Linnæus University
Sweden

Fill-in-the-Blank or Write an Original Sentence

A Comparison of Practice Materials for Vocabulary Retention

Author: John Lindström
Supervisor: Christopher Allen
Examiner: Charlotte Hommerberg
Term: VT20
Subject: English
Level: BA
Course code: 2ENÄ2E

Linnæus University
Sweden

Abstract

Despite extensive research, a “best method” for teaching vocabulary has not been identified. The present study investigates the efficacy of two different practice materials for vocabulary retention, a fill-in-the-blank exercise and an original sentence writing exercise, from the hypothesis that a fill-in-the-blank exercise will be superior in terms of vocabulary retention on an immediate posttest. By testing this hypothesis, the study aims to contribute to the understanding of what is best practice in vocabulary teaching. The participants were 20 adult EFL learners studying at *Basic Adult Education* level. A controlled experiment using a Presentation, Practice, Production (PPP) structure was used. The cartoon public service announcement *Dumb Ways to Die* was used as the medium for presentation of twelve level and frequency scanned target nouns, followed by 10 minutes of practice time with either a fill-in-the-blank exercise or an original sentence writing exercise. Finally, an immediate posttest ensued. The result shows that the participants who practised using the fill-in-the-blank exercise performed better on the posttest, albeit not significantly better. In addition, further analysis indicates that some items of target vocabulary may have been previously known by the participants or easier to guess the meaning of. The pedagogical implications point towards using fill-in-the-blank exercises in the initial stages of vocabulary learning, after target vocabulary has been presented, while original sentence writing exercises appear better suited in later stages when more aspects of word knowledge have been acquired.

Key words

Vocabulary Acquisition and Retention, Practice Materials, Adult EFL Learners, Fill-in-the-Blank, Original Sentence Writing

Acknowledgments

Mr. Rune Lindström (1925-2009)

Table of contents

Appendices

1 Introduction

"The fact is without grammar very little can be conveyed, without vocabulary nothing can be conveyed" (Wilkins 1972). This quote reflects the fact that vocabulary may be the most important part of language to learn. Without sufficient vocabulary, an individual cannot express his/her thoughts, ideas and opinions. Simply put, language is communication, and without an adequate number of words, an individual cannot communicate. What is more, no other group of learners are as lacking in vocabulary knowledge as those who begin learning English as adults (Nation & Waring 1997:6). Indeed, many English as a Foreign Language (EFL) teachers realise this and put their best efforts into facilitating vocabulary acquisition and retention. However, while the research on vocabulary acquisition is extensive, due to the many factors involved in vocabulary teaching and learning, a clear-cut "best way" to teach vocabulary has failed to manifest itself. As a result of this, syllabi fail to provide teachers with explicit instruction on how to teach vocabulary, resulting in teachers sometimes being insecure in which methods are the most effective ones (Schmitt 2008:329).

In Sweden as of 2017, approximately 15,000 adults study English at a level equivalent to secondary school (National Agency for Education 2018) However, while database searches yield several examples of vocabulary acquisition studies on Swedish high school students (for instance; Sylvén 2004 and Olsson 2015), studies on adult EFL learners appear to be far less frequent in a Swedish context. This is unfortunate, since adults tend to benefit from somewhat different teaching methods than those used to teach children and teenagers (Harmer 2007:85). Thus, the question which methods are the most effective ones for teaching vocabulary to the group most in need of vocabulary expansion, the adult beginner/intermediate EFL leaners, has not been sufficiently answered, at least not in a Swedish context with the

heterogeneous groups studying *Basic Adult Education*. Consequently, this study hopes to assist Swedish adult EFL teachers, by investigating the effectiveness of two different practice exercises designed to facilitate vocabulary retention.

The choice to focus on practice exercises rather than different methods of presenting vocabulary is due to the pivotal role that repetition plays in language learning as well the results of recent research in the field. To begin with, whereas a teacher undoubtedly will put their best efforts into presenting new vocabulary in a way that engages the students' interest, it takes more than presenting if the students are to learn new vocabulary. On the contrary, a new vocabulary item needs to be met somewhere from five to over 20 times for the students to "learn" it (Nation 2001:81). Consequently, recycling by means of practice exercises is crucial to facilitate vocabulary retention. Furthermore, historically, the prevalent perception has been that the more cognitive engagement an exercise triggers, the more effective it is (Craik & Tulving 1975, Brown & Perry 1991, Hulstijn & Laufer 2001), and that this holds true for presenting new vocabulary as well as practising already presented vocabulary. However, more recent studies by Folse (2006) and Hashemzadeh (2012) found that fill-in-the-blank exercises led to better retention than other, more creative, exercises such as an original sentence writing exercise, thus suggesting, in terms of practice materials, that the number of word retrievals means more for retention than cognitive depth. These are compelling findings, since they suggest that practice materials should above all prioritise repetition. Consequently, based on these findings, it seems that fill-in-the-blank exercises could be superior to original sentence writing exercises.

1.1 Aim and Scope

The primary aim of this paper is to compare the efficacy of two different practice exercises designed for vocabulary retention, a fill-in-the-blank exercise and an original sentence writing exercise, by testing the following hypothesis.

- After target vocabulary has been introduced, a fill-in-the-blank exercise repeated once will be superior to an original sentence writing exercise repeated once, in terms of vocabulary retention on an immediate posttest.

To test the hypothesis, a controlled experiment in a classroom environment is performed where target vocabulary is presented, practiced and analysed on a group- and individual level. Furthermore the secondary aim of this study is to have the result contribute to more general understanding of what is best practice when teaching vocabulary. However, the scope of the study is strictly limited, since the small sample size of participants all come from the same school, and are only tested on one occasion. Consequently, this study does not draw any wider conclusions based on the result.

2 Background

The background provides a general overview of key concepts involved in vocabulary research, followed by a more specific account of studies pertaining to EFL learners' vocabulary acquisition, retention and practice exercises. To begin with, the terms *word*, *lemma* and *word families* are explained. Next, vocabulary learning goals in terms of number of word families is discussed, followed by an account of the different aspects of word knowledge. Subsequently, the background details the selection of which words to teach along with information concerning the adult EFL learner, and individual differences related to language learning. Finally, a sample of

relevant studies on vocabulary acquisition and retention is presented, together with specific research on practice exercises.

2.1 Defining and Grouping words

Linguists and philosophers have long attempted to define what a word is, making it exceedingly difficult to select a single universal definition. However, this paper chooses the reasonable and oft-quoted definition of a word as "a single unit of language that has meaning and can be spoken or written" (Cambridge Dictionary 2019). Moreover, words can be grouped into either lemmas or word families. A lemma is a headword, meaning the form of a word that appears as a heading in a dictionary (i.e *walk*, *cat* or *big*), some of its inflections (*walked/walks/walking, cats/cat's* and *bigger/biggest*) and close derivations. However, words grouped under the same lemma are usually confined to the same part of speech (Nation 2001:7). Word families, on the other hand, contain a headword, all of its inflections and more of its close derived forms, taking affixes like *-ly*, *-ness* and *-un* into account (Nation 2001:8). In general, the term word families is used when discussing a learner's vocabulary size, and every piece of target vocabulary used in this study originates from a separate word family.

(1)

2.2 Learners' Knowledge of Word Families

Whereas an educated native speaker of English has a receptive vocabulary size of around 20,000 word families, the majority of adult EFL learners will never reach that number. Instead, the number of word families a learner needs is dictated by the individual needs and goals of the learner. To begin with, a learner who has a receptive vocabulary size of 5000 word families including the 3000 most high frequency words has an ample base for comprehension and production of English. (Nation & Waring 1997:6).

Moreover, to be able to read and understand authentic material such as novels or newspapers, 8-9000 word families may be required. However, the aforementioned figures are not valid if a learner does not know all of the words belonging to one word family. That is to say, a learner may know the word *person*, yet fail to fully understand the meaning and use of the derivative form *personally* (Schmitt 2008:331). Consequently, while grouping words into word families may facilitate the description of a learner's vocabulary size, there are several aspects involved in knowing a word

2.3 Knowing a Word

Knowing a word entails more than being able to provide a translation in another language. Naturally, the first step in the acquisition process is establishing the form-meaning link. In other words, a learner encounters a new word, learns its meaning either by translation, a picture or an explanation. However, this type of recognition knowledge needs to be expanded if the learner is to have full receptive and productive control over the word (Schmitt 2008:333). The table below (Nation 2001:27 cited in Schmitt 2008:334) details the different aspects of word knowledge.

Table 1. Aspects of Word Knowledge

Form	Spoken	R* What does the word sound like? P** How is the word pronounced?
	Written	R How does the word look? P How is the word written and spelled?
	Word Parts	R What parts are recognizable in the word? P What word parts are needed to express this meaning?
Meaning	Form and Meaning	R What meaning does this word form signal? P What word form can be used to express this

		meaning?
	Concepts and Referents	R What is included in the concept? P What items can the concept refer to?
	Associations	R What other words does this make us think of? P What other words can we use instead of this one?
Use	Grammatical Functions	R In what patterns does the word occur? P In what patterns must we use this word?
	Collocations	R What words or types of words occur with this one? P What words or types of words must we use with this one?
	Constraints on Use (Register… Frequency...	R Where, when and how often would we expect to meet this word? P Where, when and how often can we use this word?

** R = Receptive knowledge ** P = Productive knowledge*

Indeed, a multitude of aspects are involved in knowing a word. Naturally, the experiment performed in this study does not claim to teach the participants every aspect of word knowledge of target vocabulary. Consequently, throughout this paper, the word *learn* is written within inverted commas as "learn", when it refers to word knowledge.

2.4 Selecting which Words to Teach

Nation and Waring (1997:11-12) suggests teaching high frequency vocabulary first, due to its profound impact on comprehension and production. To begin with, the 2000 most common words in English make up around 80% of written English, and around 90% of spoken discourse (Nation 2001:17). Thus, teachers are recommended to consult frequency lists for guidance in which vocabulary to focus on. The most well-known frequency

list is the General Service List (GSL) (West 1953). It contains the (then) 2000 most frequent word families. In 2013, an updated version called New GSL was published. NGSL is expanded to 2633 word families and provides a vocabulary coverage of approximately 90% of the Cambridge English Corpus (Brown 2013:13-16). Consequently, NGSL provides an ample starting point for any learner. Furthermore, once the learners have acquired basic vocabulary, Nation and Waring (1997:11-12) suggest teaching strategies for acquiring low frequency -, academic-, and technical vocabulary. That is to say, learners need to be able to use dictionaries, guess the meaning of a word based on context and clues in word formation.

2.5 Teaching Vocabulary to Adult Learners

Granted, vocabulary can be taught indirectly (incidental vocabulary learning). However, "the main reason for an explicit focus on vocabulary is that it is effective" (Schmitt 2008:341). In other words, students learn more vocabulary from explicit methods than from incidental ones.
In terms of a vocabulary focused lesson, Arju (2011:57) suggests an order of setting the context, eliciting, drilling, concept questioning and finally a practice activity as a general template, in other words, a rather traditional structure of presentation, practice and production (PPP). Harmer (2007:67) acknowledges the critique against PPP as a rather traditional method of teaching, yet maintains that it "is a very effective way of teaching small numbers of individual words at beginner level" (2007:230). For the adult learner, Harmer highlights the value of "allow[ing] them to use their intellects to learn consciously where this is appropriate", and "to pay special attention to the level of challenge presented by [activities]" (2007:85). Consequently, an adult beginner EFL learner will most likely benefit from a rather traditional explicit approach to vocabulary teaching, combined with activities that do not exceed their current language level by too great of a margin, while still allowing them to use their intellect and life experience.

However, regardless of method, successful vocabulary acquisition and retention will also be affected by individual differences.

2.6 Individual Differences

A variety of aspects influence the proficiency level ultimately attained by an individual learner. Figure 1 illustrates a continuum of how these aspects affect the outcomes of language learning.

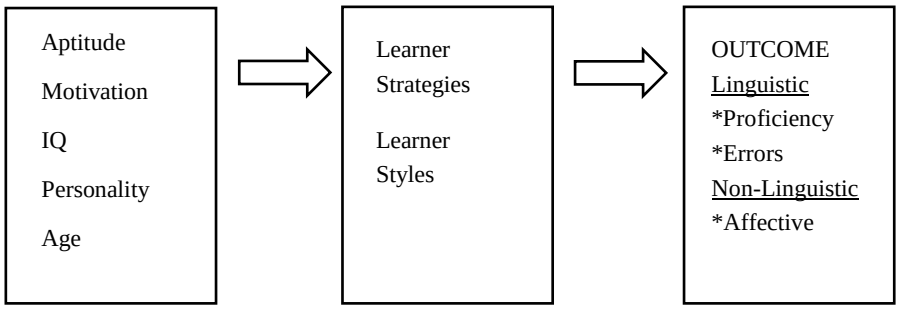

Figure 1. Influences on Language Learning (Skehan 1991:277)

To begin with, the first square in the figure shows an individual's abilities and prerequisites in terms of talent for language learning (aptitude), motivation, IQ, personality and age. Next, square two demonstrates how a learner's preferred strategies and unique learner style also affect the linguistic and non-linguistic outcomes shown in square three. In fact, due to the substantial impact of individual differences on language learning in general, extensive research has been conducted on all the aforementioned factors (see; Skehan 1991, Dörnyei, Z. 2014 for further reading). However, the relevance of those studies in regards to explaining the design of this study is limited. Suffice to say, individual differences related to language learning undoubtedly affect the result of an experiment such as the one presented in this paper.

2.7 Theories behind Vocabulary Acquisition and Research on Practice Exercises

Effective second language vocabulary acquisition and retention can be summarized in the idea that any exercise leading to more time spent on and mental energy invested in target vocabulary is likely to facilitate acquisition and retention (Schmitt 2008:339). However, not every exercise is equally effective. To begin with, different exercises are assumed to activate different levels of processing in the mind. Consequently, exercises that stimulate semantic processing of a word are more likely to facilitate retention than exercises that stimulate phonemic or structural processing of the same word (Craik & Lockhart 1972, Craik & Tulving 1975). That is to say, exercises that require the learner to consider the meaning of a word and determine whether it fits in a context or not are assumed to be superior to exercises that solely concern the form or the pronunciation of a word. However, the participants in Craik and Tulving's studies were native speakers of English and not EFL students. Studies have also been carried out to test the applicability of the levels of processing theory on EFL teaching and learning. Brown & Perry (1991) found a combined semantic and L1-keyword method to be superior to L1-keyword alone or the semantic method alone, thus to some extent confirming the idea that deeper processing leads to better retention than shallower processing. In addition, Hulstijn (1992) concluded that inferring the meaning of words (more mental effort i.e. deep processing) provided better retention than meaning given by synonym or translation (less mental effort i.e. shallow processing). Later, the importance of mental effort for vocabulary acquisition and retention was developed into the *Involvement Load Hypothesis* (Hulstijn & Laufer 2001). In short, the hypothesis provides a way to rank exercises based on their cognitive and motivational effect on the learner (Hulstijn & Laufer 2001:2). According to the ranking provided by the *Involvement Load Hypothesi*s, an original-sentence writing exercise

outranks a non-original sentence exercise for reading comprehension activity by one (1) point, when both exercises provide marginal glosses (Hulstijn & Laufer 2001:18). However, the hypothesis is more geared towards incidental learning as opposed to explicit, and more towards introducing/presenting vocabulary as opposed to practising it.

Indeed, when comparing practice exercises, Folse (2006) tested an original sentence writing exercise, a fill-in-the blank exercise repeated once and a fill-in-the-blank repeated three times, and found the fill-in-the blank exercise repeated three times to be significantly better for vocabulary retention. However, on average, the participants of Folse's study spent 50% more time of their allotted 40 minutes completing the three fill-in-the-blank exercise than they did completing the original sentence writing exercise. Thus, initially, the result appeared to be a consequence of more time on task (Folse 2006:286). However, further analysis of 31 participants who had spent a relatively equal amount of time on each exercise still found that those participants tested better on the vocabulary they had encountered in the fill-in-the-blank exercise repeated three times (Folse 2006:286). As a result, Folse (2006:287) claimed that the number of word retrievals means more for retention than a deeper level of processing, as the fill-in-the-blank exercise repeated three times offered more repetition of the target vocabulary than the original sentence writing activity within the same time frame. However, while Folse found no difference between a fill-in-the-blank exercise repeated once and an original sentence writing activity repeated once, Hashemzadeh (2012) found that a fill-in-the-blank exercise repeated once provided better retention than three other types of written exercises repeated once, both on immediate and delayed tests. In addition, Hashemzadeh used a narrower time frame than Folse did, as the participants were given 25 minutes to read a text and work with one of the exercises in four separate sessions, thus adding to the validity of a fill-in-the-blank exercise as an efficient option in terms of

number of words retained per minute. Granted, Hashemzadeh (2012:1724) admits that the fill-in-the blank exercise was the one that the participants were most familiar with, and that may have affected the result. Nevertheless, it seems that retention is facilitated by practice activities that keep the focus on the target vocabulary (i.e. fill-in-the-blank or matching pictures with words), rather than involve other areas of language proficiency such as grammar (i.e. writing original sentences/compositions or answering questions). Furthermore, it becomes evident that efficiency in terms of number of words retained per minute has to be measured when comparing practice exercises. In other words, time on task needs to be controlled by limiting practice time. Granted, an alternative would be to have the participants note time spent on task and correlate that number with their score on the posttest. However, from personal experience from years of second language teaching to adults, the time that an individual learner will take to complete a practice exercise will vary greatly. Thus, forcing early finishers to wait for every last person to complete their material runs the risk of aggravating the participants, and does not coincide with the ethical aim of this paper in providing a worthwhile and informative experience (see ethical considerations).

In summary, whereas cognitive and motivational aspects appear to play a large part when target vocabulary is presented, focus on task appears crucial for retention when target vocabulary is practiced. However, none of the aforementioned studies have put this hypothesis to the test in a Swedish setting on adult EFL learners. Thus, this study aims to adapt the premise of Folse's and Hashemzadeh's studies and compare a fill-in-the-blank activity to an original sentence writing activity in terms of efficiency on vocabulary retention.

3 Materials and Method

In order to compare the efficacy on vocabulary retention of a fill-in-the blank exercise and an original sentence writing exercise, while providing a adequate learning experience, an engaging presentation material and 12 target nouns were selected, two practice materials and a posttest were designed. Next, a controlled experiment was conducted using adult EFL learners studying *Basic Adult Education* as participants. The data was then analysed using quantitative methods. The following sections provide an exhaustive account of the participants of the study, the material and method used and their possible limitations.

3.1.1 Participants

The participants of the study were adult EFL learners studying at the municipal adult education. The participants study a program called *Basic Adult Education* which consists of courses in Swedish (alternatively Swedish as a Second Language), Civics, Mathematics and English. The subject of English is divided into 4 separate courses with levels ranging from CEFR level A1 in course one to approximately B1 at the end of course four (National Agency for Education 2019). The participants of this study were enrolled in course two and three. The reasons for this selection of participants were that the level of courses two and three corresponded best to the chosen presentation material and the aim to compare activities for vocabulary retention. That is to say, the participants needed to have basic vocabulary in order to be able to understand the presentation and complete the practice activities. Consequently, course one was not relevant, since it accommodates absolute beginners. In addition, the level of course four was deemed too high for the target vocabulary.

Table 2. Participant Sample Information[1]

N (F)*	Age** (SD)***	Number of Years in English Studies (SD)
20 (14)	29.5 (8.9)	4.8 (3.3)

** The number of females within brackets*
*** Age and Number of Years noted as Mean (average) numbers*
**** SD = Standard Deviation of the Mean, within brackets.*

In total, 20 EFL students participated in the study, 6 males and 14 females. In terms of age, native language and background in English studies, the participants were indeed a heterogeneous group. That is to say, the oldest participant was 55 while four 20-year-old participants were the youngest. The most common native language among the participants was Arabic (7) followed by Dari (4). Other languages represented were Persian (2), Swahili, Thai, Polish, Saho, Turkish and Pashto (1, respectively). Furthermore, one participant stated twelve years of previous English studies, whereas one participant had only six months of experience of learning English in a formal context.

3.1.2 Materials

In an attempt to create as engaging a presentation as possible, the short film *Dumb Ways to Die* (DumbWays2Die 2012) was selected. Originally, the film was ordered by Melbourne Metro as a public service announcement (PSA) with the purpose of reducing train related accidents. However, rather than keeping with the sombre and realistic tone prevalent in PSAs, the filmmakers opted for a surrealistic and facetious approach (Allagui & Breslow 2016:6). The movie contains a song which describes various ways to die, such as taking one's helmet off in outer space or setting fire to one's hair. The titular

[1] The terms Mean and Standard Deviation of the Mean are explained in detail in the Results section.

score of the film is illustrated by cartoon characters shaped as jelly beans who eventually all join together to sing the chorus in their disfigured forms. As an added twist, the film's purpose does not become clear until the final verse, when the jelly beans are killed due to being careless around trains, and the final message of "Stay safe around Trains" appears on screen. Indeed, the film has been a success, amassing over 183 million views on Youtube (as of 2019), and helping to reduce train related accidents in Melbourne by 21% within the first six weeks of its ensuing campaign (Allagui & Breslow 2016:6). A large part of the film's success can be attributed to the emotional response it triggers. Consequently, despite the rather gruesome theme, the film was considered optimal in terms of creating an engaging presentation.

Next, the lyrics of the song were transcribed and all the concrete nouns were extracted. The choice of concrete nouns over other parts of speech was due to the fact that concrete nouns are more easily illustrated with pictures than for instance abstract nouns or phrasal verbs, and for this study a picture based practice and test material was chosen, rather than a translation-based test which would have required resources above those available to this study. In the next step, the nouns were sorted on the basis of their absence from the New General Service List and CEFR level using the English Vocabulary Profile (Cambridge University 2019), resulting in 20 nouns not accounted for in NGSL and at the B1 level or higher being selected. However, an additional search for frequency using COBUILD (1995) found eight (8) of the 20 nouns to be too frequent for inclusion. Consequently, based on frequency and CEFR level, twelve (12) nouns thought to be unfamiliar to learners at the level of the participants were selected.

Table 3. Target Vocabulary

Nouns	CEFR - Level*	COBUILD Freq. Rating**	NGSL ***
bait	N/A***	■ ■ □ □ □	N/A
pie	B1	■ ■ □ □ □	N/A
helmet	B2	■ ■ □ □ □	N/A
dryer/drier	N/A	■ □ □ □ □	N/A
kidney	C2	■ ■ □ □ □	N/A
glue	C2	■ ■ □ □ □	N/A
moose	N/A	■ □ □ □ □	N/A
nest	C2	■ ■ □ □ □	N/A
wasp	B2	■ □ □ □ □	N/A
spine	C1	■ ■ □ □ □	N/A
ash	N/A	■ ■ □ □ □	N/A
skeleton	B2	■ ■ □ □ □	N/A

* CEFR = Common European Framework for Language

** COBUILD frequency rating

*** NGSL = New General Service List

****N/A = Not Listed

Naturally, the level and frequency scan could not guarantee that all items of target vocabulary where unknown to the participants at the time of the experiment. Consequently, an additional point of interest in the analysis was to see to what extent the selection of target vocabulary had been successful.

In the next step, practice material A, a fill-in-the blank activity and practice material B, an original sentence writing activity were designed (Appendix 1), both consisting of two pages. To begin with, an identical first page for both activities was created. The first page contained the 12 target nouns together

with a picture illustrating each noun, whereas, on the second page, practice material A contained a fill-in-the blank activity and practice material B an original sentence writing activity. Next, the test (Appendix 1) reused the pictures from the practice material and had the learners attempt to write the correct noun beside the corresponding picture. Finally, preliminary tests of the materials and test was conducted by the author of this paper, and ten minutes was chosen as the allotted practice time, along with five minutes to complete the posttest. The assumption was that most participants working with material A would complete it within five minutes, and that ten minutes would be sufficient for most participants working with the more time consuming material B. However, having never met the participants, and knowing very little regarding their individual proficiency in writing, the risk of some participants not completing their material was noted.

3.1.3 Limitations of the Material

One limitation of the material concerns the participants' possible previous knowledge of target vocabulary. Thus, a valid alternative would have been to use a pretest of the 12 nouns from the film. A pretest can either be done by a similar group of learners or by the participants themselves as a prelude to the actual experiment, a method used by Hashemzadeh (2012:1720). However, finding a similar group of learners studying Basic Adult Education is nearly impossible, due to the vast differences in background, age, native language and proficiency level prevalent at that type of school. In addition, a target vocabulary pretest on the participants of this study would possibly have resulted in a new, time-consuming selection of replacement target vocabulary, time in which the participants could well have learned new vocabulary unknown to the author of this study. Thus, due to the allotted time of this study, the choice was made not use a pretest.

3.2 Method

Due to the deductive and analytical nature of this study, an experimental design was chosen as the method for collecting data. Seliger and Shohamy (1989:136-137) describe the three components of an experiment as *the population, the treatment,* and *measurement (of the treatment).* Firstly, *the population* in this study are what Seliger and Shohamy (1989:136) refer to as a "natural group". In other words, the participants are a group of students at Basic Adult Education level, as opposed to a random group of people brought together for the purpose of this study. Moreover, *the treatment* is what is known as the independent variable in an experiment. That is to say, *the population* receives a "*controlled* and *intentional* experience, such as . . . materials presented under controlled circumstances . . ." (Seliger & Shohamy 1989:137). Consequently, the treatment, as it pertains to this study, consists of, either practice material A or, practice material B, Finally, *the measurement,* is the tool used to determine the effects of the treatment. Thus a *measurement* could be a written test, recognition test, or an oral assignment (Seliger and Shohamy 1989:137). Thus, to compare the efficacy of practice material A and B on vocabulary retention, this study used an immediate posttest as the measurement.

The experiment was carried out in a classroom environment with researcher control over the sequence of events. In light of the chosen method and the aforementioned success of the PPP structure when teaching vocabulary to adult EFL learners, the structure of the experiment was 1) presentation of target vocabulary (approximately twelve minutes) 2) time on task controlled practice with either practice material A or B (ten minutes), and finally 3) Immediate posttest (five minutes). A more detailed account of the collection of data, limitations of the method, ethical considerations and finally analysis of the data is detailed in the sections below.

3.2.1 Collection of Data

The data was collected from a total of 20 participants on two separate occasions in a Swedish municipal adult education school by the author of this paper and with assistance from the teacher present. The first group of participants studied course three of the basic adult education programme and consisted of twelve students. The second group of participants were eight in total, and studied course two. The second group initially had nine students who agreed to participate. However, one participant had to abort the study due to personal reasons, and is thus not accounted for. On both instances the collection of data followed the same order. To begin with, the purpose of the study was explained orally, before the written information (Appendix 2) was handed out. The participants were informed that participation was voluntary and anonymous. Furthermore, the participants were given the opportunity to ask questions concerning the study and received information about whom to contact if any further questions arose (Appendix 2). Next, a consent form (Appendix 2) was handed, out, read and signed by the participants, and then the presentation of target vocabulary ensued. Firstly, the purpose and success of the short film as a PSA was explained. Then the film was shown twice, once uninterrupted and once with pauses were target vocabulary was highlighted, explained and written on the whiteboard. In the two showings of the film and the explanations given, every item of target vocabulary was repeated between five to six times. After the second showing, the participants received either practice material A or B. The teacher present assisted in distributing the material so that material A and B were evenly spread amongst students of different proficiency levels. Next, the participants had ten minutes to work with their practice material, during which they were allowed a dictionary of their choice as assistance. After ten minutes, the

practice material was collected and the participants were instructed to remove their notes and cell phones.

When the participants had cleared their benches, the double-sided test (Appendix 1) was handed out with the page requesting background information face up. After all the participants had filled in their age, gender, number of years of English studies and which practice material they had completed, they were instructed to turn the paper over and take the test. The allotted time for the test was five minutes and it was completed under silence and without assistance. In total, eleven participants practiced using material A, and nine participants practiced using material B. The reasons for this uneven distribution is that the participant in the second group who had to abort the study did so in the midst of practicing with material B, and a slight miscount when the materials were distributed in the first group. In addition, an unexpectedly low completion grade of the practice materials was found as no participant practicing with material B was able to complete twelve original sentences within the allotted ten minutes. On average, those participants managed to complete five sentences using target vocabulary, and practice material A was completed by eight out of eleven participants.

3.2.2 Limitations of the Method

The major drawbacks of the method are the inherent weakness of a *between subjects design* and the threats against validity associated with classroom experiments of this kind. Firstly, a within subjects design such as the one used by Folse (2006) would have had all participants completing both activities, thus negating individual differences in proficiency level between the participants. However, in short, a within subjects design would have required a larger number of target vocabulary at the adequate difficulty and frequency level, and alternate versions of the practice materials and test, since half of the target vocabulary would only have appeared in one of the

activities without an alternate version. Thus, alternative versions would have required a larger number of participants than what was available for this study.

Furthermore, Lee (2012:28) highlights the validity issues associated with a classroom experiment such as the one used in this study. To begin with, using a *Randomised Control Group* (RCT) which would not receive any *treatment* (in this case practice material A or B) would have increased internal validity by supporting that any effect of either practice material on vocabulary retention was not a product of chance. However, Lee (2012:29) also acknowledges that RCT design is often impractical in a school setting and requires a large number of participants. Instead, small scale experiments using natural groups in a classroom setting can be a valid alternative in a "real life" setting, in order for a teacher (or future teacher) to find out what methods works the best.

3.2.3 Ethical Considerations

This study followed the ethical principles laid forth by the Swedish Research Council (SRC) (2002) concerning information, consent, confidentiality and usage. In addition, special considerations were taken in regards to information, consent and the benefit of the participants. To begin with, the principle of information states that the participants are to be informed about the aim, purpose and the possible benefits of the study. In addition, the researcher needs to make clear that participation is voluntary, and that a participant can withdraw their participation at any time (SRC 2002:7-8). Given that none of the participants had English or Swedish as their native language, with the majority being more proficient in Swedish than in

English, the information was given orally in both simplified Swedish and simplified English, before the written information (Appendix 2) was handed out. In addition, the participants had the opportunity to ask their questions about the study in both languages. Moreover, the teacher present assisted the students in understanding the written information and the consent form (Appendix 1). Furthermore, the participants were informed in detail about the proceedings of the experiment, including the presence of the posttest, knowing full well that the knowledge of an impending test would affect the result of the study (Schmitt 2008:339). In other words, participants who know beforehand that they are going to be tested tend to perform better on said test than participants who are tested without prior knowledge of the test. However, since any possible effect on the scores would be the same in both material groups, the choice was made to inform the participants of the posttest.

In addition, the ambition was also to provide a worthwhile and informative experience for the participants of the study. That is to say, the participants were adult EFL learners with a limited amount of lesson time compared to younger learners. In addition, adults finance their own studies, making their time valuable in several senses of the word. Thus, special consideration was taken to maximise the language learning benefit for the participants within the confines of the experiment, and to not occupy more of the participants' lesson time than necessary. Consequently, including the information about the study, the experiment took 45 minutes to conduct, leaving 45 minutes of the participants' regular 90 minute lesson time intact.

3.2.4 Analysis of Data

The tests were corrected by awarding one point for every answer deemed correct, with a maximum score of 12. Misspellings were allowed to the degree that it was still clear what word the participant intended. For instance, points were awarded to the participants who wrote *kedney and *kindey for *kidney*, but no points were awarded for the less intelligible *kendel and *kendy. Furthermore, misspellings that lead to the intended (read: correct) word turning into another word, such as *nets for *nest* and *warps for *wasp* were not regarded as correct. After correction, the tests were divided into two groups named material group A and material group B, depending on which practice material the participant had completed before taking the test. Next, an analysis of the posttest and practice materials was performed.

4 Results

The first priority of the analysis was to test the hypothesis in regards to the efficacy of the two different practice materials on vocabulary retention on the posttest. Thus, scores were analysed on a group level as well on an individual level. An additional point of interest was to see to what extent the level and frequency scan was successful. That is to say, if there were any words that the participants answered correctly to a higher degree than other words, thus indicating that they were previously known or somehow easier to "learn"

4.1 Posttest

This section compares how material group A and material group B performed on the posttest on a group level as well as individually, explains the calculations performed in the analysis, and details to what degree each individual item of target vocabulary was correctly answered on the posttest.

To begin with, in order to calculate the mean values (M) of number of correct answers on the posttest in the respective groups, the total number of correct answers in each material group was divided by the number of participants in that group. Next, to determine standard deviation of the mean (SD) for each group, the mean value is subtracted from each individual score, and that number is then squared, followed by calculating the mean of all the squared differences, and finally squaring that number. A low standard deviation value (≤ 2.0) indicates individual scores more closely grouped around the mean, while a higher standard deviation (>2.0) indicates a larger discrepancy. Figure 1 illustrates mean values score for material groups A and B.

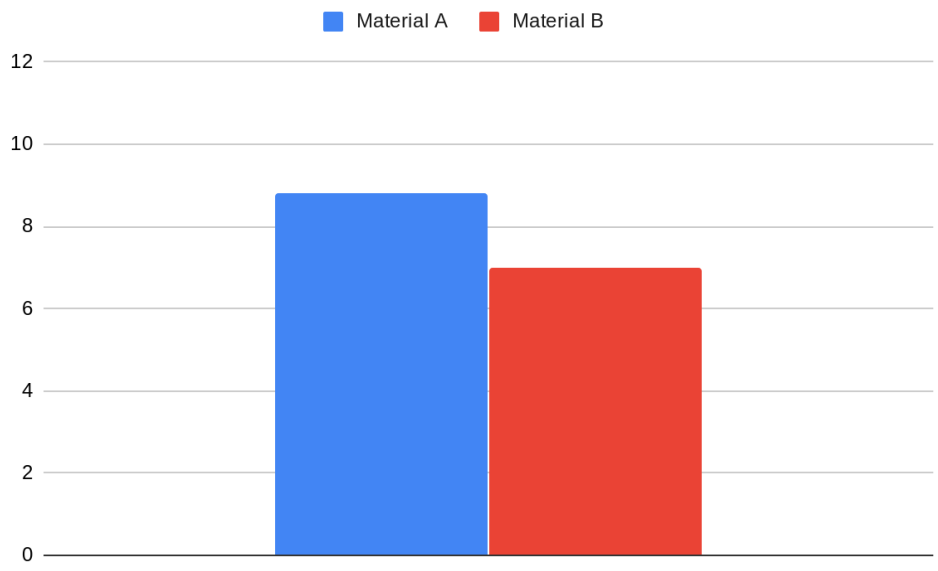

Figure 2. Mean Numbers of Correct Answers divided by Practice Material

The figure shows that the participants who practiced with material A, a fill-in-the-blank exercise, performed better on the posttest (M=8,8, SD=3,4) than the participants who practised using material B, an original sentence writing exercise (M=7 SD=2,7).

Next, to determine if there was a significant ($p \leq 0.05$) difference between the mean values of the groups, an independent samples t-test was conducted. A significant difference would have indicated that the result was not a product of chance. However, the difference was not significant, t(18)=1.30, p=0,209. Moreover, as shown by the high standard deviation values, and illustrated by figure 3, there was a large discrepancy between individual scores, in both groups.

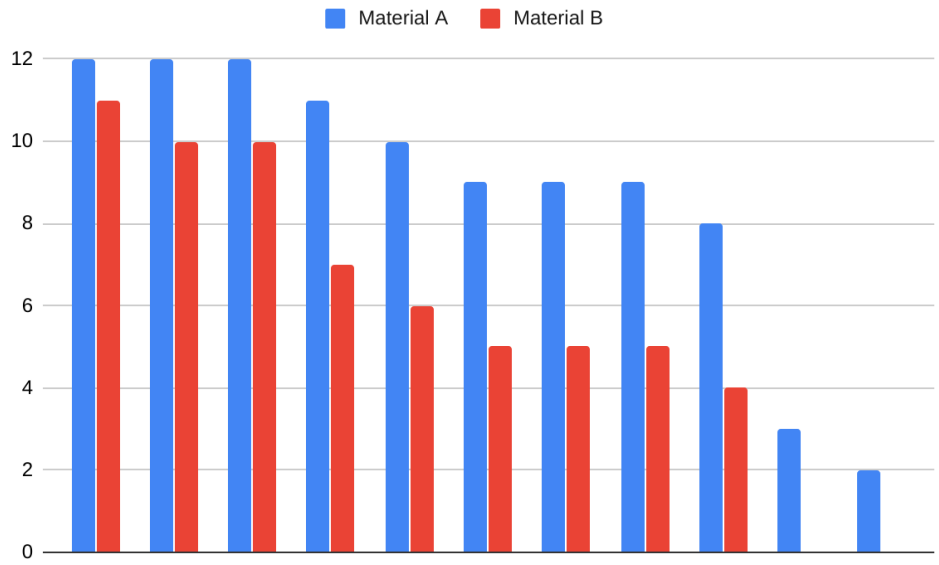

Figure 3. Individual Scores divided by Practice Material

In both groups overall, and particularly in material group A, performance on the posttest varied greatly. That is to say, in material group A, three participants recorded the maximum score of twelve, whereas one participant recorded the overall lowest score of two. In material group B, the highest score was eleven (one participant) and four was the lowest (one participant). Discrepancy was also found in the number of correct answers to each individual item of target vocabulary. As previously detailed (see Method), 12 target nouns from the film were selected using a CEFR level, COBUILD, and NGSL frequency scan. Figure 4 and the text below detail how the participants performed on each item.

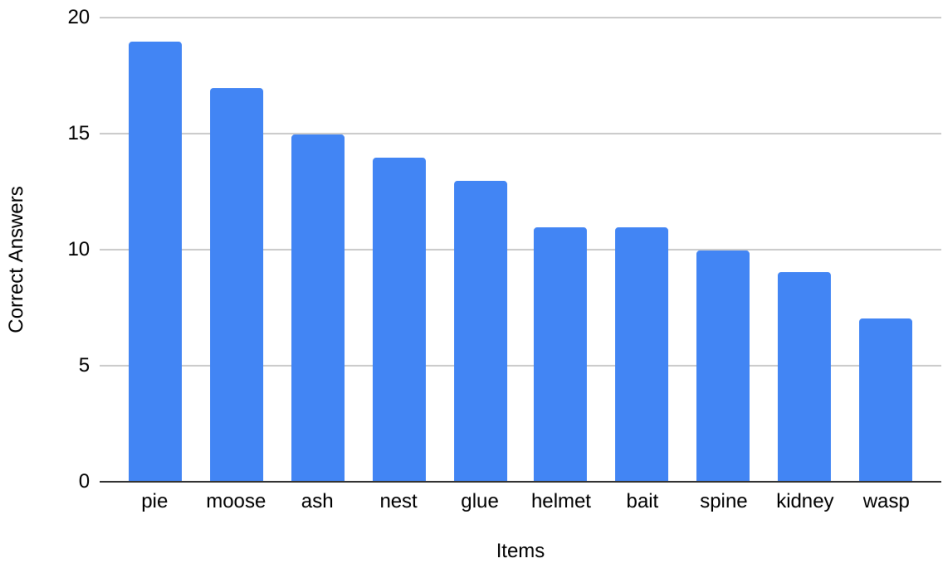

Figure 4. Number of correct answers to individual items of target vocabulary

Unfortunately, the uneven number of participants in each material group (eleven in material group A, and nine in material group B) does not lend itself to a comparison of correct answers divided by the number of participants in each material group. Nevertheless, the figure shows how the total 20 participants scored on each individual word on the posttest. The most commonly correct word on the posttest was *pie* (19/20 correct answers CEFR = B1, COBUILD = 2/5), followed by *moose* (17/20 correct, CEFR = N/A, COBUILD = 1/5), and *ash* (15/20 correct, CEFR = N/A, COBUILD = 2/5). Furthermore, the most commonly erroneous word on the posttest was *wasp* (7/20 correct, CEFR = B2, COBUILD = 1/5) before *kidney* (9/20 correct, CEFR = C2, COBUILD = 2/5) and, *spine* (10/20 correct, CEFR = C1, COBUILD = 2/5). Consequently, it appears that the CEFR-level scan (where available) served as an adequate predictor of which items would either be previously known/unknown or easier/harder to "learn". For instance, the most common correctly answered word *pie* had the lowest CEFR-rating of all items, at B1, whereas the two out of three of the most commonly incorrect words *spine* and *kidney* had a CEFR rating of C1 and C2 respectively.

In summary, although the participants who practiced with material A performed better on the posttest, the result did not reach the significance level. Furthermore, there was a large discrepancy between individual scores in both groups, particularly in material group A where three participants recorded the maximum score and one participant scored the lowest overall score of two. Finally, some items of target vocabulary (*pie* and *moose*) appear to have been either previously known by the participants in general, or easier to "learn" than the others.

5 Discussion

The results of the analysis showed that the participants who used practice material A, a fill-in-the-blank exercise performed better on an immediate posttest than the participants who used practice material B, an original sentence writing exercise. The difference, however, was not significant ($p=0.209$). The result showed a tendency towards coinciding with previous research by Folse (2006) and Hashemzadeh (2012) who both found fill-in-the-blank exercises to lead to better vocabulary retention than a variety of other exercises. Moreover, this study attempted to duplicate the findings of aforementioned international research on homogenous groups in a heterogeneous group of EFL learners, while adding the dimension of a limited time-on-task. Consequently while the hypothesis formulated on the basis of previous results could not be proven, a tendency towards better retention after a fill-in-the-blank exercise was seen. The main possible reasons for this tendency could be the discrepancy in the difficulty level of the exercises, the insufficient time on task, number of word retrievals per exercise, the different focus of the respective exercises and, the previous vocabulary knowledge and other individual differences between the participants. To begin with, the majority of the participants (8/11) who did practice material A were able to complete it within the allotted time, whereas

no participant who practiced using material B was able to write twelve original sentences. This was an unexpectedly low completion grade. That is to say, even though the risk of some participants not completing their material was noted, it was not expected that material group B, on average, would only manage to construct five sentences using target vocabulary. As a result, the number of word retrievals is likely to have been much higher in material group A than in material group B, and previous research by Folse (2006) has demonstrated the importance of a high number of word retrievals for vocabulary retention. Furthermore, an original sentence writing exercise involves more aspects of word knowledge in terms of grammatical functions and collocations (see table 2) than a fill-in-the-blank exercise in which that context is already given. Thus, learners who have just been presented with new vocabulary have merely established the initial form-meaning link and lack the other aspects of word knowledge needed to complete an original sentence writing exercise within a reasonable time. Consequently, the pedagogical implication of the result appears to be that a fill-in-the-blank exercise is better suited directly after the initial presentation of a new word, when the form-meaning link needs to be consolidated, whereas a production exercise is better suited at subsequent stages of learning, when more aspects of word knowledge have been acquired, and significantly more lesson time is available.

Naturally, the tendency found could also have been influenced by the participants' previous knowledge of target vocabulary and other individual differences related to language learning. Firstly, despite the level and frequency scan of the words, there is simply no way of ensuring that all the words were unfamiliar to every participant at the time of the experiment. On the contrary, the high frequency of correct answers on certain items overall and the perfect score on the posttest by three participants in material group A might indicate the opposite. However, it is also possible that those high

scoring participants were simply motivated, high aptitude learners who knew that they were going to be tested on their knowledge, and thus performed at the height of their abilities and, the directed focus on target vocabulary provided by practice material A, combined with a higher number of word retrievals may have helped facilitate their performance on the posttest.

In retrospect, this study was faced with two, for different reasons, difficult choices: the selection of target vocabulary and the decision to limit practice time. To begin with, besides testing the hypothesis, an ethical aim of this study was to maximise the language learning benefit of the participants. Thus, the film *Dumb Ways to Die* was selected first, as it was considered close to optimal in creating an engaging and entertaining presentation. Consequently, the choice of target vocabulary became limited to concrete nouns illustrated in the film or included in the lyrics of the song. As a result, it was difficult to extract an adequate number of low frequency, high CEFR level words thought to be unknown to participants at the current level. This resulted in the inclusion of the item *pie* which appears to have been either previously known by participants or for some reason easier to "learn". In the case of *pie*, it was the only item of target vocabulary ranked as low as CEFR level B1 and should not have been included as a thought to be unknown word to participants at the level of the ones in this study. In hindsight, a selection of target vocabulary listed as CEFR level B2 or higher should have preceded the choice of presentation material, as an engaging presentation, for instance, in the form of a language learning game could have easily been designed based on that selection. Furthermore, the benefit of the participants was also one of the underlying reasons for limiting practice time. That is to say, even though a participant might be offended by not being allowed to finish an assignment, the offence of wasting other participants' valuable lesson time by having to wait for every last person to

finish was deemed to be a larger offence. Granted, early finishers could have been given another assignment from the researcher or the teacher present that they could work on in their own time while other participants finished the practice materials. However, this would have added a complex variable to the experiment, in terms of bringing in other items of vocabulary not part of the experiment. The other reason for limiting practice time was to compare the efficacy of the practice materials in terms of number of words acquired within a given time. Nevertheless, even though the choice was justified, it was a difficult one to make. In hindsight, however, ten minutes of practice time was simply too short. Given the low completion grade of material B, ten minutes could have been increased to 15 without causing too much discomfort to the early finishers while also increasing the validity of the result.

Consequently, in conclusion, a tendency was found, pointing towards a fill-in-the-blank exercise as a better alternative in terms of vocabulary retention, at the early stages of target vocabulary acquisition and learning, in particular when a limited amount of practice time is available. Needless to say, the small sample size of this study nullifies the ability to draw any wider conclusions from this tendency. Future research could duplicate this study using a larger population sample, a more adequate time on task, and preferably using a within subjects design.

5.1 Conclusion

The purpose of this study was to examine the efficacy of two practice exercises for vocabulary retention not to argue one's superiority over the other. Recognition exercises such as fill-in-the-blank and production exercises such as original sentence writing both have merit as valuable tools in a teacher’s arsenal. The true difference in efficacy however, may lie in the context in which such exercises are used. Thus, it is of key importance that an EFL teacher is aware of the pros and cons of different practice materials in terms of the challenge level presented, time consumption and stage of learning, as to best scaffold the learners’ vocabulary acquisition and retention.

6 References

Allagui, I., & Breslow, H. (2016). Social media for public relations: Lessons from four effective cases. *Public Relations Review*, *42*(1), 20-30.

Arju, S. (2011). A study on ESL vocabulary acquisition needs and classroom practice: a Bangladeshi context. *Stamford Journal of English*, *6*, 54-71.

Brown, T. S., & Perry Jr, F. L. (1991). A comparison of three learning strategies for ESL vocabulary acquisition. *Tesol Quarterly*, *25*(4), 655-670.

Browne, C. (2013). The new general service list: Celebrating 60 years of vocabulary learning. *The Language Teacher*, *37*(4), 13-16.

Cobuild, C. (1995). English language dictionary, 1987. *London: Collins*.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of verbal learning and verbal behavior*, *11*(6), 671-684.

Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of experimental Psychology: general*, *104*(3), 268.

Dictionary.cambridge.org. (2019). *WORD | meaning in the Cambridge English Dictionary*. [online] Available at: https://dictionary.cambridge.org/dictionary/english/word [Accessed 18 Dec. 2019].

DumbWays2Die. (2012). *Dumb Ways to Die*. [Online]. Available at: https://www.youtube.com/watch?v=IJNR2EpS0jw [Accessed 15 Dec. 2019].

Dörnyei, Z. (2014). The psychology of the language learner: Individual differences in second language acquisition. Routledge.

Englishprofile.org. (2019). *English Profile - English Vocabulary Profile*. [online] Available at: https://www.englishprofile.org/wordlists [Accessed 18 Dec. 2019].

Folse, K. S. (2004). Myths about teaching and learning second language vocabulary: What recent research says. *TESL reporter*, *37*(2), 1-13.

Folse, K. S. (2006). The effect of type of written exercise on L2 vocabulary retention. *TESOL quarterly*, *40*(2), 273-293.

Hashemzadeh, M. (2012). The Effect of Exercise Types on EFL Learners' Vocabulary Retention. *Theory & Practice in Language Studies*, *2*(8).

Harmer, J. (2007). *The practice of English language teaching*. Harlow: Pearson Longman.

Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In *Vocabulary and applied linguistics* (pp. 113-125). Palgrave Macmillan, London.

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied linguistics*, *22*(1), 1-26.

Lee, J. H. (2012). Experimental methodology in English teaching and learning: Method features, validity issues, and embedded experimental design. *English Teaching: Practice and Critique*, *11*(2), 25-43.

Nation, P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. *Vocabulary: Description, acquisition and pedagogy*, *14*, 6-19.

Nation, I. S. (2001). *Learning vocabulary in another language*. Ernst Klett Sprachen.

National Agency for Education (2018). Basic Adult Education continues to Grow. [online]. Available at: https://www.skolverket.se/skolutveckling/statistik/arkiverade-statistiknyheter/statistik/2018-06-19-komvux-pa-grundlaggande-och-gymnasial-niva-fortsatter-att-vaxa [Accessed 15 Dec. 2019].

National Agency for Education (2019). Basic Adult Education - English . [online]. Available at:https://www.skolverket.se/undervisning/vuxenutbildningen/komvux-grundlaggande/laroplan-for-vux-och-kursplaner-for-komvux-grundlaggande/kursplaner-for-komvux-pa-grundlaggande-niva?url=1530314731%2Fsyllabuscw%2Fjsp%2Fsubject.htm%3FsubjectCode%3DGRNENG2%26tos%3Dvuxgr&sv.url=12.4fc05a3f164131a741820f6 [Accessed 15 Dec. 2019].

National Agency for Education (2019). English. [online]. Available at: https://www.skolverket.se/download/18.4fc05a3f164131a74181056/1535372297288/English-swedish-school.pdf [Accessed 30 Oct. 2019].

Olsson, E. (2015). Progress in English academic vocabulary use in writing among CLIL and non-CLIL students in Sweden. *Moderna språk*, *109*(2), 51-74.

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language teaching research*, *12*(3), 329-363.

Schmitt, N., & McCarthy, M. (1997). *Vocabulary: Description, acquisition and pedagogy*. Cambridge university press.

Seliger, H. W., Seliger, H., Shohamy, E. G., & Shohamy, E. (1989). *Second language research methods*. Oxford University Press.

Skehan, P. (1991). Individual differences in second language learning. *Studies in second language acquisition*, *13*(2), 275-298.

Swedish Research Council, V. (2002). Ethical Principles in Social-Science Research

Sylvén, L. K. (2004). *Teaching in English or English teaching? On the effects of content and language integrated learning on Swedish learners' incidental vocabulary acquisition* (Doctoral dissertation, Göteborg University).

West, M. P. (Ed.). (1953). *A general service list of English words: with semantic frequencies and a supplementary word-list for the writing of popular science and technology*. Longmans, Green.

Wilkins, D. A. (1972). *Linguistics in language teaching*. E. Arnold, 1973.

Appendix 1 Practice Material A/B and Test

Practice Material A - Words to Learn

	a nest		a dryer
	a spine		glue
	a moose		a helmet
	bait		a pie
	ash		a skeleton
	a kidney		a wasp

Practice Material A / Fill-in-the-blank

Complete the sentences by using one (1) word from the box below. You may use a dictionary. The maximum time for this exercise is ten (10) minutes.

nest	ash	skeleton	spine	wasp	moose	
bait	pie	helmet	dryer	kidney	glue	~~track~~

Example: Do not walk on the railroad tracks! It is extremely dangerous.

Yesterday, when I was walking in the forest, I saw _____

I need to buy some _____before we go fishing

Would you like to try the _____I just made?

If you're going biking, don't forget your_____.

Your sweater is in the _____It will be ready soon.

Too much salt can be bad for your_____.

The chair is broken. I need some _____to hold it together.

Too much sitting down is bad for your _____. You need to move around more.

On Halloween, kids sometimes dress up as_____, and go asking for candy.

If a _____stings you, it will hurt.

Look!, there's a bird's-_____up in that tree.

Practice Material B - Words to Learn

	a nest		a dryer
	a spine		glue
	a moose		a helmet
	bait		a pie
	ash		a skeleton
	a kidney		a wasp

Practice Material B / Original Sentence Writing

Write one (1) original sentence to every word in the box below. You may use a dictionary. The maximum time for this exercise is ten (10) minutes.

nest	ash	skeleton	spine	wasp	moose	
bait	pie	helmet	dryer	kidney	glue	~~track~~

Example: Do not walk on the railroad <u>tracks</u>! It is extremely dangerous.

Linnæus University
Sweden

Test

Write the correct word to the right of the picture. Please write only one (1) word to every picture.

 _____  _____

 _____  _____

 _____  _____

 _____  _____

 _____  _____

 _____  _____

Thank you so much for your participation in this study!

General Information

Age _____

Gender: Male ☐ Female ☐

Mother tongue (first language) _____

Years studying English _____

I did practice material: A ☐ B ☐

Appendix 2 - Information and Consent Form

Information about the Study

Hello

My name is John Lindström, and I am a student at the upper secondary teacher training program at Linnaeus University. Right now, I am writing my Bachelor's thesis on vocabulary.

The study compares two different ways of practicing vocabulary, a fill-in-the blank activity and an original sentence writing activity. The aim of the study is to find out which of these two activities is the most effective one. Hopefully, the result of this study and others like it helps teachers in selecting the best exercises for you, the student. In this study, words will be presented, practised with the two different activities and finally tested.

If you choose to participate in this study, you will be completely anonymous. Your name or person will not appear in any way in the material or the finished paper.

You are free to end your participation at any point during the project, and if you have any questions, you may ask them now. If you have any further questions after today you can contact the researcher at:

john-rune.lindstrom@ksgyf.se

Consent form

I hereby agree to participate in this study on vocabulary

- I understand the information about the purpose and aim of this study
- I have been given the opportunity to ask questions about the study, and I know who to contact if I have any further questions
- I participate willingly in this study, and I am aware of that I can end my participation at any point
- I allow the researcher and the Linnaeus University to store and process the material gathered in this study

Kalmar, _____/_____ 2019