

Age/Order of Acquisition effects and massed versus distributed learning



Bachelor Thesis BA Linguistics

Utrecht University

Iris van der Wulp

Student number: 4295609

E-mail: i.m.vanderwulp@students.uu.nl

Supervisor: Shalom Zuckerman

Second reader: Sergey Avrutin

Table of contents

Abstract	2
Introduction	3
Theoretical Background	4
The Age of Acquisition Effect	4
Age of Acquisition in L2	5
Massed versus Distributed Practice	6
OoA and Nonwords	7
Research questions and hypotheses	8
Method	9
Participants	9
Materials	10
Procedure	10
Picture Naming Task	11
Word Search Puzzle	11
Results	12
Participants and means	12
Scoring	12
Is there an effect of version?	12
Is there an effect of method?	13
Is there an OoA effect visible in this study?	13
Is there an effect of (massed and distributed) practice?	15
Do some words have higher scores than others?	16
Can the AoA effect be (partially) explained by an effect of distributed practice?	16
Discussion	18
Conclusion	20
References	21
Appendix	23
Appendix 1. Materials	23
Appendix 2. PowerPoint Presentations	24
Appendix 2.1. <i>Set A</i>	24
Appendix 2.2. <i>Set B</i>	28
Appendix 2.3. <i>Set A & B</i>	32
Appendix 3. Information letters for the participants	38
Appendix 3.1. <i>Letters for group 1</i>	38
Appendix 3.2. <i>Letters for group 2</i>	40
Appendix 4. Test sessions	42
Appendix 4.1. <i>Word Search Puzzles</i>	42
Appendix 4.2. <i>Picture Naming Tasks</i>	46

Abstract

The Age of Acquisition effect (AoA) shows that words learned earlier in life are remembered better than words acquired later. The goal of this experiment was to teach participants non-words in a specific order, to manipulate their AoA. When AoA is manipulated in an experimental setting, it is called Order of Acquisition (OoA). This method of using a word-training study is based on the work of Izura et al. (2011). Also, the effect of massed versus distributed practice was tested, by assigning participants to a massed or distributed group, who learned the words in one day (massed) or spread over three weeks (distributed). The main question of this thesis was if the effect of practice and OoA would show interaction, in a way that distributed practice might explain the OoA effect. Two different test methods were used: a picture naming task and a word search puzzle. A total of 38 participants took part in this study. The participants who made the WSP had significantly higher scores than participants who made the PNT. Unfortunately, because of this effect of method the OoA could not be statistically tested, as the sets were not counterbalanced within each method. There was no effect of practice. The current results seem to show no interaction between practice and OoA, but since both effects were not significant or not testable, this remains a question for further research.

Introduction

This Bachelor Thesis is based on a research proposal, which was written by Bart Verlangen and me for the course Experimental Psycholinguistics at Utrecht University (van der Wulp & Verlangen, 2017). Bart gave me his permission to continue with the subject for this BA thesis. This thesis was written during block 3 and 4 of the course year 2017-2018, because this thesis consists of experimental research. Block 3 was used for the preparation of the materials, studying the literature and finding participants. Block 4 was used for the execution of the experiment, the processing of the results, and writing this thesis.

An important notion for this thesis is the Age of Acquisition (AoA) effect, which shows that earlier acquired words are remembered better than words that are acquired later in life. This effect is mainly shown in word processing tasks, where earlier acquired words are processed faster and remembered better than later acquired words (Juhasz, 2005).

As an attempt to isolate this effect and control for other factors it correlates with, word training experiments can be used. In these experiments, participants are taught new words and the order in which these words are learned is controlled. This could then generate Order of Acquisition effects (OoA). In this case it is called the Order of Acquisition effect, because AoA is a natural effect of language acquisition, that is researched by subjective questionnaires to determine when words were acquired (as in Hirsh, Morrison, Gaset, & Carnicer (2003); where participants had to rate their AoA of certain words of their mother language on a 7-point scale).

It is not clear if AoA and OoA generate the exact same effects. However, research on OoA effects is very important for the field of AoA research, because it allows researchers to control not only the order in which words are learned, but also other factors that correlate with AoA like word frequency, length, and cumulative frequency. These notions will be explained later in this thesis.

Many other explanations for the AoA effect have been proposed. Izura et al. (2011) also mention in their discussion that AoA effects in word learning could possibly be explained by distributed learning effects: words that are acquired earlier in life are by default acquired more distributed over time than later acquired words. The aim of this study is to research if distributed practice effects interact with OoA/AoA effects.

The structure of this Bachelor thesis is as follows:

First, the theoretical background of the subject will be described. Then, the research questions and hypotheses are presented. After that, the method used for this experiment is explained. Then, the results of the experiment are presented, followed by the discussion and conclusion.

Theoretical Background

The Age of Acquisition Effect

Juhasz (2005) provides an extensive review of research that focusses on the Age of Acquisition (AoA) effect in word learning and several theories that follow from this research. She mentions that AoA has a high correlation with other factors – such as word frequency (how frequently a word is used in a certain language), word length, concreteness of the word, word familiarity for the subject, and imageability – and provides an overview of methods and statistical analyses that try to isolate the AoA effect from these factors. Then, Juhasz (2005) explains several theories that aim to account for the AoA effects. The four most relevant hypotheses of AoA for this thesis will now be summarized.

The Phonological Completeness Hypothesis states that children acquire words as a whole phonological form, whereas later acquired words would be memorized as phonetic segments that have to generate the phonetic form each time the word is spoken. However, there are also AoA effects found in adults who acquire a second language (L2), as we will see later on in this thesis as well. Adults do not acquire words as whole phonetic forms, but segmented in phonetic parts. Therefore, this theory is not favoured over other hypotheses of AoA.

The Cumulative Frequency Hypothesis states that early acquired words have more occurrences in someone's lifetime than later acquired words. Therefore, they are memorized better than later acquired words. So instead of an AoA effect, the occurring effect is actually an effect of cumulative frequency. It will be shown later in this thesis that cumulative frequency is controllable and will therefore probably not be the cause of AoA effects.

The Semantic Locus Hypothesis states that new words are connected to the concept they represent. These concepts form a network in the brain. New concepts build upon the earlier acquired concepts. Therefore, early acquired concepts are integrated more deeply into the mental lexicon. However, this hypothesis predicts no AoA effects for the L2 (Izura et al., 2011). The idea that the AoA effects arise from semantic concepts building on each other predicts that new words that are learned for existing concepts would generate AoA effects only when they correspond to the same concepts that generate AoA effects in the L1. However, AoA effects were observed in the L2 (Izura et al., 2011; Juhasz, 2005).

The Network Plasticity Hypothesis states that a network becomes less plastic/changeable when it grows (Juhasz, 2005). This hypothesis is based on a connectionist model built by Ellis and Lambon Ralph (2000), which was trained to learn words. Ellis and Lambon Ralph (2000) trained the model on an 'early' set and a 'late' set of words. When the early set was not repeated after training on the late set began, the early set was completely forgotten. This is called "catastrophic interference." However, when the early set was still trained (although maybe less often than the late set), it was always remembered better by the model than the late set. This showed AoA effects (Ellis & Lambon Ralph, 2000; Izura et al., 2011; Juhasz, 2005). Therefore, Ellis and Lambon Ralph (2000) hypothesized that a network loses plasticity as it grows, as an explanation for AoA effects.

The Mapping Hypothesis is a subpart of the Network Plasticity Hypothesis. It states that AoA effects only occur when mappings from input to output (for example orthography to semantics)

are arbitrary. These effects have been observed (Izura et al., 2011; Lambon Ralph & Ehsan, 2006; Mermillod, Bonin, Méot, Ferrand, & Paindavoine, 2012).

Juhász (2005) concludes that the AoA effect is very robust and real, whereas some researchers explain it as frequency of familiarity effects. The most interesting hypotheses for this study are the Semantic Locus Hypothesis and the Network Plasticity Hypothesis (containing the Mapping Hypothesis), because they concern AoA in the L2 as well. The next session will continue about this subject.

Age of Acquisition in L2

Izura et al. (2011) used the second language (L2) as a method to control the AoA of words learned by their participants. They controlled for other factors that correlate with AoA like frequency, length, and cumulative frequency. Izura et al. (2011) explain that the Semantic Locus Hypothesis predicts no AoA effects in the L2, whereas the Mapping Hypothesis does. They place these hypotheses as opposed to each other.

In three different experiments, Izura et al. (2011) trained participants to learn words in a language they never learned before. They controlled the Order of Acquisition (OoA) by splitting the materials into a set A and set B and the participants into two groups. One of the groups would learn set A as the 'early' set and set B as the 'late' set. For the other group it was the other way around. The 'early' set is trained first. Then the 'late' set is taught either in combination with the early set (experiment 1) or first by itself and then together with the 'early' set (experiments 2 & 3), controlling for cumulative frequency, word frequency, length and cognates.

In the first two experiments, participants were monolingual English native speakers with no experience with the Spanish language. Spread over several weeks, they learned Spanish words corresponding to pictures in the following way: first, a fixation asterisk appeared on a computer screen for 1000 ms. After this, one picture of an object appeared with the Spanish name written underneath it for 2000 ms. During these two seconds, the Spanish name was also produced as a spoken word by a voice through speakers. After this, the screen went blank for 500 ms and then the asterisk appeared again for 1000 ms before the next stimulus was presented. The difference between experiment 1 and 2 is that in experiment 1, the 'early' set was presented on its own during the first two learning sessions and the 'late' set only occurred in combination with the 'early' set. In the second experiment, the 'late' set was also presented alone during two later training sessions. Both experiments controlled for cumulative frequency by presenting both sets an equal amount of times. Experiment 2 also controlled for frequency trajectories by training the 'late' set on its own for an equal amount of times as the 'early' set. At the end of each training session, a picture naming task with feedback was carried out.

Several test sessions were held during the experiments. These test sessions were picture naming tasks in the first two experiments, to measure OoA effects. There was a significant effect of OoA in experiments 1 and 2, providing evidence for the AoA effect to be independent of cumulative frequency and frequency trajectories. It also provides evidence of AoA effects in the L2 (Izura et al., 2011).

To further test the influence of OoA effects in word recognition, a third experiment was conducted. In this experiment, native speakers of Spanish who spoke English as a second language were taught words in Welsh. The learning sessions were similar to those in experiments 1 and 2,

although different words and pictures were used due to the different training language. Also, during some of the learning sessions and in addition to the learning method that was used in experiment 1 and 2, participants were asked to repeat the words out loud, write them down or produce the word when only the first half of it was shown. Instead of a picture naming task with feedback after each training session, two acquisition tests were carried out. These tests were picture naming and writing with feedback (Izura et al., 2011).

The test sessions of experiment 3 consisted of a picture naming task, a lexical decision task and a semantic categorization task. These were always carried out in this order. Also, a delayed picture naming task was done in the first test session. For the picture naming tasks, there was no significant effect of OoA. For the statistically corrected scores of the picture naming tasks, there was a significant effect of OoA, but only for one order of presentation of the sets. The lexical decision task and the semantic categorization task did show significant effects of OoA. Because the tests were always presented in the same order, a priming effect might have occurred.

Izura et al. (2011) conclude that there is an effect of OoA/AoA in the L2 and that cumulative frequency and frequency trajectories do not explain these effects. They also favour the Mapping Hypothesis over the Semantic Locus Hypothesis, because the Semantic Locus Hypothesis does not account for AoA effects in the L2.

Massed versus Distributed Practice

In their general discussion, Izura et al. (2011) mention a possible alternative explanation for AoA effects. Words that are acquired earlier in life are by default acquired more distributed over time. Therefore, this *distributed practice* might be the cause of the AoA effects. Izura et al. (2011) refer to Cepeda et al. (2009), who conducted two experiments to test the effects of massed and distributed practice on memory tasks. However, these experiments are focussed on finding the *optimal gap* between learning sessions with the goal of gaining more knowledge about the learning process and not about AoA.

In the first experiment done by Cepeda et al. (2009), participants learned words from an unfamiliar L2. The participants were native speakers of English, who learned words in Swahili. There were two training sessions and one test session. During the training sessions, the Swahili words were shown one by one, with their English translations. After that, a test-with-feedback trial started. Participants saw the Swahili words and had to type the English equivalents to those words. If a response was made, the Swahili word with the correct English translation was presented on the screen. If two correct responses were made for a word, it did not appear again. This went on until two consecutive correct responses were given for each word.

There were two variables: gap and test delay. The gap is the time between the two learning sessions. This varied between 5 minutes and 14 days. Participants were randomly divided over the gap conditions. The test delay was held constant at 10 days. Cepeda et al. (2009) found that a gap of one day seemed optimal for a test delay of 10 days.

In the second experiment the test delay was 6 months. Participants also had two learning sessions. This time, participants learned names of unfamiliar visually presented objects and unknown facts, because pilot data suggested that word pairs would produce floor effects with a test delay as long as 6 months. The two learning sessions were separated by a gap that varied

from 20 minutes to 6 months. Participants were again randomly divided over the conditions. In this case, the optimal gap turned out to be 28 days.

Cepeda et al. (2009) concluded that the optimal gap increases as the test delay increases. However, the ratio of the optimal gap also decreases as a function of test delay. For example: in experiment 2, participants who trained with a gap of 6 months performed less accurately on the test than participants with a gap condition of 28 days. In other words: the optimal gap is not too short and not too long, with respect to the test delay.

Sobel, Cepeda, & Kapler (2011) also did an experiment to test the effects of massed versus distributed practice in vocabulary learning. They did a classroom experiment where children who had English as their native language had to learn unfamiliar English vocabulary.

There were two conditions. The first condition – the ‘massed’ condition – consisted of two learning sessions separated by less than one minute. The ‘distributed’ condition consisted of the same two learning sessions separated by one week. In the training sessions the novel words and their definitions were shown to the children. Then, the children had to fill in the words and their corresponding definitions in a booklet. The children also had to create a novel sentence for each word in which the word was used in a correct manner.

Five weeks after training, a vocabulary test was done. The children had to write down the definitions for the words they learned. Sobel et al. (2011) concluded that children who had learned the vocabulary in one day (the ‘massed’ condition) scored significantly lower in percentage correct than children who had two training sessions spaced over one week (the ‘distributed’ condition).

The review about spacing effects by Carpenter, Cepeda, Rohrer, Kang, & Pashler (2012) provides a lot of evidence as well that distributed learning enhances memory. However, most studies concerning spaced learning are carried out using only two training sessions and one test session. Carpenter et al. (2012) mention that therefore, the optimal gap for more than two training sessions is not determined yet. Moreover, as Izura et al. (2011) mention in their discussion, there is a need for OoA and distributed practice research with more than two learning sessions.

OoA and Nonwords

Catling, Dent, Preece, & Johnston (2013) carried out a laboratory experiment with nonwords assigned to pictures, to control for any possibility that a participant already knows some of the words or pictures. Also, cumulative frequency was controlled.

Their participants were undergraduate students. The materials were 3D Greeble pictures of novel shapes and 3-letter nonwords of a consonant-vowel-consonant structure. They used five training sessions, showing the Greeble and the nonword on a screen, with a fixation asterisk in between. The test session consisted of a picture naming task. They found a significant difference in the reaction times of correct naming, where participants responded faster to early learned pictures than later learned pictures. Because it is a laboratory study, where participants were trained instead of naturally acquiring the words, Catling et al. (2013) call the effect they found an OoA effect, like just Izura et al. (2011) did for the L2. Still, they argue that this is the only way to show that frequency effects are independent of AoA effects, as frequency was controlled in the experimental design.

Research questions and hypotheses

Izura et al. (2011) propose a research similar to theirs, with more than two learning sessions but with one condition in which all learning sessions are massed together and one in which they are distributed over time. That is what this thesis tries to accomplish. For this research, multiple research questions are important. The main research question of this thesis is:

1. Can the AoA effect be (partially) explained by an effect of distributed practice? In other words: is there an interaction between a practice effect and an OoA effect and which one is stronger?

The two sub-questions which are needed to answer the question above are:

2. Is there an OoA effect visible in this study?
3. Is there an effect of (massed or distributed) practice?

Other relevant questions are:

4. Is there an effect of method?
5. Is there an effect of version?
6. Do some words have higher scores than others? Could this be explained by the complicatedness/length/imageability etc. of the original word?

The hypothesis for questions 2 and 3 is yes, hopefully there is an effect of OoA and an effect of practice. The hypothesis for question 4 is that the group who did the word search puzzle shows a better performance than the group who did the picture naming task, because word recognition would be easier than word production. The hypothesis for questions 5 and 6 is that there is no effect of version and that all words are equally difficult.

There is no hypothesis for question 1. As this has not been researched yet, both options are possible: there could be an interaction or not. Moreover, when no OoA and no practice effects can be found, question 1 can possibly not be answered in this thesis, but perhaps in further research.

Method

This part of the thesis will explain the method that was used in this experiment. The participants learned words in a non-existent language. Therefore the OoA, cumulative frequency, frequency trajectories, length, word frequency, massed or distributed learning cognates, and AoA of the original concepts were controlled. Moreover, two different test methods were used and compared.

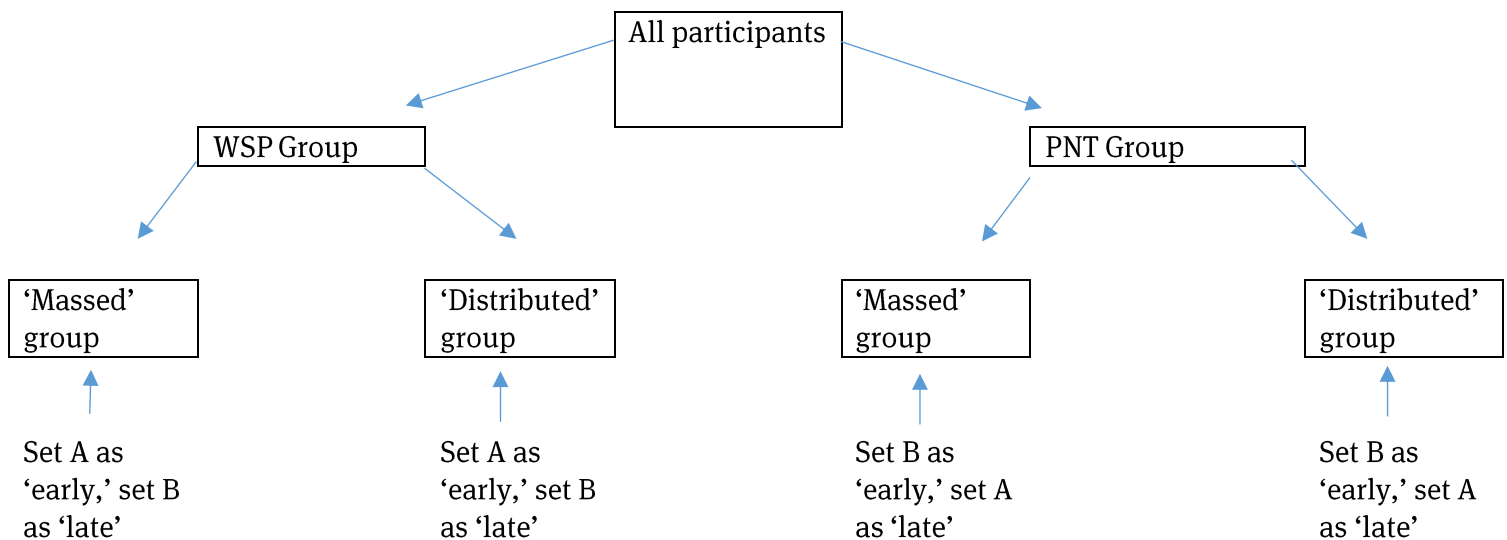
The experiment consisted of two conditions: the ‘massed’ and the ‘distributed’ condition. ‘Massed’ stands for ‘massed practice’ and ‘distributed’ stands for ‘distributed practice.’ The goal of the experiment was to study the effect of massed versus distributed practice on the performance of participants in retrieving nonwords they learned as a part of the experiment, and to research if this interacts with a possible Order of Acquisition effect (OoA). The participants learned the words during learning sessions. Because of the learning sessions, the timing of learning (‘massed’ or ‘distributed’) and the Order of Acquisition were controlled. After the training sessions, a test session was held. Two different test methods were used: a Picture Naming Task (PNT) and a Word Search Puzzle (WSP).

Participants

The participants were undergraduate students of the courses “Variation and Universality in Natural Language” and “Taalverwerving en Taalstructuur” at Utrecht University. The participants were divided into two groups: the PNT group and the WSP group, based on the test method that was used for that group. Each participant either completed the PNT or the WSP during the test session. They did not do both because of possible transfer effects.

Those groups were also further divided into two groups; for each method there was a ‘massed’ group and a ‘distributed’ group. Moreover, one set of words was always taught earlier than the other. For the WSP group, set A was the ‘early’ set and set B was the ‘late’ set (see Appendix 1 for the words in each set, see also “materials” below). For the PNT group, set B was the ‘early’ set and set A was the ‘late’ set. Figure 1 visualizes this set-up.

Figure 1: *Set-up of the experiment.*



Materials

The stimuli were 20 coloured pictures of objects along with 20 nonwords. The pictures were taught as the *meaning* of the nonwords. Izura et al. (2011) argue that teaching words and assigning a meaning to them gives better results than when the words are taught without an assigned meaning.

The pictures were chosen from the MultiPic database (Duñabeitia, J.A., Crepaldi, D., Meyer, A.S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M., in press) (See Appendix 2). The Dutch words corresponding to the objects in the pictures all have an AoA of 5 years old (see Appendix 1) (Brysbaert, Stevens, De Deyne, Voorspoels, & Storms, 2014)¹. Therefore, there will not be a natural AoA effect, which could possibly interfere with the OoA effect that was created in this study.

The nonwords were chosen from Keuleers, Diependaele, & Brysbaert (2010). All chosen nonwords had an accuracy of 1, which means that all participants of Keuleers et al. (2010) classified the word correctly as a nonword. Also, the chosen words were among the words with the fastest reaction times, which means that the participants accurately classified these words as nonwords the fastest compared to other nonwords (see Appendix 1). The chosen words also follow the average Dutch phonology and typology rules. For example, none of the chosen nonwords end with a “z.”

In assigning the nonwords to the pictures, phonological similarity between the nonwords and the Dutch words corresponding to the objects in the pictures was held into account to prevent any cognate-effects. The 20 stimuli were divided into two sets of each 10 stimuli (set A and set B; see Appendix 1).

Procedure

The experiment consisted of three learning sessions and one test session. During the first learning session only the ‘early’ set was shown, the second learning session showed the ‘late’ set only – controlling for frequency trajectories (Izura et al., 2011) – and the third session showed both the early and the late set (see figure 2). Cumulative frequency was controlled by presenting all items an equal number of times. In each session, each item was presented three times in random orders. Each item was therefore presented a total of nine times. The test session was held one week after the last learning session. For the ‘massed’ condition, all learning sessions took place on the same day. For the ‘distributed’ condition, the sessions were each one week apart.

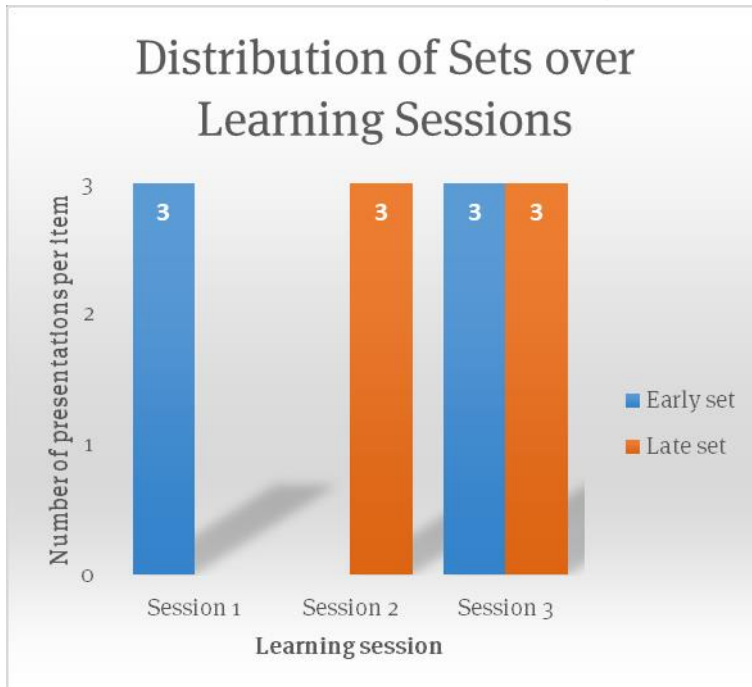
During the learning sessions, the participants were required to learn the 20 nonwords corresponding to the objects in the pictures. The PNT group had set B as the ‘early’ set. Therefore, these groups both learned the words from set B in the first session, without the presence of words from set A during this training session. Only words from set A were presented during the second training session. In the remaining training session, both set A and set B were presented. The WSP group had set A as their ‘early’ set, and therefore learned the words of set A in the first session without any words of set B. They learned set B in the second learning session and both in the remaining session. Handouts of the PowerPoint presentations used during the learning

¹ The website <http://crr.ugent.be/archives/1602> offers a download link for the Excel file with all AoA data collected in Brysbaert et al. (2014).

sessions are in Appendix 2. The information letters that were given to the participants at the start of the first learning session are in Appendix 3.

The stimuli were projected on a big screen at the front of the classroom. Before the stimuli were presented, a fixation asterisk appeared on the screen for one second. After this, a picture appeared with the nonword written underneath it for two seconds. During these two seconds, the nonword was also produced as a spoken word by a recorded voice.² The voice was heard through the speakers in the classroom. Then, the asterisk appeared again for one second before the next stimulus was presented.

Figure 2. *Distribution of the sets over the learning sessions*



Word Search Puzzle

In the test session of the WSP group a word search puzzle was used. Participants were handed a word search puzzle which contained all 20 nonwords. They had to find the nonwords they learned in the training sessions as fast as they could. Then they wrote the nonwords on the paper in the order they found them. They got 5 minutes to complete the test. There were two versions of the puzzle, to prevent that the words that were in the upper left corner of the puzzle would be found most frequently (see Appendix 4.1).

Picture Naming Task

During the test session of the PNT group a written picture naming task was carried out. The 20 pictures were presented in a random order in a booklet, and the participants had to write the corresponding nonword under each picture. The participants were given 10 minutes to complete the test session. There were two versions of the booklet, to prevent participants to cheat on the test (see Appendix 4.2).

² The recording of the audio was done by Iris van der Wulp at the Uil-OTS phonology lab in Utrecht.

Results

Participants and means

In total, 38 participants completed all training sessions and one test session. 26 participants took the PNT, while 12 participants completed the WSP. This last group suffered quite a high drop-out because during the test session, many students were absent and some of the students who were present had missed one or more training sessions. Otherwise, the WSP group would have been around 25 participants as well.

Of the 26 participants in the PNT group, 8 were male and 18 were female; 13 were in the distributed group and 13 were in the massed group. For the 12 participants in the WSP group, 3 participants were male and 9 were female; 7 were in the distributed group, while 5 were in the massed group. This is summarized in the table below, with the mean scores and standard deviation of each group:

Table 1. *Participant information with means and standard deviations.*

Number of participants(M, SD)

Group	PNT	WSP	Total
Male	8 (M = 6.25, SD = 5.83)	3 (M = 70.00, SD = 22.91)	11 (M = 23.64, SD = 31.87)
Female	18 (M = 16.39, SD = 20.85)	9 (M = 49.44, SD = 24.68)	27 (M = 27.41, SD = 26.90)
Distributed	13 (M = 15.00, SD = 21.31)	7 (M = 59.29, SD = 23.17)	20 (M = 30.50, SD = 30.43)
Massed	13 (M = 11.54, SD = 14.91)	5 (M = 48.00, SD = 28.64)	18 (M = 21.67, SD = 25.15)
Total	26 (M = 13.27, SD = 18.11)	12 (M = 54.58, SD = 25.00)	38 (M = 26.32, SD = 28.04)

(M = mean percentage correct, SD = standard deviation)

Five participants did not have Dutch as their first language. Although the nonwords were based on Dutch phonology, these participants were not excluded from the analysis. Their mean percentage correct on the tests did not differ significantly from the other participants (M non-Dutch = 25.00, SD = 31.22; M Dutch = 26.52, SD = 28.05); $t(36) = 0.11, p = 0.912$.

Scoring

The scoring was done as follows. When a participant wrote the correct name under the correct picture in the PNT, one point was given to the participant and the word. When the word was not found, or an incorrect answer was filled in, the participant and the word got zero points. This was roughly the same for the WSP, where participants just had to find the word instead of writing it under the matching picture. When an answer differed by one letter from the correct answer (for example: “bepim,” instead of “bepi” was written down), the answer was also scored as a correct answer. Especially in the WSP method, where the letter “m” was next to the word “bepi,” for example, this was considered to be necessary.

The next session will focus on answering the research questions. First, the sub-questions will be answered. Then, the main question will be answered. Scores were converted to percentage correct (min = 0.00, max = 100.00) to control for group size differences and to aid readability (raw scores ranged between 0.00 and 20.00, because 20 items were tested).

Is there an effect of version?

There should be no effect of version. Otherwise, this effect might interfere with the results of the tests. The hypothesis for this question was that there would be no effect of version, as they would be counterbalanced. Each item was tested in both versions of the test sessions for both

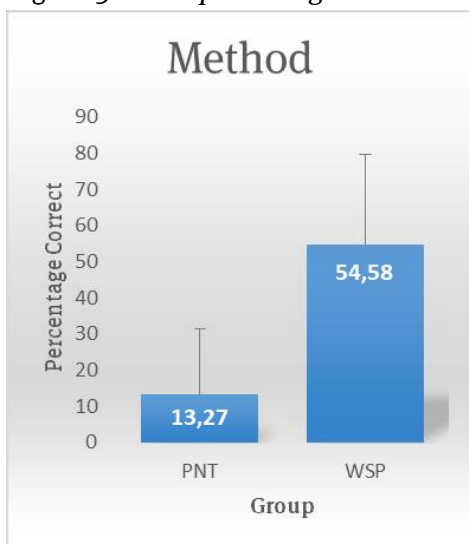
methods. The only difference was the order in which the items were tested. Participants were given long enough to find all words in both methods. Therefore, the order of testing should not matter.

There was indeed no effect of version. An independent-samples t-test was conducted to compare the mean percentage correct in participants who made the test versions A (N = 17) and versions B (N = 21). There was no significant difference in the scores for version A (M = 25.00, SD = 29.53) and version B (M = 27.38, SD = 27.46); $t(36) = -0.27, p = 0.799$. Moreover, for each method separately there was also no effect of version (PNT version A (M = 9.09, SD = 14.80), version B (M = 16.33, SD = 20.13): $t(24) = -1.01, p = 0.324$; WSP version A (M = 54.17, SD = 27.82), version B (M = 55.00, SD = 24.49): $t(10) = -0.55, p = 0.957$).

Is there an effect of method?

The hypothesis for this question was that the group who did the WSP would show a better performance than the group who did the PNT. This was indeed the case. A Generalized Linear Mixed Model (GLMM) was used. This model allows to determine the effects of method (and practice, as we will see later) on the mean percentage correct per participant, while accounting for the repeated testing of each participant for each item. Method and practice were included as fixed effects. Random effects were included for participant number and items. The reference category was zero: the score when an item was not found. There was a significant difference in the scores for the PNT (M = 13.27, SD = 18.11) and the WSP (M = 54.58, SD = 25.00) conditions found in the GLMM (coeff. = -2.354, SE = 0.778, $t = -3.024, p = 0.003$). The WSP group scored significantly higher than the PNT group. Figure 3 visualises this difference.

Figure 3. Mean percentage correct with SD for the PNT and WSP groups.



Is there an OoA effect visible in this study?

The hypothesis for this question was yes, hopefully there is an effect of OoA.

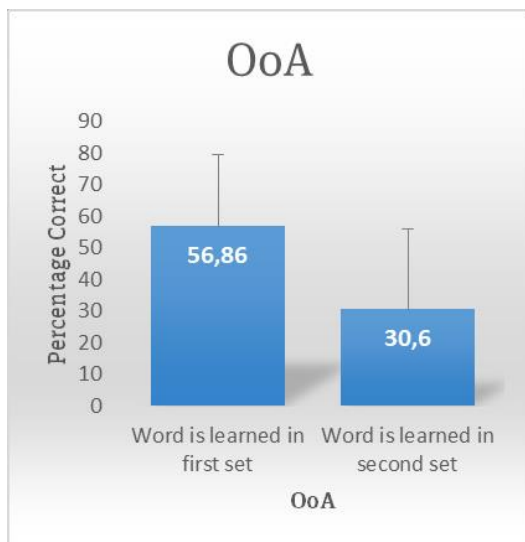
The 20 items were acquired in four different conditions:

1. OoA1 Practice1: the item was learned in the first set, with distributed practice.
2. OoA1 Practice2: the item was learned in the first set, with massed practice.
3. OoA2 Practice1: the item was learned in the second set, with distributed practice.
4. OoA2 Practice2: the item was learned in the second set, with massed practice.

OoA has two conditions: the word was learned in the first set, and the word was learned in the second set. Practice also has two conditions: massed, and distributed. The groups who learned the each item in a particular condition are different for the two sets of words (set A and set B, see Appendix 1). For each word in set A, participants from the WSP group were in condition 1 (N = 7) and 2 (N = 5). Participants from the PNT group were in condition 3 (N = 13) and 4 (N = 13). For each word in set B, participants from the PNT group were in condition 1 (N = 13) and 2 (N = 13). Participants from the WSP group were in conditions 3 (N = 7) and 4 (N = 5). In summary: participants from the WSP group were in condition 1 and 2 for each word in set A and conditions 3 and 4 for each word in set B. Participants from the PNT group were in conditions 3 and 4 for each word in set A and conditions 1 and 2 for each word in set B.

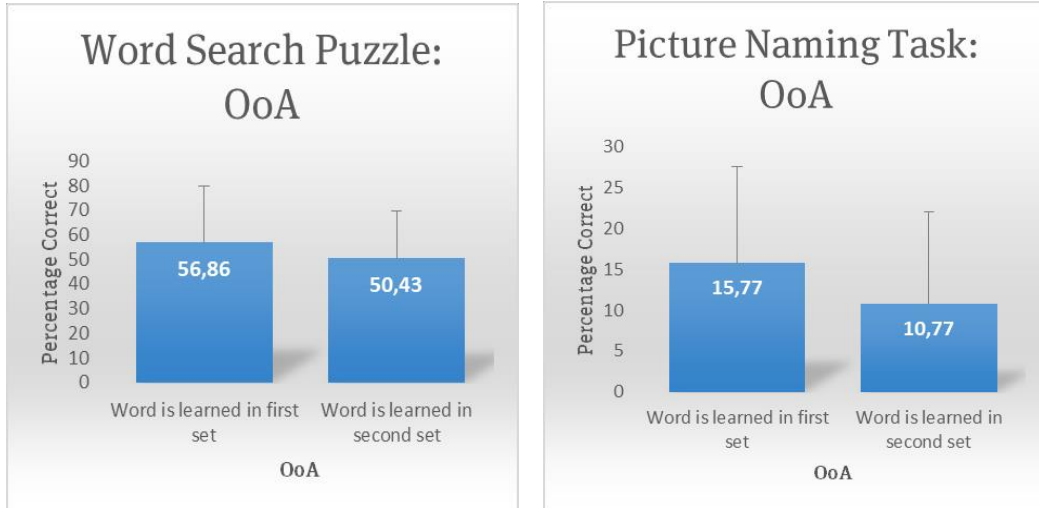
Due to the effect of method, OoA could not be included in the GLMM. For each word, two of the conditions above (with OoA 1) consisted of participants from one method and the other two conditions (with OoA 2) consisted of participants with the other method. Therefore, the effect of method masks the possible OoA effects. However, the mean percentage correct for words that were learned first is higher than for words that were learned later. Figure 4 visualises this difference. This shows a trend in the right direction, which can unfortunately not be tested statistically.

Figure 4. Mean percentage correct for the words in each OoA condition.



Moreover, for each word in one of the sets that was learned first in one method, there was a word in the other set that was learned later in the same method condition. Therefore, the data could be restructured to show mean percentages correct for items in the different OoA conditions per method. However, in this case two different items were analysed, instead of one item. The GLMM showed that there was no significant difference between the items (Estimate = 0.088, SE = 0.095, z = 0.929, p = 0.353). OoA for each word could not be calculated or statistically tested. Figure 5 shows the mean percentages correct in each OoA condition for each method.

Figure 5. Mean percentage correct for the words in each OoA condition in each method



Is there an effect of (massed and distributed) practice?

The hypothesis for this question was yes, hopefully there is an effect of practice. Overall, the GLMM found no effect of practice (coeff. = 0.586, SE = 0.833, t = 0.703, p = 0.482). However, participants in the distributed groups did have a higher mean percentage correct than participants in the massed groups. Figure 6 shows this difference.

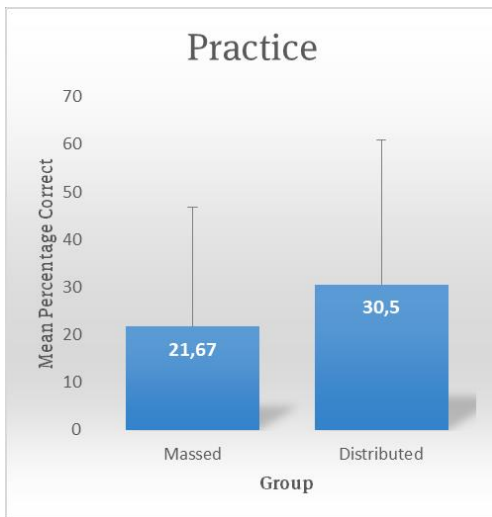
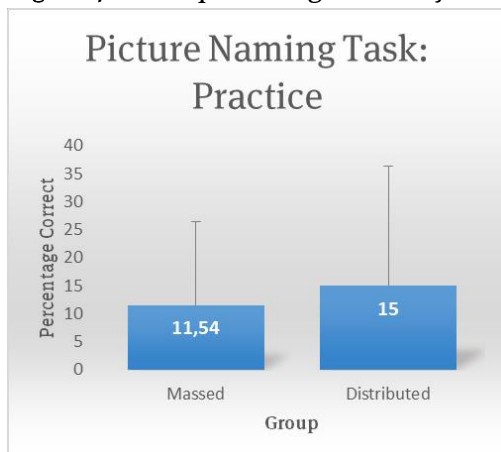


Figure 6. Mean percentage correct for the massed and distributed conditions.

Figure 7. Mean percentage correct for the massed and distributed conditions in the PNT group.



There was also no effect of practice in the PNT group. A paired-samples t-test was conducted to compare the means of the distributed (M = 15.00, SD = 21.31) and the massed (M = 11.54, SD = 14.91) conditions; $t(24) = 0.48$, $p = 0.636$ (see figure 7).

Figure 8. Mean percentage correct for the massed and distributed conditions in the WSP group.



There was also no effect of practice in the WSP group. A paired-samples t-test was conducted to compare the means of the distributed ($M = 59.29$, $SD = 23.17$) and the massed ($M = 48.00$, $SD = 28.64$) conditions; $t(10) = 0.76$, $p = 0.467$ (see figure 8).

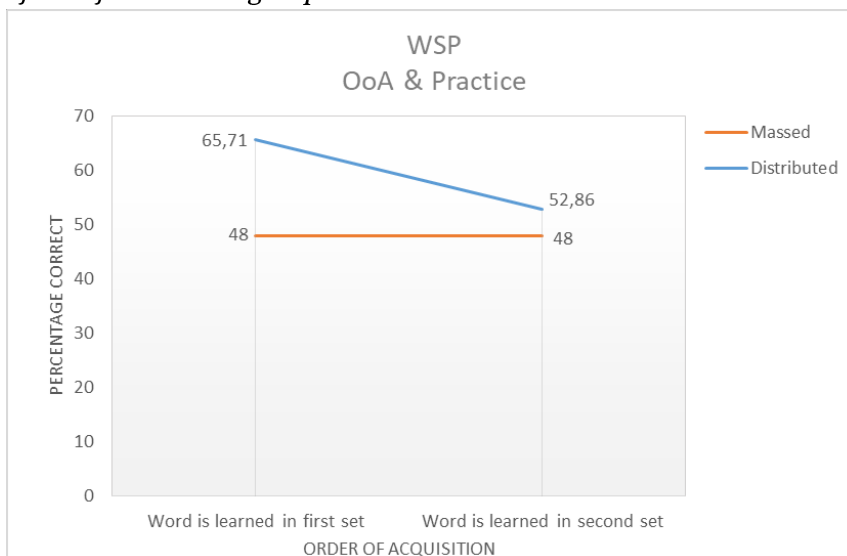
Do some words have higher scores than others?

Some items did have a higher score than others, but this difference in scores was not significant. The GLMM showed that there was no significant difference between the items (Estimate = 0.088, SE = 0.095, $z = 0.929$, $p = 0.353$). That shows that the items are comparable, which was the goal when they were chosen.

Can the AoA effect be (partially) explained by an effect of distributed practice?

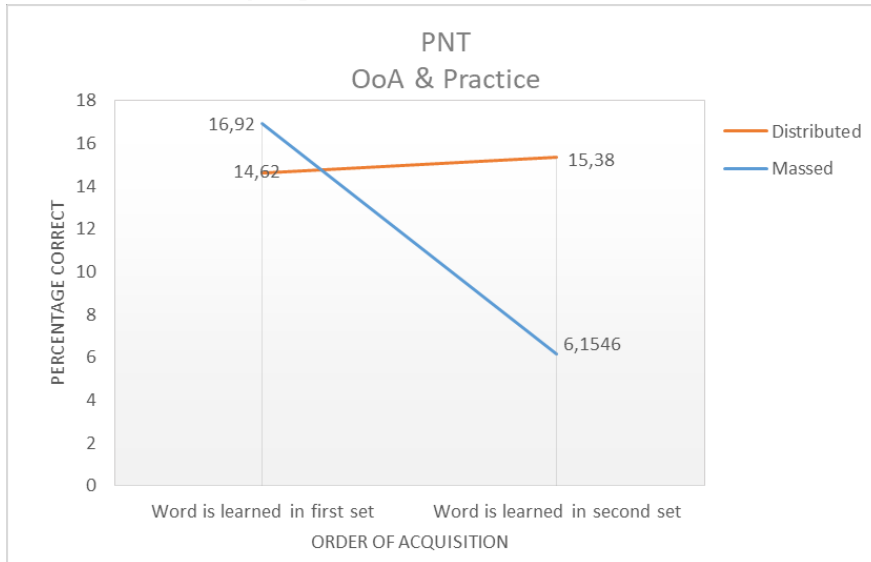
This is the main question of this thesis. Since OoA and Practice did not give significant effects, it is impossible to answer this question fully. However, for each method the means of the practice variable were plotted against the OoA. This was done with the corrected dataset per item per method, because otherwise OoA could not be calculated. For the WSP group, the graph in figure 9 shows that the mean percentage correct of the massed group was the same for each OoA condition. The means of the distributed condition did differ, where the mean percentage correct for a word that was learned in the second set was lower than the mean percentage correct for a word that was learned in the first set.

Figure 9. Massed and distributed practice mean percentages correct as a function of OoA for the WSP group



For the PNT group, it was the other way around. The massed group had a lower mean percentage correct when a word was learned in the second set than in the first set, and the distributed group scores roughly the same in both OoA conditions. This group even scored a little bit higher on words that were learned in the second set (see figure 11).

Figure 11. *Massed and distributed practice mean percentages correct as a function of OoA for the PNT group*



Discussion

The effect of method was expected. However, the fact that it would interfere with the OoA was not. The groups should have been splitted once more, so the total would be eight groups. Each massed or distributed group from each method should have had a group that learned set A first and a group that learned set B first. This would counterbalance the OoA, which was unfortunately not the case in this study. The recommendation for further research is to repeat this study with eight groups per method and with bigger groups. The current results seem to show no interaction between practice and OoA, but since both effects were not significant or not testable, this remains a question for further research.

Also, the number of items might have been too small. Because of time limits, there were only 3 training sessions of maximally 10 minutes. Further research should try more items, as that would generate more data which might show an OoA or practice effect more clearly.

The scores on the PNT test were very low, especially compared to the scores on the WSP. Many participants struggled to remember the words. Unfortunately, there was no time to do an additional test, like a translation test or a lexical decision task. Many participants said that they would have known more words if they could translate them to Dutch, for example. A recommendation for further research is to not use the PNT test method but a recognition task like a WSP, a translation task, lexical decision task or a task where participants need to match the words with the right picture (instead of coming up with the words themselves). Izura et al. (2011) did use a PNT, which is why that method was chosen for this study. However, unlike the study by Izura et al. (2011), this was a classroom experiment instead of an individual experiment. Maybe an individual experiment (where participants complete all sessions alone in front of a computer) would generate better results in general, and also on a PNT test.

Moreover, the way of learning in this research was quite static. Participants just saw and heard the words while trying to remember them. Other research used more interactive methods like making the participants repeat the word out loud or writing them down during the training sessions. This could also benefit the results (Carpenter et al., 2012; Cepeda et al., 2009; Izura et al., 2011; Sobel et al., 2011). This could also benefit the motivation of the students, as they were now quite bored while looking at the screen during the learning sessions.

In this study, there was no effect of practice, while other classroom experiments did show this effect (Carpenter et al., 2012; Sobel et al., 2011). Also, the mean scores of the groups for OoA and practice did show the expected directions of earlier words being remembered better and distributed practice benefitting word learning. This might also be due to the way of learning and the small number of participants, items and training sessions.

An argument for discussion that would remain if OoA and practice were significant and did not interact is the question of similarity between OoA and AoA. It is not clear if OoA and AoA generate the same effects. Maybe there is an interaction between distributed practice and AoA, and not between distributed practice and OoA, because AoA is by default distributed over a lifetime. More research on the similarities and differences between AoA and OoA needs to be done in order to determine if they generate the same effects.

Izura et al. (2011) favoured the Network Plasticity Hypothesis over the Semantic Locus Hypothesis, as they found OoA effects in the L2. However, if distributed practice would generate

stronger effects than OoA, this could complicate the Network Plasticity Hypothesis (and also the Semantic Locus Hypothesis). Maybe the distribution of word learning would be the cause of the OoA/AoA effect. The origin of the Network Plasticity Hypothesis, as stated in the Theoretical Background of this thesis is the connectionist model built by Ellis and Lambon Ralph (2000):

“Ellis and Lambon Ralph (2000) trained the model on an ‘early’ set and a ‘late’ set of words. When the early set was not repeated after training on the late set began, the early set was completely forgotten. This is called “catastrophic interference.” However, when the early set was still trained (although maybe less often than the late set), it was always remembered better by the model than the late set.”

This “catastrophic interference” happened when the early set was not repeated. This could be due to the fact that the learning of the early set was massed. Further research is necessary to find out if that was the case.

Conclusion

The aim of this research was to answer this question: *Can the AoA effect be (partially) explained by an effect of distributed practice? In other words: is there an interaction between a practice effect and an OoA effect and which one is stronger?*

To test this, participants learned nonwords in combination with pictures. These nonwords were split into two sets: set A and set B, to manipulate the order of acquisition. The distribution of learning was also manipulated. There was a massed group, which learned all words in one session, and a distributed group, which learned the words distributed over three weeks. Two test methods were used: a PNT and a WSP. Participants who made the WSP showed significantly higher scores than participants who made the PNT. With this kind of classroom experiment it is recommended to not use a PNT. Unfortunately, because of this effect of method OoA could not be statistically tested. There was no effect of practice. The current results seem to show no interaction between practice and OoA, but since both effects were not significant or not testable, this remains a question for further research.

References

- Brybaert, M., Stevens, M., De Deyne, S., Voorspoels, W., & Storms, G. (2014). Norms of age of acquisition and concreteness for 30,000 Dutch words. *Acta Psychologica, 150*, 80–84. <https://doi.org/10.1016/j.actpsy.2014.04.010>
- Carpenter, S. K., Cepeda, N. J., Rohrer, D., Kang, S. H., & Pashler, H. (2012). Using spacing to enhance diverse forms of learning: Review of recent research and implications for instruction. *Educational Psychology Review, 24*(3), 369–378.
- Catling, J., Dent, K., Preece, E., & Johnston, R. (2013). Age-of-acquisition effects in novel picture naming: A laboratory analogue. *The Quarterly Journal of Experimental Psychology, 66*(9), 1756–1763. <https://doi.org/10.1080/17470218.2013.764903>
- Cepeda, N. J., Coburn, N., Rohrer, D., Wixted, J. T., Mozer, M. C., & Pashler, H. (2009). Optimizing distributed practice: Theoretical analysis and practical implications. *Experimental Psychology, 56*(4), 236–246.
- Duñabeitia, J.A., Crepaldi, D., Meyer, A.S., New, B., Pliatsikas, C., Smolka, E., & Brybaert, M. (in press). MultiPic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology*. Retrieved from <http://www.bcbl.eu/databases/multipic/>
- Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: insights from connectionist networks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1103.
- Hirsh, K. W., Morrison, C. M., Gaset, S., & Carnicer, E. (2003). Age of acquisition and speech production in L2. *Bilingualism: Language and Cognition, 6*(2), 117–128.
- Izura, C., Pérez, M. A., Agallou, E., Wright, V. C., Marín, J., Stadthagen-González, H., & Ellis, A. W. (2011). Age/order of acquisition effects and the cumulative learning of foreign words: A word training study. *Journal of Memory and Language, 64*(1), 32–58.

- Juhasz, B. J. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin*, 131(5), 684.
- Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice Effects in Large-Scale Visual Word Recognition Studies: A Lexical Decision Study on 14,000 Dutch Mono- and Disyllabic Words and Nonwords. *Frontiers in Psychology*, 1.
<https://doi.org/10.3389/fpsyg.2010.00174>
- Lambon Ralph, M. A., & Ehsan, S. (2006). Age of acquisition effects depend on the mapping between representations and the frequency of occurrence: Empirical and computational evidence. *Visual Cognition*, 13(7–8), 928–948.
- Mermillod, M., Bonin, P., Méot, A., Ferrand, L., & Paindavoine, M. (2012). Computational Evidence That Frequency Trajectory Theory Does Not Oppose But Emerges From Age-of-Acquisition Theory. *Cognitive Science*, 36(8), 1499–1531. <https://doi.org/10.1111/j.1551-6709.2012.01266.x>
- Sobel, H. S., Cepeda, N. J., & Kapler, I. V. (2011). Spacing effects in real-world classroom vocabulary learning. *Applied Cognitive Psychology*, 25(5), 763–767.
- van der Wulp, I. M., & Verlangen, B. (2017). *Massed versus distributed practice and Age of Acquisition effects* (Paper).

Appendix

Appendix 1. Materials









List of Dutch words and nonwords corresponding to the pictures, with AoA of the Dutch words (Brysbaert et al., 2014), and reaction times and accuracy for the nonwords (Keuleers et al., 2010). The items were divided in Set A and Set B (see left column).

Set A	Word	Translation	No. list ³ (Brysbaert et al., 2014)	AoA (Brysbaert et al., 2014)	Nonword (Keuleers et al., 2010)	Reaction Time (ms) (Keuleers et al., 2010)	Accuracy (%) (Keuleers et al., 2010)
1	Schaar	Scissors	453	5,03	Vebe	551,7692308	1
2	Pet	Cap	203	5,21	Adim	547,8717949	1
3	Peer	Pear	42	5,06	Hatef	550,5897436	1
4	Pan	Pan	369	5,21	Eloes	540,5384615	1
5	Olifant	Elephant	579	5,08	Segwis	552,3421053	1
6	Vlinder	Butterfly	351	5,34	Asho	548,4615385	1
7	Tomaat	Tomato	326	5,11	Bepi	552,4736842	1
8	Slee	Sled	625	5,36	Zebbes	541,1794872	1
9	Telefoon	Telephone	509	5,12	Merpa	557,5128205	1
10	Bril	Glasses	733	5,38	Wigro	547,4864865	1
Set B							
1	Aardbei	Strawberry	381	5,12	Vrotin	552,5526316	1
2	Hoed	Hat	690	5,39	Derla	542,3589744	1
3	Snor	Moustache	494	5,12	Udep	514,3589744	1
4	Sleutel	Key	78	5,41	Epor	555,3421053	1
5	Fee	Fairy	7	5,15	Olret	542,7948718	1
6	Slak	Snail	152	5,44	Govon	549,5641026	1
7	Leeuw	Lion	15	5,16	Osfen	546,7894737	1
8	Rits	Zipper	223	5,45	Lodro	558,5263158	1
9	Spook	Ghost	139	5,17	Erdi	524,3846154	1
10	Honing	Honey	610	5,46	Mawel	555,1578947	1

³ The website <http://crr.ugent.be/archives/1602> offers a download link for the Excel file with all AoA data collected in Brysbaert et al. (2014).

Appendix 2. PowerPoint Presentations

Appendix 2.1. Set A. Presentation from left to right .

Set A	*	 vebe
*	 adim	*
 hatef	*	 eloes
*	 segwis	*
 asho	*	 bepi
*	 zebbes	*



merpa

*



wigro

*



zebbes

*



adim

*



bepi

*



segwis

*



eloes

*



merpa

*



vebe

*



asho

*



wigro

*



hatef

*



adim

*



merpa

*



vebe

*



wigro

*



segwis

*



zebbes

*



bepi

*



hatef

*



asho

*











eloes

*

*

Einde

Appendix 2.2. Set B. Presentation from left to right

Set B	*	 vrotin
*	 derla	*
 udep	*	 epor
*	 olret	*
 govon	*	 osfen
*	 lodro	*



erdi

*



mawel

*



erdi

*



udep

*



lodro

*



olret

*



vrotin

*



mawel

*



govon

*













epor

*



osfen

*	 derla	*
 lodro	*	 epor
*	 olret	*
 govon	*	 udep
*	 mawel	*
 vrotin	*	 osfen
*	 erdi	*











derla



*

Einde

Appendix 2.3. Set A & B. Presentation from left to right.

Sessie 3	*	 olret
*	 asho	*
 vebe	*	 wigro
*	 merpa	*
 govon	*	 lodro
*	 zebbes	*



segwis

*



mawel

*



eloes

*



epor

*



derla

*



hatef

*



adim

*



bepi

*



udep

*



erdi

*



osfen



vrotin



merpa



udep



eloes



govon



wigro



osfen



vrotin



epor



bepi



olret

*



hatef

*



adim

*



derla

*



erdi

*



asho

*



mawel

*



zebbes

*



vebe

*



segwis

*



lodro



asho



eloes



adim



epor



lodro



zebbes



bepi



segwis



govon



merpa



vebe

*



vrotin

*



osfen

*



olret

*



udep

*



hatef

*



derla

*



erdi

*



wigro

*



mawel

*

Einde

Appendix 3. Information letters for the participants

Appendix 3.1. Letters for group 1.

Informatiebrief onderzoek woordleren (English on the other side of this sheet)

Beste participant,

Bedankt voor je deelname aan mijn onderzoek over woordleren! Dit onderzoek is voor mijn Eindwerkstuk voor de Bachelor Taalwetenschap.

Wat houdt het in?

Je gaat in totaal 20 woordjes leren uit een niet-bestaande taal. Dit gebeurt via een PowerPoint presentatie, waarbij je telkens een plaatje ziet met daaronder een bijbehorend woord. Het plaatje geeft de betekenis van het woord aan. Terwijl het plaatje en het woord in beeld zijn, hoor je een stem het woord uitspreken via de speakers. Elk woord komt meerdere keren voorbij. Probeer de woorden en hun betekenis (het plaatje) voor jezelf te onthouden (ga dus niet overleggen). Verder hoeft je niets te doen.

Het aanleren van de woordjes gebeurt telkens op bovenstaande manier, in drie delen. Het eerste deel zal je vandaag leren, het volgende deel over een week en het laatste deel over twee weken. Over drie weken zal er een testsessie gehouden worden, om te zien wat je onthouden hebt.

Wanneer vindt het onderzoek plaats?

Het onderzoek wordt geheel tijdens de college-uren gehouden. Als je één van de vier colleges waarin dit onderzoek wordt uitgevoerd gemist hebt, kan je helaas niet meer meedoen.

Anoniem en vrijwillig

De gegevens die tijdens dit onderzoek verzameld worden zijn geheel anoniem en zullen nooit te herleiden zijn naar individuele personen.

Deelname aan dit onderzoek is vrijwillig. Mocht je om een of andere reden besluiten niet (meer) mee te willen doen, dan kan je dat aangeven bij de experimentleider (Iris). Voor overige vragen of opmerkingen met betrekking tot het onderzoek kan je me ook altijd mailen.

Participantengroep

Wanneer je deze informatiebrief hebt ontvangen betekent dat dat je in **groep 1** zit. Onthoud dit. Op de testsessie zal gevraagd worden in welke participantengroep je zit.

Nogmaals hartelijk dank voor je deelname!

Met vriendelijke groeten,

Iris van der Wulp

i.m.vanderwulp@students.uu.nl

Information letter research about word learning (Nederlands op de andere kant van dit blad)

Dear participant,

Thank you for taking part in my study about word learning. I am conducting this research for my Bachelor Thesis for the Bachelor Linguistics.

What does it involve?

You are going to learn a total of 20 words from a non-existing language. This will be done through a PowerPoint Presentation. You will see pictures with words below them. Each picture symbolizes the meaning of the word below it. You will also hear the pronunciation of the words through the speakers. Try to remember the words and their meanings (the pictures) for yourself (don't consult with your classmates about them). You don't have to do anything else.

The word-learning sessions will be as described above, in three parts. You will learn the first part today, the second part one week from now, and the third part in two weeks. In three weeks there will be a test session, to see what you remembered.

When does the research take place?

This research will be conducted during the classes of this course. If you missed one of the four classes during which this research is conducted, you can no longer take part in the study.

Anonymous and voluntary

The data that is collected with this study will always be anonymous and will never be traceable to an individual.

Taking part in this study is voluntary. If you no longer wish to participate, you can let the experiment leader (Iris) know. For further questions or remarks, feel free to send me an e-mail.

Participant group

When you have received this information letter, it means you are in **group 1**. Please try to remember this. At the test session, you will be asked to write down in which group you are.

Non-native speaker of Dutch

The words you are going to learn are designed with the Dutch phonology in mind. If you are not a native speaker of Dutch, or don't speak Dutch at all, you can still participate in this study. Please make sure to write **Non-native** on your answering sheet during the test session.

Thanks again for taking part in this study!

Kind regards,

Iris van der Wulp

i.m.vanderwulp@students.uu.nl

Appendix 3.2. Letters for group 2.

Informatiebrief onderzoek woordleren (English on the other side of this sheet)

Beste participant,

Bedankt voor je deelname aan mijn onderzoek over woordleren! Dit onderzoek is voor mijn Eindwerkstuk voor de Bachelor Taalwetenschap.

Wat houdt het in?

Je gaat in totaal 20 woordjes leren uit een niet-bestaande taal. Dit gebeurt via een PowerPoint presentatie, waarbij je telkens een plaatje ziet met daaronder een bijbehorend woord. Het plaatje geeft de betekenis van het woord aan. Terwijl het plaatje en het woord in beeld zijn, hoor je een stem het woord uitspreken via de speakers. Elk woord komt meerdere keren voorbij. Probeer de woorden en hun betekenis (het plaatje) voor jezelf te onthouden (ga dus niet overleggen). Verder hoeft je niets te doen.

Het aanleren van de woordjes gebeurt telkens op bovenstaande manier, in drie delen. Alle drie de delen zullen vandaag achter elkaar aangeleerd worden. Over één week zal er een testsessie gehouden worden, om te zien wat je onthouden hebt.

Wanneer vindt het onderzoek plaats?

Het onderzoek wordt geheel tijdens de college-uren gehouden. Als je één van de twee colleges waarin dit onderzoek wordt uitgevoerd gemist hebt, kan je helaas niet meer meedoen.

Anoniem en vrijwillig

De gegevens die tijdens dit onderzoek verzameld worden zijn geheel anoniem en zullen nooit te herleiden zijn naar individuele personen.

Deelname aan dit onderzoek is vrijwillig. Mocht je om een of andere reden besluiten niet (meer) mee te willen doen, dan kan je dat aangeven bij de experimentleider (Iris). Voor overige vragen of opmerkingen met betrekking tot het onderzoek kan je me ook altijd mailen.

Participantengroep

Wanneer je deze informatiebrief hebt ontvangen betekent dat dat je in **groep 2** zit. Onthoud dit. Op de testsessie zal gevraagd worden in welke participantengroep je zit.

Nogmaals hartelijk dank voor je deelname!

Met vriendelijke groeten,

Iris van der Wulp

i.m.vanderwulp@students.uu.nl

Information letter research about word learning (Nederlands op de andere kant van dit blad)

Dear participant,

Thank you for taking part in my study about word learning. I am conducting this research for my Bachelor Thesis for the Bachelor Linguistics.

What do I need to do?

You are going to learn a total of 20 words from a non-existent language. This will be done through a PowerPoint Presentation. You will see pictures with words below them. The pictures symbolize the meaning of the words. You will also hear the pronunciation of the words through the speakers. Try to remember the words and their meanings (the pictures) for yourself (don't consult with your classmates about them). You don't have to do anything else.

The word-learning sessions will be as described above, in three parts. You will learn all these parts today, in a row. In one week there will be a test session, to see what you remembered.

When does the research take place?

This research will be conducted during the classes of this course. If you missed one of the two classes during which this research is conducted, you can no longer take part in the study.

Anonymous and voluntary

The data that is collected with this study will always be anonymous and will never be traceable to an individual.

Taking part in this study is voluntary. If you no longer wish to participate, you can let the experiment leader (Iris) know. For further questions or remarks, feel free to send me an e-mail.

Participant group

When you have received this information letter, it means you are in **group 2**. Please try to remember this. At the test session, you will be asked to write down in which group you are.

Non-native speaker of Dutch

The words you are going to learn are designed with the Dutch phonology in mind. If you are not a native speaker of Dutch, or don't speak Dutch at all, you can still participate in this study. Please make sure to write **Non-native** on your answering sheet during the test session.

Thanks again for taking part in this study!

Kind regards,

Iris van der Wulp

i.m.vanderwulp@students.uu.nl

Appendix 4. Test sessions

Appendix 4.1. Word Search Puzzles

Versie A

Testsessie (English below)

Draai dit blad pas om als dat gezegd wordt.

Groep: _____ Geslacht: _____

Moedertaal: _____

Instructies:

Op de achterkant van dit blad staat een woordzoeker.

Probeer alle 20 woorden die je hebt geleerd te vinden.

Alle woorden staan van links naar rechts.

Je krijgt 5 minuten om alle woorden te vinden.

Schrijf de gevonden woorden in de juiste volgorde onderaan op de lijntjes.

Test session

Don't turn this page until you are told to do so.

Group: _____ Gender: _____

Native language: _____

Instructions:

On the back of this page is a word search puzzle

Try to find all 20 words that you learned

All words are written in the puzzle from left to right

You get 5 minutes to find all words

Write the words you found in the right order on the lines below the puzzle.

versie

A

i	f	r	i	y	r	f	a	u	b	n	g	v	t	v	m	p	r	y	k	n	j	s
c	x	s	f	e	o	a	m	p	l	h	z	v	d	e	r	l	a	h	h	t	b	e
g	a	l	k	y	e	l	o	e	s	e	y	t	t	u	v	a	m	x	b	a	c	x
n	s	b	v	l	t	u	k	g	z	y	e	c	y	h	n	u	o	z	u	q	t	k
g	v	e	b	e	v	s	v	u	k	i	l	z	p	a	s	h	o	c	k	i	i	o
i	d	f	i	q	v	f	r	v	i	v	z	k	s	b	k	v	h	x	w	q	x	e
v	r	o	t	i	n	n	o	z	k	v	v	l	o	d	r	o	c	a	y	c	m	k
h	s	d	a	d	c	n	a	w	p	w	l	m	b	e	m	z	d	g	s	b	p	u
y	b	f	f	i	h	m	v	k	f	m	m	f	d	q	u	e	g	g	j	o	q	n
o	g	f	h	q	b	g	i	m	e	r	p	a	u	a	m	r	c	t	j	b	j	j
s	g	q	s	e	g	w	i	s	p	p	v	t	x	w	g	x	d	r	k	u	g	u
q	f	d	f	e	v	a	s	d	t	f	e	p	o	r	j	b	r	g	k	g	w	j
v	q	c	m	l	y	y	e	r	d	i	j	e	t	o	o	a	x	c	i	w	g	m
q	a	u	g	a	g	g	s	g	b	a	h	a	t	e	f	r	e	j	r	f	f	h
e	z	e	b	b	e	s	h	h	k	w	n	j	q	j	s	x	j	u	q	y	u	t
z	i	j	e	w	p	z	b	e	p	i	o	w	w	e	k	r	p	x	j	r	p	k
r	w	o	s	f	e	n	u	h	x	f	t	x	b	k	g	s	p	s	w	y	o	z
h	o	y	s	q	v	h	n	i	q	q	n	m	a	w	e	l	g	n	w	x	d	f
b	v	w	i	g	r	o	q	l	f	o	t	l	s	h	e	t	k	p	y	q	y	v
j	q	q	p	t	h	v	a	o	l	r	e	t	l	t	j	a	i	f	o	w	l	b
x	h	w	d	m	f	g	o	v	o	n	h	f	e	k	w	a	q	o	g	c	d	a
y	i	g	f	r	u	v	b	i	d	h	a	z	u	c	y	k	s	h	c	b	j	y
j	g	c	a	d	i	m	a	o	z	v	v	x	t	u	d	e	p	p	x	a	o	o
k	h	l	m	l	w	x	m	h	t	k	e	b	o	y	j	z	i	q	c	t	v	m

1 _____

11 _____

2 _____

12 _____

3 _____

13 _____

4 _____

14 _____

5 _____

15 _____

6 _____

16 _____

7 _____

17 _____

8 _____

18 _____

9 _____

19 _____

10 _____

20 _____

Versie B

Testsessie (English below)

Draai dit blad pas om als dat gezegd wordt.

Groep: _____ Geslacht: _____

Moedertaal: _____

Instructies:

Op de achterkant van dit blad staat een woordzoeker.

Probeer alle 20 woorden die je hebt geleerd te vinden.

Alle woorden staan van links naar rechts.

Je krijgt 5 minuten om alle woorden te vinden.

Schrijf de gevonden woorden in de juiste volgorde onderaan op de lijntjes.

Test session

Don't turn this page until you are told to do so.

Group: _____ Gender: _____

Native language: _____

Instructions:

On the back of this page is a word search puzzle

Try to find all 20 words that you learned

All words are written in the puzzle from left to right

You get 5 minutes to find all words

Write the words you found in the right order on the lines below the puzzle.

ver-
sie B

i	b	k	u	k	o	r	x	z	y	r	a	r	k	a	c	s	y	e	w	l	t	w
t	m	l	h	j	r	g	g	s	d	r	h	f	u	d	e	p	y	e	h	m	x	t
e	f	s	f	e	a	d	i	m	f	i	c	n	w	r	f	r	x	i	u	d	b	q
d	k	t	l	r	s	s	c	i	j	u	x	v	l	e	u	u	a	h	y	u	m	b
m	g	o	v	o	n	u	l	n	r	g	g	n	w	o	l	r	e	t	h	b	k	l
g	u	a	d	o	q	q	y	s	k	i	d	g	n	m	l	u	s	c	k	x	f	r
w	i	g	r	o	x	s	z	c	l	v	d	m	a	w	e	l	v	a	j	k	i	o
d	p	l	o	i	z	o	g	w	r	n	l	y	s	y	j	r	o	c	j	s	m	e
s	h	i	e	o	c	c	u	g	y	w	m	n	v	r	e	q	y	t	q	a	i	k
f	u	g	e	g	z	v	p	o	s	f	e	n	z	z	y	f	v	q	x	g	v	u
e	u	x	b	e	p	i	m	l	h	n	l	m	e	q	j	z	x	q	b	j	l	i
q	e	d	x	a	c	p	r	l	x	v	z	e	b	b	e	s	m	y	g	e	p	f
i	o	i	c	q	v	p	h	a	t	e	f	c	s	a	z	q	b	j	n	o	x	f
a	r	i	g	q	a	l	s	p	m	u	e	r	d	i	n	o	y	g	f	r	v	y
p	e	p	o	r	l	k	t	l	t	g	s	z	r	i	m	s	u	c	a	j	x	e
i	t	o	s	e	a	d	s	e	g	w	i	s	c	q	c	t	t	v	j	x	q	u
g	h	m	e	r	p	a	c	a	f	m	l	i	s	i	g	p	e	i	y	p	z	m
v	i	o	k	p	g	c	o	j	p	f	g	l	o	d	r	o	e	f	r	e	v	b
j	d	v	r	o	t	i	n	h	s	f	d	h	g	s	m	z	w	z	o	k	x	p
l	d	k	n	f	s	a	k	a	s	h	o	r	s	s	j	o	h	b	m	f	y	w
k	a	m	r	a	j	v	e	b	e	k	o	e	z	f	y	f	z	c	p	u	x	b
j	t	b	y	z	l	v	i	x	q	w	l	l	q	d	p	i	z	y	i	m	u	t
e	j	v	e	l	o	e	s	p	q	r	z	o	p	d	e	r	l	a	u	r	w	q
g	i	g	i	d	m	b	z	s	c	x	g	t	g	i	j	r	i	j	u	y	p	k

1 _____

11 _____

2 _____

12 _____

3 _____

13 _____

4 _____

14 _____

5 _____

15 _____

6 _____

16 _____

7 _____

17 _____

8 _____

18 _____

9 _____

19 _____

10 _____

20 _____

Testsessie (English below)**Open dit boekje pas als dat gezegd wordt.**

Groep: _____ Geslacht: _____

Moedertaal: _____

Instructions:

In dit boekje staan de plaatjes van de trainingssessie(s).
Schrijf op het lijntje onder elk plaatje in dit boekje het
bijbehorende woord dat je geleerd hebt in de
trainingssessies. Je krijgt hiervoor 10 minuten.

Test session**Don't open this booklet until you are told to do so.**

Group: _____ Gender: _____

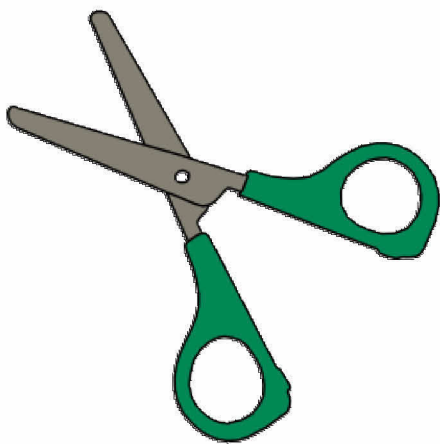
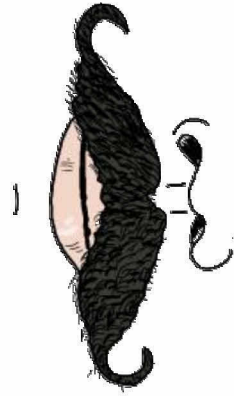
Native language: _____

Instructions:

This booklet contains the pictures you saw during the
training session(s). Write the word you learned in
combination with the picture on the line under each
picture in this booklet. You get 10 minutes to do so.

Appendix 4.2. Picture Naming Tasks

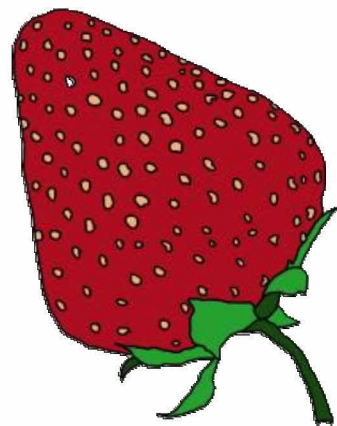
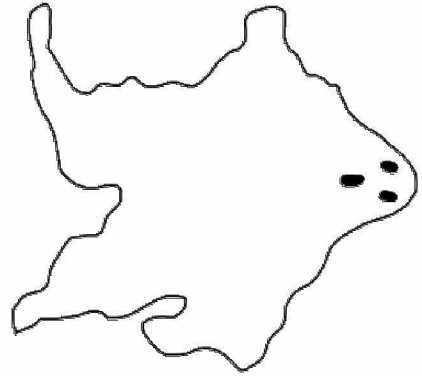
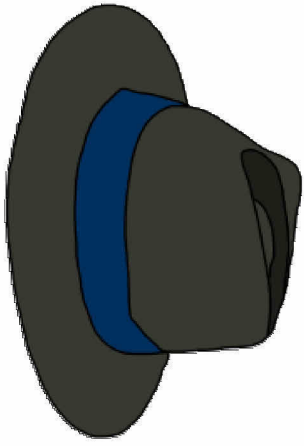
Versie A



Versie A

Appendix 4.2. Picture Naming Tasks

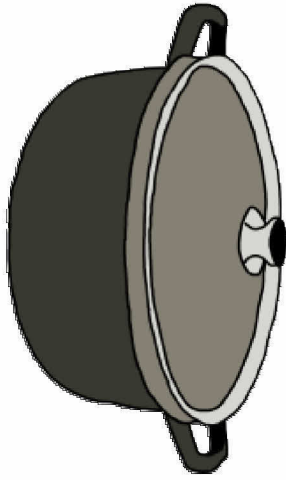
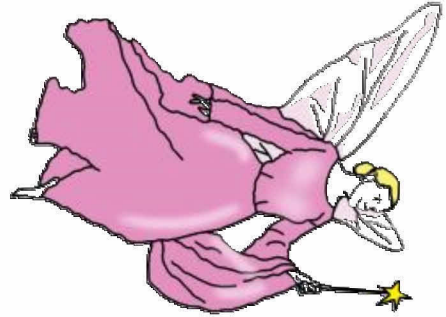
Versie A



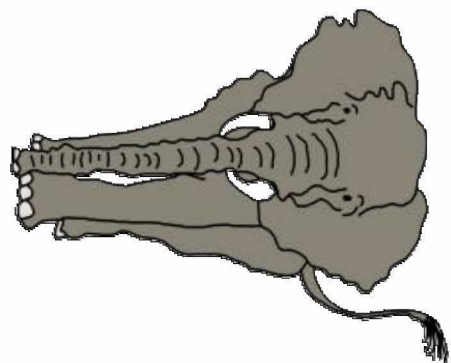
Versie A

Appendix 4.2. Picture Naming Tasks

Versie A

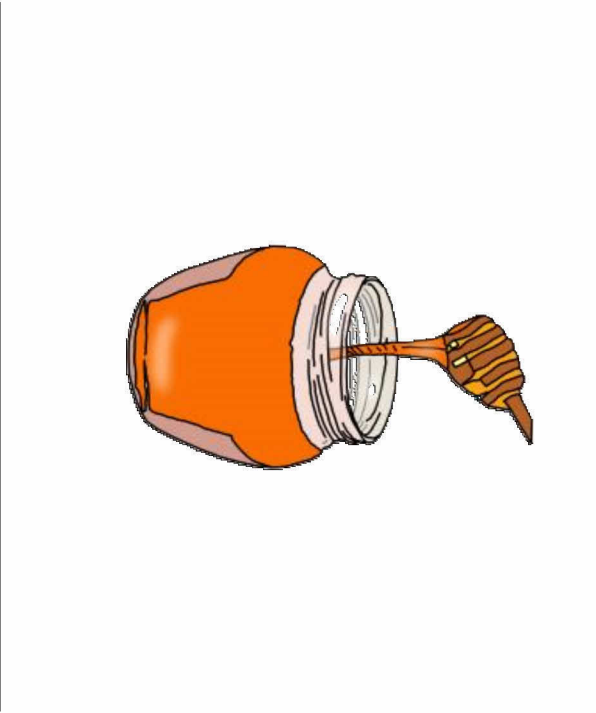
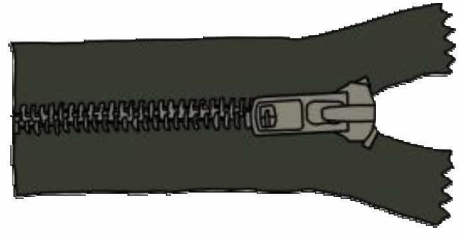


Versie A

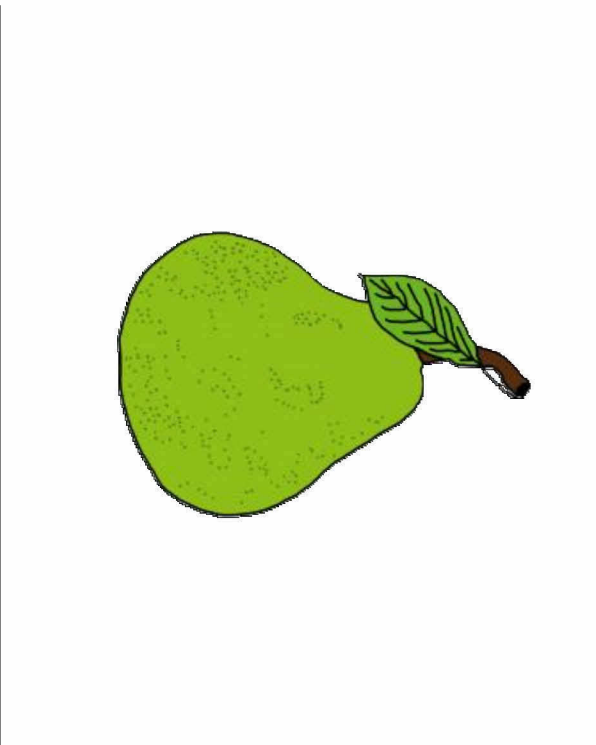
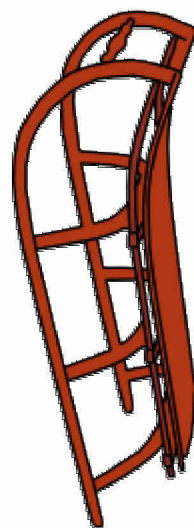


Appendix 4.2. Picture Naming Tasks

Versie A

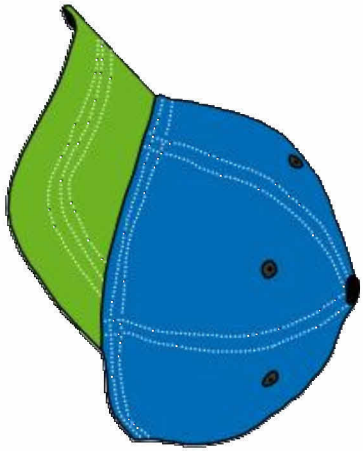
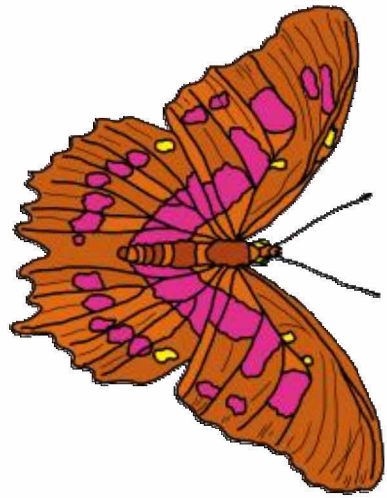


Versie A

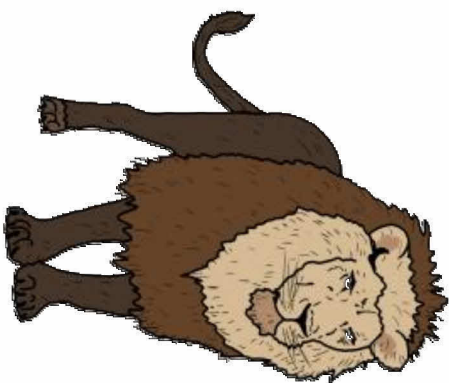
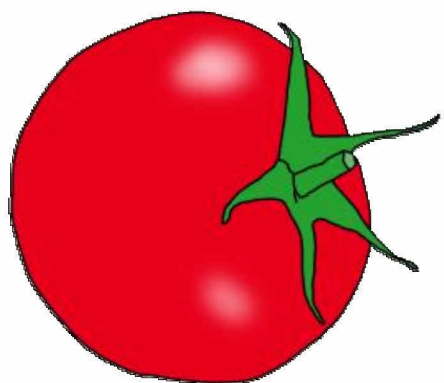


Appendix 4.2. *Picture Naming Tasks*

Versie A



Versie A



Testsessie (English below)**Open dit boekje pas als dat gezegd wordt.**

Groep: _____ Geslacht: _____

Moedertaal: _____

Instructies:

In dit boekje staan de plaatjes van de trainingssessie(s).
Schrijf op het lijntje onder elk plaatje in dit boekje het
bijbehorende woord dat je geleerd hebt in de
trainingssessies. Je krijgt hiervoor 10 minuten.

Test session**Don't open this booklet until you are told to do so.**

Group: _____ Gender: _____

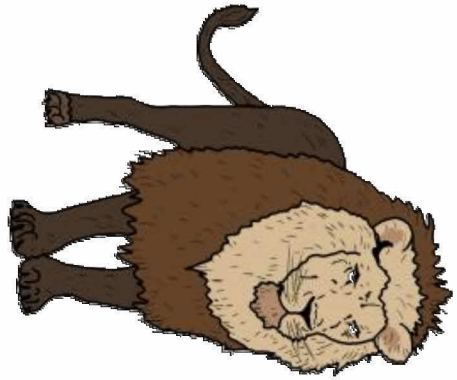
Native language: _____

Instructions:

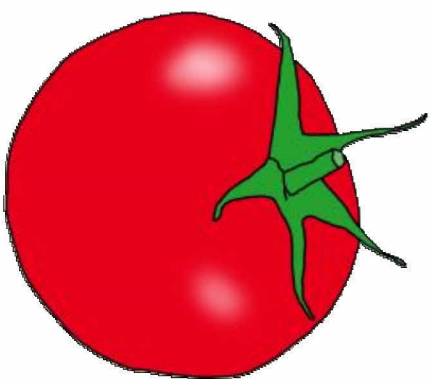
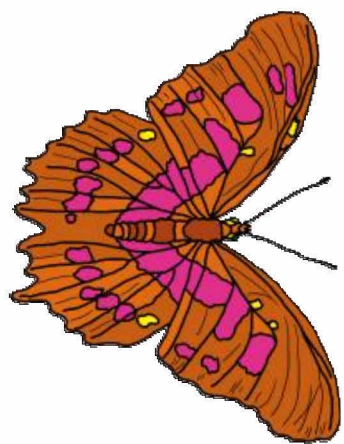
This booklet contains the pictures you saw during the
training session(s). Write the word you learned in
combination with the picture on the line under each
picture in this booklet. You get 10 minutes to do so.

Appendix 4.2. Picture Naming Tasks

Versie B

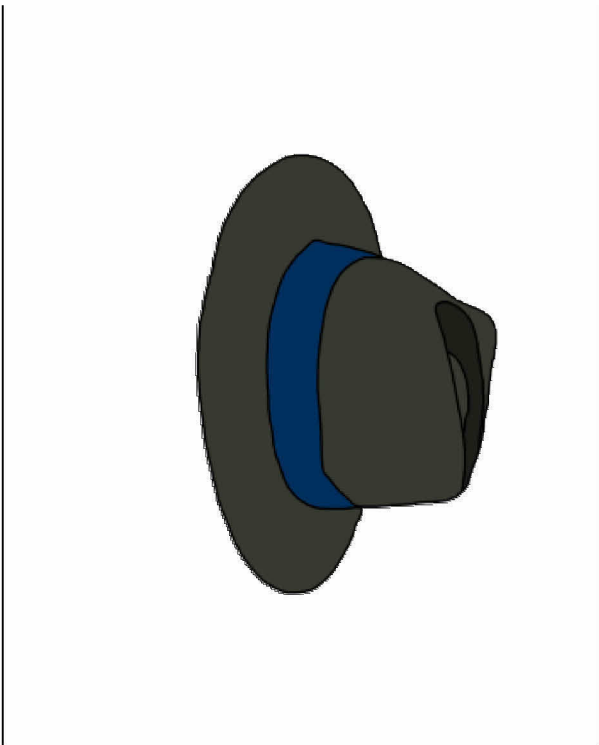
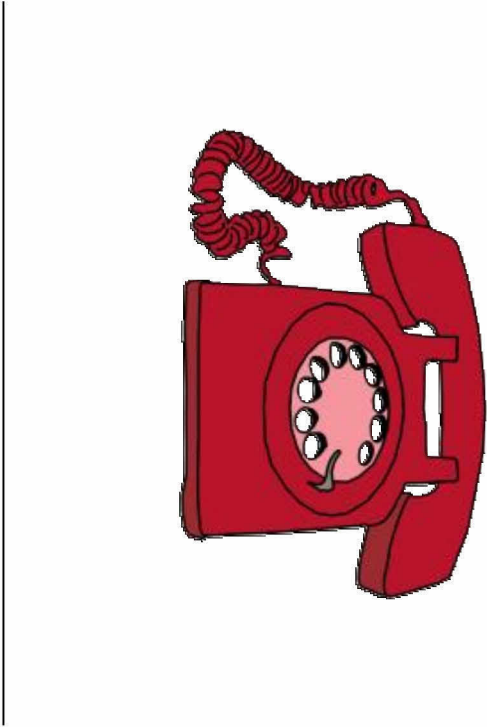
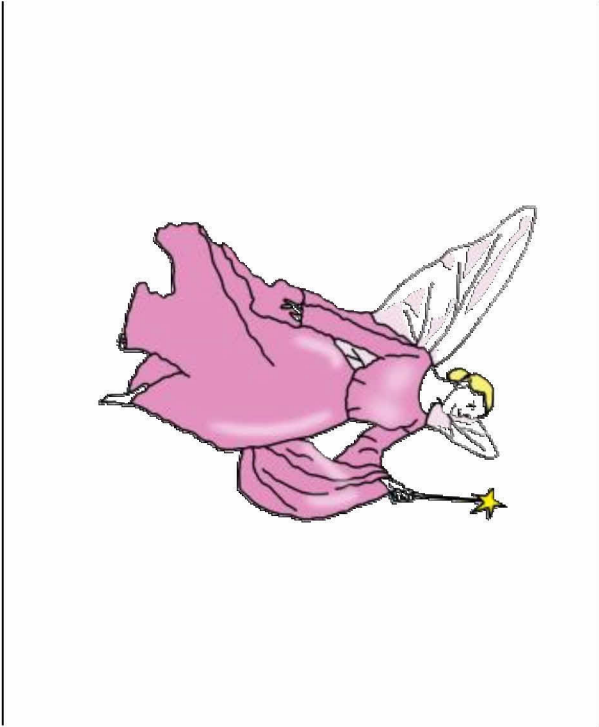


Versie B



Appendix 4.2. *Picture Naming Tasks*

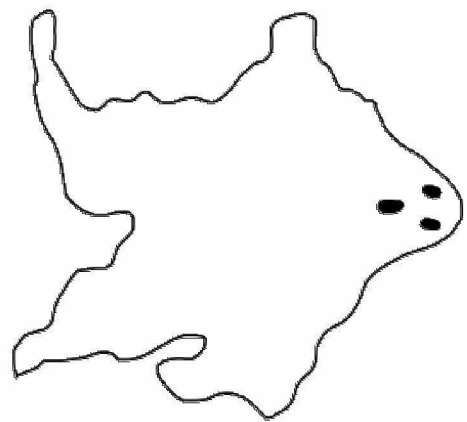
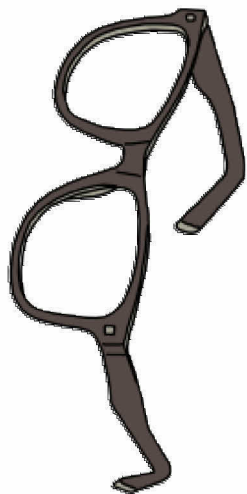
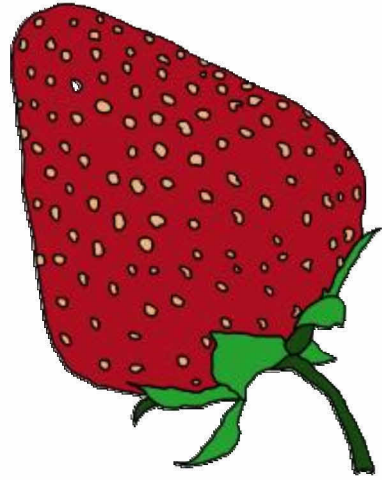
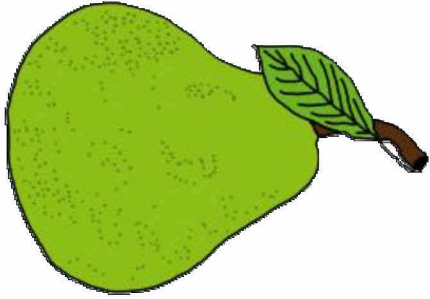
Versie B



Versie B

Appendix 4.2. Picture Naming Tasks

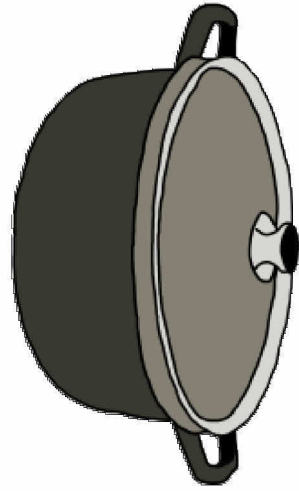
Versie B



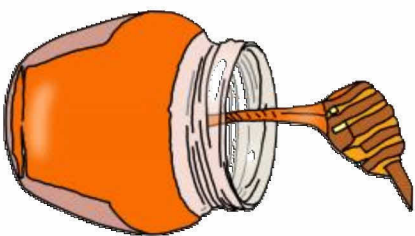
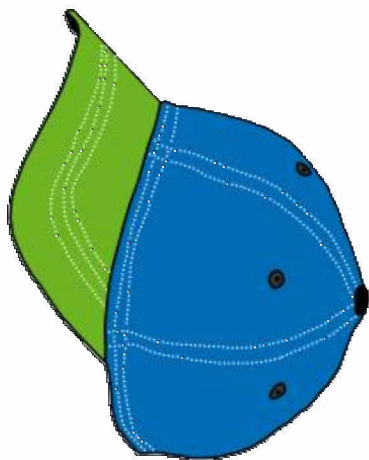
Versie B

Appendix 4.2. Picture Naming Tasks

Versie B

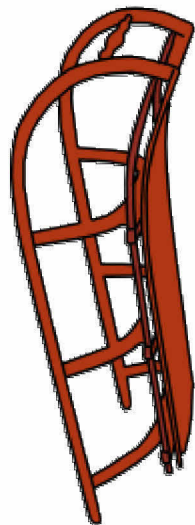
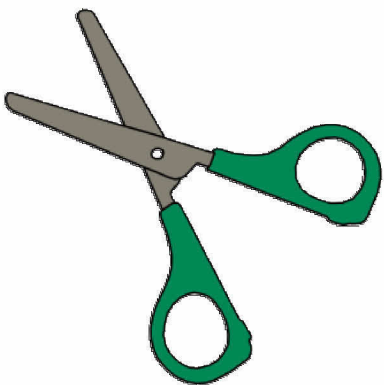
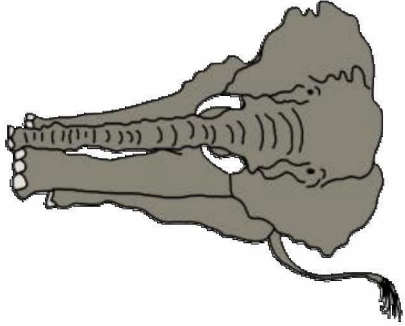


Versie B



Appendix 4.2. Picture Naming Tasks

Versie B



Versie B