieval? A critical review of the extant literature. In J. Nairne (Ed.), *The Foundations of Remembering* (pp. 83-105). London: Psychology Press.
- Barcroft, J. (2009). Effects of synonym generation on incidental and intentional vocabulary learning during second language reading. *TESOL Quarterly*, 43, 79-103.
- Carpenter, S.K., & DeLosh, E.L. (2005). Application of the testing and spacing effects to name learning. *Applied Cognitive Psychology*, 19, 619-636.
- Cook, V., & Singleton, D. (2014). *Key Topics in Second Language Acquisition*. Bristol: Multilingual Matters.
- Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyaki & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 62-101). Cambridge: Cambridge University Press.
- Cull, W. (2000). Untangling the benefits of multiple study opportunities and repeated testing for cued recall. *Applied Cognitive Psychology*, 14, 215-235.
- Ellis, N. (1995). Vocabulary acquisition: Psychological perspectives. *The Language Teacher*, 19(2), 12-16.
- Gathercole, S.E., & Alloway, T.P. (2008). *Working Memory and Learning. A Practical Guide for Teachers*. London: Sage Publications.
- Gathercole, S.E., Pickering, S.J., Hall, M., & Peaker, S. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. *Quarterly Journal of Experimental Psychology*, 54A, 1-30.
- Heift, T. (2010). Developing an Intelligent Language Tutor. *CALICO Journal*, 27(3), 443-459.
- Hulstijn, J. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In P. Arnaud & J. Béjoint (Eds.), *Vocabulary and Applied Linguistics* (pp. 113-125). Houndmills: Macmillan.
- 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.
- Jones, R.L., & Tschirner, E. (2006). *A Frequency Dictionary of German*. London: Routledge.
- Landauer, T.K., & Bjork, R.A. (1978). Optimal rehearsal patterns and name learning. In M. Gruneberg, P. Morris & R. Sykes (Eds.), *Practical Aspects of Memory* (pp.625-632). London: Academic Press.
- Leitner, S. (1972). *So Lernet man Lernen: Der Weg zum Erfolg*. Freiburg: Herder.
- Milton, J. (2009). *Measuring Second Language Vocabulary Acquisition*. Bristol: Multilingual Matters.
- Milton, J., Jonsen, S., Hirst, S., & Lindenburn, S. (2012). Foreign language vocabulary development through activities in an online 3D environment. *The Language Learning Journal*, 40(1), 99-112.
- Mondria, J.A., & Mondria-De Vries, S. (1994). Efficiently memorizing words with the help of word cards and "hand computer": Theory and applications. *System*, 22(1), 47-57.
- Nation, I.S.P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Nation, I.S.P. (2009). *Teaching Vocabulary. Strategies and Techniques*. Boston: Heinle.
- Oxford, R.L. (1990). *Language Learning Strategies: What Every Teacher Should Know*. Boston: Heinle.
- Pimsleur, P. (1967). A memory schedule. *The Modern Language Journal*, 51, 73-75.
- Raichle, M. (2011). The restless brain. *Brain Connectivity*, 1, 3-12.
- Roediger, H.L., & Karpicke, J.D. (2010). Intricacies of spaced retrieval: A resolution. In A.S. Benjamin (Ed.), *Successful Remembering and Successful Forgetting: Essays in Honor of Robert A. Bjork* (pp. 1-36). New York: Psychology Press.
- Schmitt, N. (2010). *Researching Vocabulary*. Basingstoke: Palgrave Macmillan.
- Schuetze, U. (2017). *Language Learning and the Brain*. Cambridge: Cambridge University Press.
- Schuetze, U. (2015). Spacing techniques in second language vocabulary acquisition: Short-term gains vs. long-term memory. *Language Teaching Research*, 19(1), 28-42.
- Schuetze, U., & Weimer-Stuckmann G (2010). Virtual vocabulary: Research and learning in lexical processing. *CALICO Journal*, 27, 517-528.
- van Zeeland, H., & Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening. *System*, 41, 609-624.

Appendix A. Word List

Content words (nouns / verbs / adjectives): 12 total

Nouns:

der Turm [the tower]

das Sonnenschutzmittel [the sunscreen lotion]

die Liebe [the love]

der Berg [the mountain]

Verbs:

malen [to paint]

beabsichtigen [to intend]

weinen [to cry]

rauchen [to smoke]

Adjectives:

tief [deep]

notwendig [necessary]

verrückt [crazy]

kühl [cool]

Function words (adverbs / conjunctions / prepositions): 9 total

Adverbs:

vielleicht [perhaps]

schon [already]

wirklich [really]

Conjunctions:

nachdem [after]

deshalb [thus]

obwohl [although]

Prepositions:

seit [since]

zwischen [between]

jenseits [beyond]