The Effect of Images and their Memorability on Foreign Language Vocabulary Learning

Merel van Duivenboden

s4355571

Department of Psychology, University of Groningen PSB3E-BT15: Bachelor Thesis

Group 28

Supervisor: Dr. Mark Nieuwenstein

Second evaluator: Dr. Stefanie Enriquez Geppert

In collaboration with: Marjolijn Drenthen and Jip Konings.

February 27, 2023

Abstract

Although multimedia learning has become more popular over the years, current research on the benefits of using multimedia with vocabulary learning is still all but conclusive. This bachelor thesis aims to provide a practical answer to the question whether images, especially memorable images, should be used to enhance vocabulary learning. In a 2 X 2 experimental design, participants were asked to learn Dutch translations of Finnish words. Participants were all Dutch speaking first-year Psychology students ($N = 40$). Participants were shown Finnish words, their Dutch translations, and, depending on the assigned condition, an image that represented the words shown. These images had high or low memorability scores. Short-term and long-term retention scores were assessed as measure for vocabulary learning. Repeated Measures ANOVA was used to analyze the data. Our data showed that for concrete words, image presence and high image memorability did not have beneficial effects on vocabulary learning. Therefore, we would not advise to use (memorable) images for enhancement of vocabulary learning. One reason why no effect was found could be based on the concreteness of words. Perhaps, vocabulary learning of more abstract words could benefit from image presence and memorability. Therefore, further research should include more abstract words.

Keywords: Foreign vocabulary learning, image memorability, multimedia learning

The Effect of Images and their Memorability on Foreign Language Vocabulary Learning

Learning new vocabulary is part of daily life. Obviously, young children will encounter lots of new words in their native language during their development. In high-school, teenage students will still encounter new vocabulary in their native language as well, but they probably also will be expected to learn vocabulary in a foreign language. Second language acquisition is a process that becomes more difficult as people age, as shown by Hartshorne et al. (2018). Specifically, children who learn a new language can become as proficient as a native speaker if they start learning the age before the age of 10 years old. After that age, children are still able to learn a new language quite rapidly, but they will not become as proficient as a native speaker. After the age of 17 or 18, learning a new language becomes more difficult. Therefore, it is interesting to investigate if there is a way to make second language learning more efficient for younger and older students.

There are many different strategies that students can use to study second language vocabulary. The task of learning second language (L2) vocabulary can be considered as a paired-associate learning task in which students need to learn the association between an L2 word and its translation. Although there are many strategies that students can apply to study paired-associate learning tasks, not all strategies are equally effective for long-term memory. In general, deep processing of study material is more effective for learning than shallow processing (Reisberg, 2019). One example of shallow processing is the drill-and-practice method (Hald et al., 2015). This form of studying is characterized by maintenance rehearsal: A form of rehearsal where items are repeated over and over again in the working memory system to try to achieve long-term memory storage (Reisberg, 2019). A more effective method relies on elaborative rehearsal, whereby connections are made or strengthened between new and existing knowledge. According to Carpenter et al. (2009), elaborative retrieval creates richer memory representations and additional retrieval routes. This would make it easier to retrieve knowledge, and therefore easier to learn and retrieve foreign vocabulary. One form of elaborative rehearsal is deep processing. When trying to learn novel words, thinking about the meaning of these words is a form of deep processing. In contrast to shallow processing, where a person for instance only looks at the form (i.e. spelling) of words, deep processing typically leads to better long-term memory retention (Reisberg, 2019). In short, to make vocabulary learning more effective, students should engage in elaborative rehearsal. But how can this be applied in practical settings?

One method to enhance deep encoding during vocabulary learning is by using multimedia. Multimedia learning can be defined as learning through words and visual illustrations together (Mayer, 2002). Words can be represented either through text or through auditive stimuli. Visual illustrations can be presented either statically or dynamically, meaning that all kinds of visual information like photos, tables, videos or animations can be used. With the rise of technological advancement, the use of multimedia has become more prevalent in educational settings. In the Netherlands, many high-school students are for example introduced to applications that are designed to enhance vocabulary leaning by adding multimedia. These applications include WRTS (2022), SlimStampen (Van Rijn et al., 2009; Wilschut et al., 2021), and Quizlet (2023). One explanation why multimedia learning might provide its benefits derives from the Dual-Coding Theory by Clark and Paivio (1991). According to this theory, pictures are remembered better than words, because they activate both visual and verbal codes, in contrast to words, which are more likely to only activate verbal codes. By activating two codes, a visual and a verbal code, a picture would generate a strong memory trace. As explained by Van den Broek et al. (2021), combining words and pictures could enhance learning, when compared to learning based on words alone. The combination of different resources of information requires the learner to more actively encode the material that needs to be studied. When done efficiently, this would lead to more effortful organization of relevant information and it allows the learner to integrate the new information with existing knowledge (Ainsworth & Loizou, 2003; Butcher, 2006). Consequently, the addition of images next to text could lead learners to develop a more integrated concept in a richer network of knowledge (Butcher, 2006, 2014; Fiore et al., 2003) Therefore, multimedia learning is a method that can increase deep processing.

Support for the Dual-Coding Theory is mixed in studies that show that foreign words were better remembered by showing pictures than words in annotations. Abkulut (2007) conducted a study in which Turkish English Foreign Language (TEFL) students were asked to read a text that was annotated with either definitions of words, definitions coupled with associated pictures or definitions coupled with associated short videos. Participants who were shown both text and visuals performed better on later incidental vocabulary tests than participants with annotated text only. The study of Shahrokni (2009) and showed similar results, with benefits for a combination of text and images compared to textual or pictorial glosses only. Furthermore, the study of Yeh and Wang (2003) showed that an annotation with text and images was more effective than annotations with only text, or text, images and auditive stimuli combined. These findings suggest that there is an optimum in the amount of information that is helpful in vocabulary learning. However, the study of Acha (2009) cannot support this idea. Similar to the design of Akbulut (2007), participants in the study of Acha (2009) were provided a text in which key words were presented with verbal or visual annotations, or both. In this study, visual annotations were not helpful at all. Recall was better for participants that were shown annotations with verbal information only. Based on these results, it seems to be not beneficial to only show illustrations in text annotations. The effects of combining text and illustrations in annotations are mixed across different studies using a similar methodology, and cannot fully support the Dual-Coding Theory.

Regarding ‘simple’ vocabulary learning (i.e. the presentation of the L2 by itself instead of in-text), effects for multimedia learning have been mixed as well. There seems to be a general tentative finding that adding images is only beneficial for vocabulary learning of abstract words. Firstly, support for this statement comes from studies that included concrete words and found no significant benefits for presenting the foreign word with an image. For instance, Lotto and de Groot (1998) asked Dutch participants to learn concrete Italian words. These words were paired with either an image that represented the word, or the Dutch translation. There were three trials in which participants were able to study each word, before they were asked to take the first retention test. During the retention, one of the cues was presented, either the Dutch translation or the image. Then, 3 more learning trials were used. Then, a final retention test was taken. The results showed that there were marginally significant results on the analysis of participants when comparing the learning strategy. Participants were better at retention when they learned the Italian word based on the Dutch translation, than through the presence of the image. Other research with similar methodology did not find effects for multimedia learning either (e.g. Chen,1990; Cohen & Johnson, 2011). Furthermore, other studies included both concrete and abstract words (Farley et al., 2012; Farley et al., 2014) and found only positive effects for abstract words. Participants were shown the L2 word with either an image that represented the meaning of the L2 word, or with the L1 translation of the L2 word. During cued-recall, the L2 was supplied and the L1 translation was to be written down. These studies showed that recall of abstract words was better when pictures were shown during encoding, but this effect was not found for concrete words. Consequently, these findings combined suggest that multimedia learning could be beneficial, but only for abstract vocabulary learning.

These findings are consistent with the Dual-Coding Theory, since the presentation of abstract words activates only verbal codes. When an image is presented instead, the image can activate more codes, both verbal and visual, and could thereby facilitates better elaborative encoding. In contrast, concrete words are highly imaginable and can therefore activate both verbal and visual codes. Therefore, an image would not activate more codes than the word itself, which could explain why there is no significant difference in showing images versus L1 translations during encoding.

One limitation of the fellow studies mentioned is that participants were only presented either L1 translations, or images (Chen,1990; Cohen & Johnson, 2011, Lotto & De Groot, 1998, Farley et al., 2012; Farley et al., 2014). It remains unclear what the effects of combining pictures and L1 translations are on 'simple' vocabulary learning. As one might predict, the combination of text and images might be beneficial in some cases, as it turned out to be in some studies concerning text annotations (e.g. Akbulut, 2007; Shakroni, 2009; Yeh & Wang, 2003). On the contrary, the addition of images next to L1 translations could potentially have detrimental effects as well. One explanation for this prediction derives from the redundancy effect (Kalyuga & Sweller, 2014). The redundancy effect is based on the cognitive load theory, that supposes that human cognition has two parts that are related to information processing: working memory and long-term memory. Working memory has a limited capacity and limited duration in which information can be remained active. Long-term memory has no such limitations (Sweller, 2006). When the working memory is exposed to redundant and useful information at the same time, coordination is necessary, which will increase the cognitive load. Since the working memory only has a limited capacity, the increase of cognitive load might have detrimental effect on the encoding of useful information. In the case of images that are presented next to concrete words, , use of images is often not necessary since the original source already requires all the information that is necessary. This way, the cognitive load might unnecessarily increase, which could lead to poorer encoding of the actual translation, especially when limited attention is divided between the image and the word pair. In these cases, adding pictures is not beneficial for learning, which is in line with research findings (Sweller et. al, 1998). Overall, research on the benefits of using multimedia with vocabulary learning is all but conclusive.

The aim of the present research is therefore to further investigate this question. Instead of varying the properties of the words that will be encoded, this research focuses on the properties of the images. In specific, we examine the effect of memorability of images. According to the picture superiority effect, pictures are remembered better than words (Carpenter & Olson, 2011; Durso & Johnson, 1980) However, not every image is remembered as well as others. According to Isola et al. (2011), some pictures are remembered consistently better than others. In this study, a large number of participants ($N = 665$) were asked to perform a task where they were asked to look at a sequence of images and indicate whether they had seen an image before. Based on hit-rates of correctly recognized target images, memorability of these images was established. The pictures in this sample represented all kinds of semantic categories. Memorability of images turned out to be largely consistent between participants. Bainbridge et al. (2013) also investigated memorability, in this case of images of faces. With a similar method as Isola et. al (2011), Bainbridge et al. (2013) also found large consistency in how well or poorly images of faces were remembered by different participants. These robust findings suggest that memorability is an objective property of images. Attempts to explain and predict memorability through other image properties were unsuccessful. Since memorability is a highly consistent and stable property, but difficult to be reduced to other properties, it is thought to be an intrinsic property of an image. In a large-scale project called THINGS, researchers have created an open accessible dataset with 26,107 images (Hebart et al., 2019). Kramer et al. (2022) established memorability sores for these images. This database allows us to use memorability as a manipulated variable within our design. But how would memorability influence vocabulary learning?

Using a memorable image, as compared to a non-memorable image, might have a beneficial effect on multimedia vocabulary learning. Madan et al. (2010) showed that a factor that determines the memorability of a word (its imageability) influences the strength of the association with another word. Assuming that elements of associations are dependent (e.g. they are encoded and retrieved in a holistic fashion, see Horner & Burgess, 2003), it follows that a factor that influences the memorability of one element in the association would influence the strength of the association between all elements in the memory representation. By analogy, it could be hypothesized that a memorable picture would also strengthen the association between an L2 word and its L1 translation. Memorability could therefore be a property that makes retrieval of associated knowledge easier.

In the present study, the effect of images and their memorability on vocabulary learning is tested. By doing so, we try to contribute to research on the implications of using multimedia sources for vocabulary learning. Dutch first year psychology students were asked to learn Finnish words. Using a 2 x 2 experimental design, we will assign the participants to different encoding conditions where they are presented a Finnish word, the Dutch translation and a memorable picture, a non-memorable picture or no picture, in random order, depending on the assigned version. Participants are then asked to recall the Dutch translations of the words with only the Finnish word cue. They will be provided with feedback and the same image they were shown with encoding (or no image if there was no image provided in the encoding phase). After a few blocks, all Finnish words will be tested for a second time. This time, no feedback will be given afterwards. Performance on the first and second test will be used to establish whether memorability of images and images in general have an effect on vocabulary learning. Existing literature does not allow us to make very strong predictions, since the existing research is not consistent in whether there is an effect of images, and whether this effect is positive or negative for vocabulary learning. Within the perspective of Dual Coding Theory, image presence would have a positive effect on vocabulary learning. Presenting an image next to a word that needs to be learned could lead to better retention than only presenting a word with its translation (Akbulut, 2007; Shakroni, 2009; Yeh & Wang, 2003). An image could activate more codes (both visual and verbal) than a word. An underlying assumption of this theory is that the number of codes that is activated during encoding and retrieval is predictive of memory recall. This would suggest that presenting an image next to a word that needs to be learned could lead to more deep processing compared to presenting only a translation and therefore enhance vocabulary learning. However, there are also several studies that found no effect or even negative effects for adding pictures in methods to study foreign language words. This suggests that it is also possible that adding images to the encoding stimuli can cause a redundancy effect, especially within our study in which we only use concrete nouns.

Regarding memorability of pictures, it is even harder to make predictions since little research has been conducted in which the practical application of memorability was tested. From a theoretical perspective, memorability of images could lead to better vocabulary retention. As mentioned, Madan et al. (2010) suggested that associative-memory can be enhanced even when one part of the association is easier to retrieve. Therefore, we could argue that memorability of images has a beneficial effect on the retrieval of foreign language words. This research may contribute the improvement of education and learning materials, and may contribute to more consistent theoretical support.

Method

Participants

In the current study, first-year Psychology students from the University of Groningen constituted the participants. The participants received so-called SONA points based on participation, which they needed to complete a course, as compensation. A total of 40 participants participated in the experiment, of which 33 identified as female and 17 as male.

Participants' ages ranged between 17 and 26 years (*M* = 20.2; *SD* = 2.1). The participants were all fluent in Dutch and they did not speak Finnish, Estonian, Danish, Hungarian, Icelandic, Norwegian or Swedish.

Materials and equipment

The experiment was designed using *OpenSesame* (Version 3.3.12; Mathôt, 2022; Mathôt et al., 2012), and could be executed online. Therefore, the participants were able to execute the experiment on their own computer at home. The participants in the current study used their own laptop with a Latin keyboard. Participants used their keyboard to give answers. The experiment was displayed on a screen with a resolution of 1366 x 768 pixels. In this study, Finnish nouns and their Dutch translation were displayed. Depending on the condition, these words were accompanied by images that represented the words. The images that were used were selected from the database of Hebart et al. (2019). These images represented 40 nouns, including the names of objects, animals and food. The selection of the words was based on different criteria, as shown in Appendix A Table 1 and 2. All words that were selected had a percentage known of 100, a concreteness score of 5 and a SUBTLEX-score ≥ 100 (Hebart et al., 2019). This respectively means that the words that were selected should have all been familiar to the participants in our sample, the meaning of the words was very concrete and the words were frequently used in subtitles of movies. The exact scores of the criteria for each word are displayed in Appendix A Table 1 and 2.

After making an initial selection, the remaining words were translated from English to Dutch. The words were also translated to Finnish with help of Google Translate. After using Google Translate, the translation was checked with help of Google Images to see if the translation corresponds to the Dutch meaning. After translation, several steps were taken to further select words. Firstly, words of which the Dutch and Finnish spelling was too similar, were deleted from the word pool. These deleted words were completely similar or differed with only one letter. Then, selection was based on the concreteness score of the Dutch word (Brysbaert et al., 2014). All words with a concreteness score of 4 or less were deleted from the pool, this means we only used very concrete words with a maximum of 5. Next, selection was based on word length of the Finnish word. All words with 8 or fewer letters were selected.

Finally, selection was based on memorability of the images that were found for the remaining words in the database of Hebart et al. (2019). This database included several images for each word, with memorability scores. We selected images with highest memorability. These scores were .8 or higher (Kramer et al., 2022). Low-memorability images were selected with memorability scores of 0.75 or lower. To make a final selection of words, we selected words that had highly memorable and low memorable image pairs. The words that were selected were represented by images that differed with .18 or more on their memorability scores (Hebart et al., 2019).

Three different types of screens are used in this experiment, a screen for the encoding phase, a screen for the test phase and a screen for the feedback phase. On these screens was a short description, the Finnish word and potentially the Dutch word, shown in Figures 1, 2, and 3.

Figure 1

Screen in Encoding phase in the condition with Image

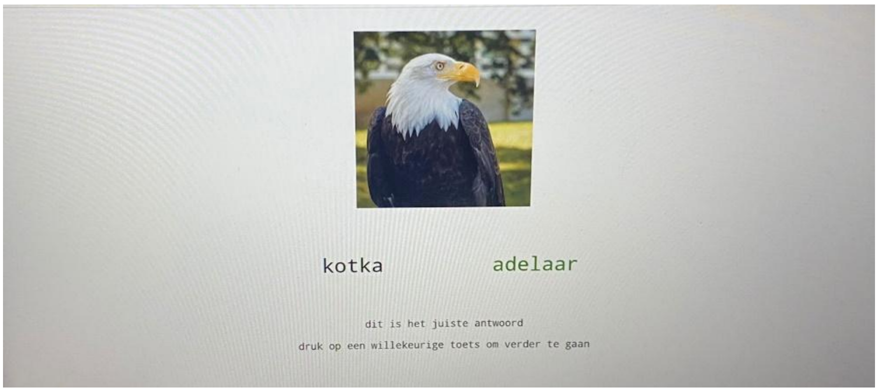
Note. The lay-out of the screen in the encoding phase was nearly identical to the condition without image, except that no image was shown in the condition without an image.

Figure 2

Screen in Testphase for the Condition with and without Image

Figure 3

Screen in Feedback Phase in the condition with Image



Note. The lay-out of the screen in the feedback phase was nearly identical to the condition without image, except that no image was shown in the condition without an image.

Design

The current study was based on a 2 x 2 design, with two manipulated variables. These independent variables were the presence of an image during the encoding phase, and the memorability of these images. Image presence was a between-subject variable, memorability was a within-subject variable. The dependent variables were the average proportion of correct responses given during the second and third test phases. The first test phase followed directly after the encoding of a specific word, in which the participant is asked to type the translation of the word that they had just learned. This test phase was mainly implemented to make sure that participants were paying attention during the experiment. The second test phase follows directly after the encoding of one single block of 8 words, in which the translations that were presented in that specific block were tested in random order. The third and therefore last test phase follows after the encoding of all learning blocks, in which all Finnish words are presented in random order. The second and third test phase scores were used as a measure of short-term and long-term learning effects.

The aim of this study was to test whether images in general have an effect on vocabulary learning, as well as whether memorability of images influences vocabulary learning. In order to answer these questions, four different versions of the experiment were designed. The first two versions included memorable and non-memorable images in mixed order, and the other two versions did not include images.

Other possible variables that might have an influence on learning were controlled for as well. For example, factors such as word length, word frequency and concreteness were kept the same across different blocks. By counterbalancing these factors, their possible influences on learning are controlled for. The difficulty of each block was the same within each version and across all versions. In the versions in which an image was presented during the encoding phase, there was an even balance between memorable and non-memorable images that were shown. In all different versions, the order of the blocks are the same. The order of words presented in each block however, is random. After collecting all data, Repeated Measures ANOVA-analyses, dependent t-tests and regression analyses are used to compare mean averages.

Procedure

First-year psychology students at the Rijksuniversiteit Groningen were asked to participate in the experiment. The students were able to sign up for one of the four versions of the experiment through SONA which automatically made them non eligible for participation in one of the other three versions. Participants who signed up for one of the first two versions were shown memorable and non-memorable images in mixed order, and participants who signed up for the other two versions were shown no images. The participants were first informed about the procedure, that they would learn the Dutch translation of Finnish vocabulary, and that the experiment would take around thirty minutes. First, the participants were asked to answer some selection questions about gender, age, and spoken languages, followed by giving consent to take part in the experiment and collect data. When consent was given, participants were informed that they would be asked to learn Finnish vocabulary. This would happen in five different blocks, which consisted of eight words each. When a participant had walked through the encoding and testing for all five blocks, a final test followed in which all learned vocabulary was presented in random order. The blocks were presented in fixed order, but the word order within these blocks would be randomly arranged so that the order would be counterbalanced.

Starting with the first part of the encoding phase, the first Finnish words, the Dutch translation and, depending on the version, the image was or was not shown. By pressing the spacebar, the Dutch word and image disappeared. Then, the participant was asked to type in the correct corresponding Dutch word. This was repeated for all words and at the end of the first block, all Finnish words, without an image, were tested again, followed by feedback. The same process of encoding and testing followed for the subsequent four blocks. After all blocks, another testing phase followed, where all 40 words were tested. The experiment concluded with the debriefing.

Results

In this study the influence of memorability from images on vocabulary learning was researched. First, it was expected that the presence of images would lead to better performance when learning vocabulary compared to learning vocabulary without images. Second, it was expected that the difference in memorability of an image would influence the learning process of vocabulary. Our specific hypothesis was that memorable images would contribute to better vocabulary learning than non-memorable images. The measure used to determine the extent of how vocabulary learning was improved, was based on the accuracy of the second and third retention tasks in the experiment. The accuracy was determined by checking whether the participant's answer corresponded completely or almost completely to the correct translation. The words counted correctly were identical to the correct answer, differed by one letter or corresponded phonetically to the correct answer. The retention tasks provided insight into short- and long-term performance and were thus a measure of learning performance. The hypotheses were tested using Repeated Measures Analysis of Variance (ANOVA). In addition, ANOVA, pairwise t-tests and regression analysis were used to test whether other variables within this experiment had affected performance and response time, such as difficulty, influence of retention time, response time in encoding phase 1 and the influence of word length.

Accuracy

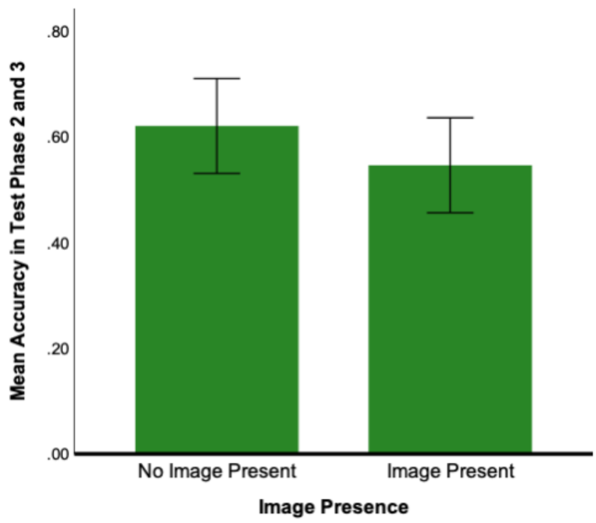
Presence of Image. To test if the presence of an image had an effect on the accuracy, a repeated measures ANOVA was used. First, it was checked whether the assumptions underlying the ANOVA were met, with the mean accuracy as the dependent variable. Independence of observations was guaranteed, because the observations were from the averages per participant. These averages were by definition independent of each other, so the measurements met the assumptions of independence. There were also no indications that the assumptions of normality had not been met. No strong deviations for normality could be found, as shown in Appendix B Figures 1 and 2 (Shapiro-Wilk[40] = .97, p = .256). The mean accuracy therefore was normally distributed. Finally, the assumption of sphericity was automatically met, because in this ANOVA there was only one within-factor with just two levels. This means that there was no correction for sphericity needed and there were no indications of violations of ANOVA assumptions.

Participants seemed to have performed better on the retention test when they had not seen images (*M* = .62, *SD* = 0.49) than when they had seen images (*M* = .55, *SD* = 0.50), see Figure 4. However, this difference was not significant ($F[1, 38] = 1.41$; $p = .24$, partial $\eta^2 = .04$). This result indicates that there was insufficient support for the hypothesis that images have an effect on accuracy.

Memorability. To measure the effect of memorability on the accuracy, we looked into the group of participants who had seen images and looked at the difference in accuracy based on memorability. For this analysis, we used an ANOVA-analysis and a pairwise t-test. The assumptions of the latter test were met. The measures were independent and paired within a person. The condition that the results were normally distributed also seemed to be met, as shown in Appendix B, Figures 3 and 4. A normality test also showed no reason to assume that the results were not normally distributed (Shapiro-Wilk[20] = .969, $p = 0.726$). Although the

Figure 4

Bar graph of Proportion of Correct Answers as a Function of the Absence or Presence of an Image



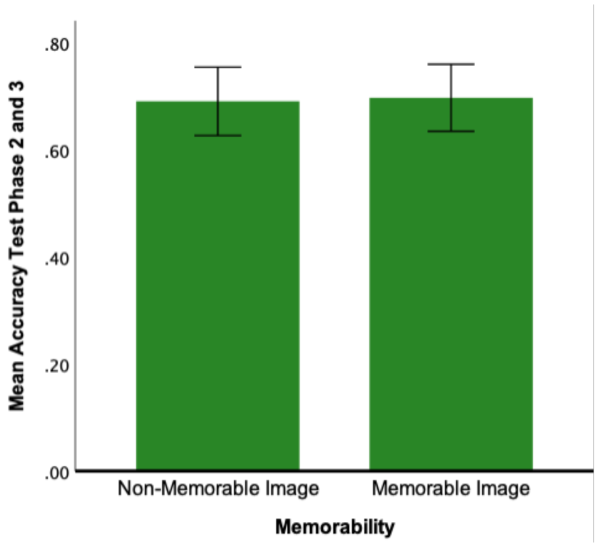
Note. The error-bars were based on a 95% confidence interval results were not normally distributed (Shapiro-Wilk[20] = .969, p = 0.726). Although the sample was small, based on these results the assumption of a normal distribution did not seem to be violated and thus it was justified to use a t-test to conduct the analysis.

Mean accuracy for both the short-term and long-term retention tests seemed to differ only a little for memorable images (M = .59, SD = .49) and non-memorable images (M = .58, SD = .49), as shown in Figure 5. In addition, there was no interaction effect between 'Memorability' and 'Presence of Image' (F[1,39] = 0.83, p = .367, partial η^2 = .02). A pairwise t-test only analyzing scores from participants who had seen images confirmed this (t[19] = -.40, p = .694, Cohen's d = -.09). Thus, based on this data there was no evidence that memorability affected performance on retention tests.

Retention time. Another factor that could affect vocabulary learning and recall was retention time. Retention time is the time between learning and having to retrieve the

Figure 5

Bar graph of Proportion Correct Answers as Function of Memorable or Non-Memorable Image



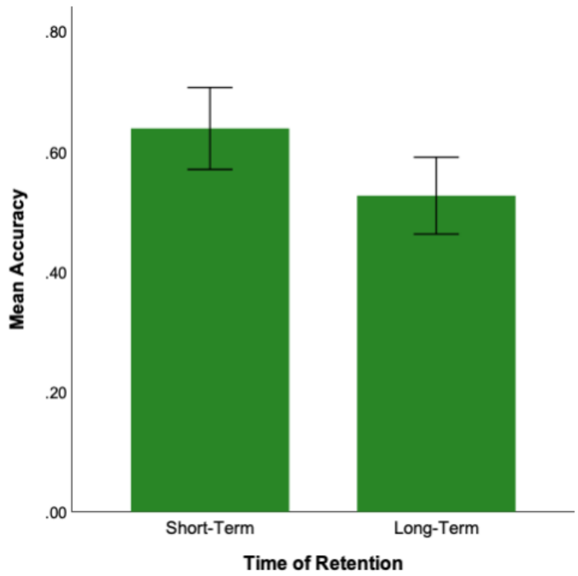
Note. The error-bars were based on a 95% confidence interval. Only scores from participants in the condition 'Image present' were included with n = 20.

retention time. Retention time is the time between learning and having to retrieve the vocabulary. Here, a distinction was made between short-term and long-term. The analysis for retention time was based on the same ANOVA used to test the effects of the presence of an image on accuracy. The corresponding assumptions were therefore met. Figure 6 shows that performance to retrieve the word was worse on average with longer retention time (M = .53, SD = .20) than with a shorter retention time (M = .64, SD = .22). Retention time was indeed found to have a significant effect on performance during retention tests ($F[1, 38] = 36.20$; $p <$ .001, partial η^2= .49). Moreover, this effect size indicated that the effect was very strong and thus could largely explain the variance in accuracy.

Word length. Finally, it was also researched whether word length of the Finnish words affected accuracy on the retention tests. For this, a similar ANOVA analysis was used as the one that tested the effects of presence of an image on accuracy. The corresponding assumptions were therefore met. There seemed to be a difference in how well a word was retrieved as a

Figure 6

Bar graph of Correct Answers as a Function of Retention time (Short-term and Long-term)



Note. The error-bars were based on a 95% confidence interval.

function of word length, as shown in Figure 7. The 3-letter words (*M* = .80, *SD* = .40), 8-letter words (*M* = .73, *SD* = .45) and 9-letter words (*M* = .69, *SD* = .46) seemed, in fact, to have higher accuracy at short-term and long-term than the 4-letter words (*M* = .43, *SD* = .50), the 5-letter words (*M* = .53, *SD* = .50) and the 6-letter words (*M* = .45, *SD* = .50). The effect of word length on proper word retrieval was significant ($F[5,195] = 45.30$; $p < .001$, partial $\eta^2 = .68$). However, these results should be interpreted with some caution as the number of words per category was very small, making the results not very reliable.

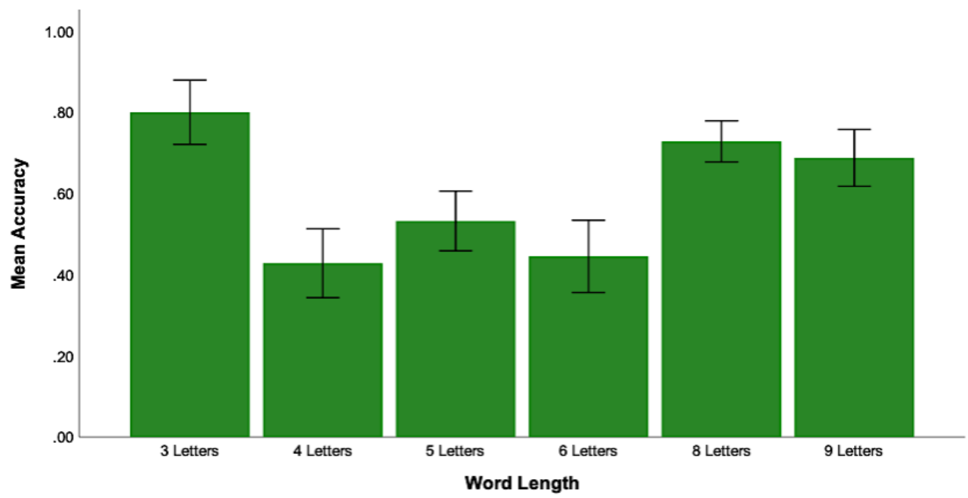
Response time

The next analysis was used to look at the time it took participants to study the items during the encoding phase, and whether this time depended on the presence of an image, and if the image was memorable or not.

Image presence. To test if response time was affected by image presence, we again used repeated measures ANOVA. The measures were again independent. The assumption of normality seemed to be violated, however. The distribution of the data seemed to be right

Figure 7

Bar graph of Correct Answers and Word Length

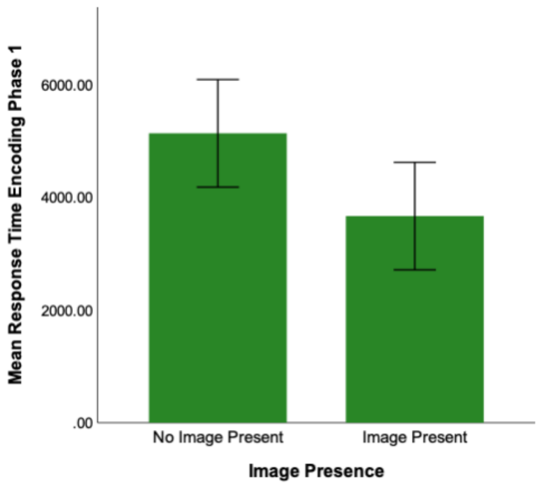


Note. The error-bars were based on a 95% confidence interval skewed, as shown in Appendix B figure 5. A Q-Q plot showed a deviation from a normal distribution, as shown in Appendix B Figure 6. Consequently, based on the Shapiro-Wilk test, we could not assume that the data was normally distributed (Shapiro-Wilk[40] = .91, p = .005). Nevertheless, our sample size (N = 40) was large enough to make sure that the analysis is sufficiently robust for violations of normality. Finally, the assumption of sphericity was met, since this ANOVA also included only one within-factor with two levels. Therefore, correction for sphericity was not necessary and there were no other indications that an ANOVA-analysis would be unsuitable. The results should nonetheless be interpreted with caution, since the assumption of normality was violated.

Our analysis showed that participants who did not see an image during encoding, took more time for encoding (M = 5131.48, SD = 5871.14) than the participants who did see an image (M = 3663.76, SD = 3096.42), as shown in Figure 8. This effect turned out to be significant (F[1, 38] = 4.85, p = .03, partial η^2 = .11) with an average effect size.

Figure 8

Bar graph of Effect of Image Presence on Mean Response Time During Encoding Phase 1.



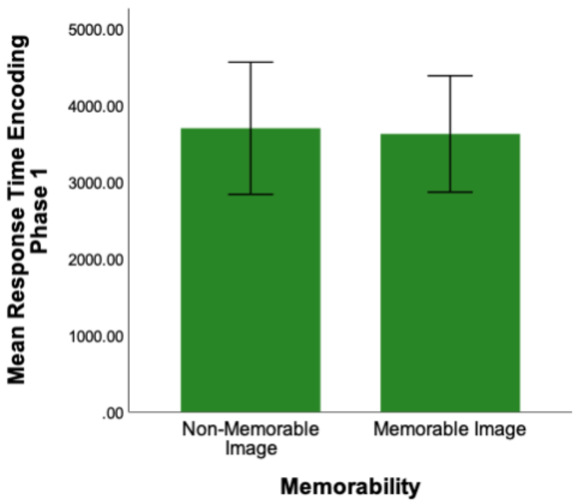
Note. The error-bars were based on a 95% confidence interval. Only scores of participants in the condition where an image was present were included, n = 20.

Memorability. To analyze the influence of memorability on response time, a paired t-test was performed and the ANOVA-analysis that was mentioned earlier was used. The assumption of independent measures and the assumption of dependent measures within the same person were met. The assumption that the results were normally distributed also seemed to be met, as shown in Appendix B, Figure 7 and 8. A normality test did not indicate that the results were not normally distributed (Shapiro-Wilk[20] = .912, p = 0.071). However, the sample size was relatively small (n = 20). Therefore, the results, again, need to be interpreted with caution.

Based on the results of ANOVA, the mean response time was barely different for memorable images (M = 4307.44, SD = 4231.49), compared to non-memorable images (M = 4487.80, SD = 5216.77), as shown in Figure 9. The interaction-effect between memorability and image presence was also not significant (F[1.38] = .002, p =.968, partial η^2 = .00). A paired t-test based on scores of participants in the image present condition further confirmed this finding (t[19] = .36, p = .726, Cohen's d = .08). Based on these results, there seems to be no

Figure 9

Bar graph of Effect of Image Memorability on Mean Response Time During Encoding Phase 1.



Note. The error-bars were based on a 95% confidence interval. Only scores of participants in the condition where an image was present were included, n = 20.

finding (t[19] = .36, p = .726, Cohen's d = .08). Based on these results, there seems to be no indication that memorability influences the time that was taken to encode the words.

Relationship Between Response Time and Accuracy

Finally, we investigated whether there was a relationship between response time and accuracy. This relationship was determined with Pearson's correlation coefficient. This correlation can be used if there seems to be a linear relationship between two continuous variables. As shown in Figure 10 and 11, this seemed to be the case for our data. It is important to note that response time scores were not normally distributed and included a few scores that were particularly high. Since it was unknown why these scores were particularly high, they were not excluded from the dataset. For this reason, the correlation coefficient should be interpreted with caution.

There seemed to be a strong correlation between response time and accuracy. Especially for short-term retention, the correlation between response time and accuracy was strong (r[38]= .68, $p < 0.001$). This result shows that participants who took more time to look at the encoding screen, had higher scores on the short-term retention test, see also Figure 10. Response time and accuracy on long-term retention were also correlated, however, this correlation was moderate (r[38] = .51, p = .001), as shown in Figure 11. This suggests that participants also scored better on the long-term retention tests when they took more time to encode the item. As mentioned before, these results should be interpreted with caution.

Discussion

In the current study, we looked at foreign language word retention as a function of image memorability and image presence. Firstly, we expected that images with a high memorability score would lead to better retention than images with a low memorability score. Our data does not support our hypothesis, since no significant effects were found. This means that we found no effect for picture memorability on retention of foreign language vocabulary.

Figure 10

Scatterplot of Relationship between Response Time and Short-Term Accuracy

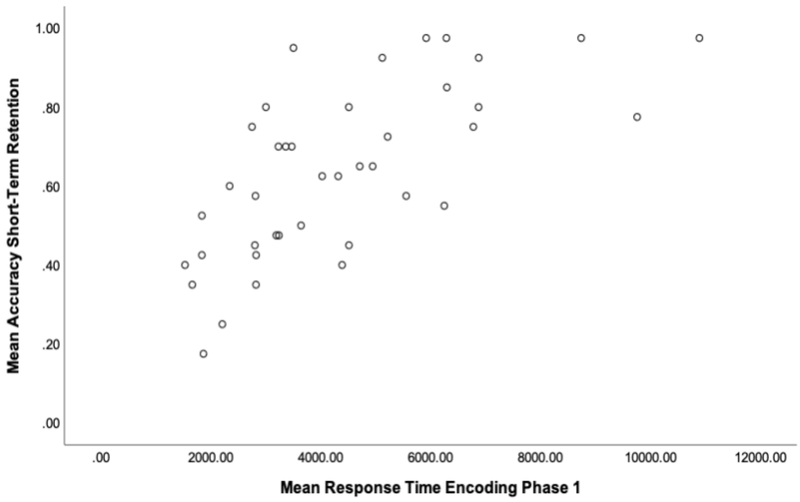
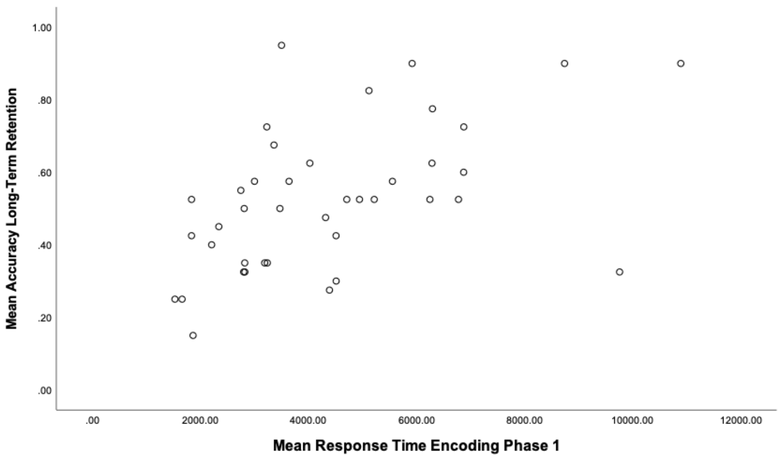


Figure 11

Scatterplot of Relationship between Response Time and Long-Term Accuracy



Our second hypothesis stated that the presence of images during encoding would have beneficial effects on retention compared to the absence of images during encoding. This hypothesis was not supported by our data either, since again, no significant effects were found. Hence, our data suggests that showing a picture during encoding has no effect on later retention and therefore will not improve vocabulary learning.

Explanations, Limitations and Future Directions for Memorability

Memorability is a relatively new concept within cognitive science and especially practical applications of memorability are sparse. To investigate whether memorability could have practical implications, we investigated whether memorability might influence vocabulary learning. In the current study, no effect was found for memorability on vocabulary learning. There could be several reasons why we did not find an effect for memorability. One reason could be that our experiment had a lack of power because the manipulation on memorability was not strong enough. Within our word selection process, we aimed at selecting words of which the high- and low-memorable images had the largest difference scores on memorability. Still, the difference between memorable and non-memorable images was not very large, with a mean difference in memorability score of 0.28 . This might explain why we found no difference between retention scores for memorable and non-memorable images. In other thesis projects, memorability was maximally different and showed significant effect on recall. These theses included studies on face-name recall (Hulsewiesche & Nieuwenstein, 2022), and brand-name recall recall (Kostova & Nieuwenstein, 2021) However, the latter study could not be replicated by Koiter and Nieuwenstein (2022) and another study on brand-name recall was not able to find significant effects of memorability either (Wit & Nieuwenstein, 2019). These findings suggest that memorability of the image or product has no influence on brand-name recall in advertisement settings. In other situations such as face-name recall, memorability does have a positive effect. This does not necessarily mean that the setting itself is responsible for the effect of memorability. Rather, the biggest difference between the studies on brand-name recall and face-name recall were the way in which stimuli were presented to the participants. In the advertisement studies, brand names were presented as words next to the images that represented the product for which the advertisement was made. In the face-name recall study, participants were presented faces on a computer screen and received the corresponding name of the person through audio. These findings, in addition to our own findings, suggest that there might be a boundary condition on the effect of memorability for which memorability of images only has an effect if two stimuli of different sensory modalities are presented. This could be the second reason why we were unable to find an effect for memorability. Since there has not been much research on the application of memorability in more practical settings, further research is necessary to test whether memorability has practical implications at all, and if so, what the boundary conditions are. Therefore, further research should include two distinct stimuli, preferably auditive and visual stimuli simultaneously, for which the visual stimuli has maximally different levels in memorability. This should then be compared to the presentation of only visual stimuli only. This way, boundary conditions for sensory modalities can be investigated.

Explanations, Limitations and Future Directions for Image Presence

A broader question is whether multimedia learning in the form of presenting images next to study material is beneficial for later recall. Existing literature consists of mixed results, as mentioned earlier. In our study, we were not able to find significant effects for image presence on vocabulary learning. This research supports other findings for which no effects were found for multimedia benefits on vocabulary learning (Boers et al. 2009; Cohen & Johnson, 2011; Lotto and de Groot, 1998; Chen,1990; Carpenter & Olson, 2011). It should be mentioned, that although our results were not significant, we did in fact find a substantial difference between both conditions (image presence versus image absence), with participants scoring lower when images were present than when images were absent. However, the variability between participants' scores was large. A replication of this study should therefore include a larger sample size, to increase power which in turn would increase the ability to find a significant effect, if there is one. This is important, since multimedia learning is gaining popularity. If presenting images really has a potential negative effect on learning, as was found by Sweller et. al, (1998) (see also, Acha, 2009; Harp & Mayer, 1998), this should be addressed and be acted upon.

Another interesting effect that was found was that participants in the condition in which no images were present during encoding, took a significantly longer time to look at the encoding screen, compared to participants that were shown a picture during encoding. These results did indirectly predict why participants in the image present condition scored lower on average accuracy, because encoding time significantly and strongly predicted of accuracy in retention. This means response time might function as a mediator variable for the relation between that image presence and accuracy in retention. This is supported by findings of Olson & Carpenter (2011), who also found strong correlations between response time and accuracy. Based on their findings, once could also suggest that learners are biased in thinking how well they encode words when images are present. With regard to deep processing, it important that not only rich information networks are formed (by activating multiple codes), it is also important that the learner puts in effort to organize new information and integrate this with existing knowledge. Perhaps learners do not engage in active learning when they are shown an image, as they might suspect that learning is easier. When learners only activate, but not efficiently organize the new information, encoding might not be as successful. Since organizing knowledge takes time, it seems plausible that the shorter encoding time reflects a lack of deep processing.

Another reason why there are mixed results on image presence and vocabulary learning might arise from the words that are included in the experiment. One specific limitation in our thesis is the choice to only select words with high concreteness scores (Hebart et al., 2019, Brysbaert et al., 2014). The reason to choose high concrete words was based on the fact that the words that were selected had to be paired with images that represented the meaning of that word. For more abstract words, there could be more debate about whether the images are representative. For example, an image that has to represent the word 'breakfast' can be interpreted as breakfast for some people, whereas others would not consider the image representative for breakfast. To avoid this issue, only concrete words were selected. Supporting the Dual-Coding Theory of Clark and Paivio (1991), Yui et al. (2017) found that concrete words are better recalled than abstract words. Another study found similar effects where concrete words lead to better vocabulary learning (Hiebert et al, 2019). Therefore, our results could have been biased because the experiment might have been too easy in advance. However, our results suggest that the difficulty level of our experiment appropriate, with an average accuracy between 50% and 60%. This level of difficulty allowed us to analyze our results without bias since no floor or ceiling effects could interfere. From this point of view, the use of highly concrete words was not problematic.

There is however, another reason why the use of only highly concrete words might have been a limitation in our study: Images of concrete words may cause a redundancy effect, as mentioned in the introduction of this article (Clark & Mayer, 2003; Kalyuga & Sweller, 2014). According to Yui et al. (2017), concrete words activate both verbal and imagery codes, whereas abstract words only activate verbal codes. In our study, we present images next to concrete words. These images might be not necessary to activate the visual code corresponding to the word, since the word itself already activates this code. Therefore, the image becomes redundant and might unnecessarily increase the cognitive load, which would have a negative influence on the encoding later retention of the word. Perhaps, the addition of images distracts learners from forming the association between the L2 word and the L1 translation. By activating the same number of codes, the association between L1 word and the image is very easy to understand, which could lead learners to have quick response styles. Their prior knowledge of the L1 word and the image is congruent with the current information, and therefore learners might have shorter encoding times. However, when the image distracts attention from the L2 word, this receives less attention or no attention at all, which could explain why shorter retention time in the picture present condition predicts poorer accuracy. For this reason, future research should include more abstract words. If concreteness of words seems to be boundary condition for the effect of image presence, it could explain why research outcomes are mixed.

Conclusion

In the last decade, multimedia learning has gained a more popularity in educational settings and more attention in research. To improve vocabulary learning, using multimedia can be considered. This bachelor thesis aimed to answer the question whether multimedia can have beneficial effects on vocabulary learning. Furthermore, memorability has become a subject of interest in the last years of research. Fundamental research on memorability has been done, but more practical applications of memorability are sparse. Therefore, we also looked at the possible benefits of using memorable pictures for vocabulary learning. These beneficial effects were not found. Instead, there was no effect for memorability, and the effect of image presence, although not significant, would more likely have a negative effect on vocabulary learning than a beneficial effect. Therefore, based on our results, we would not advise to use multimedia in the form of images to enhance vocabulary learning.

References

Akbulut, Y. (2007). Effects of multimedia annotations on incidental vocabulary learning and reading comprehension of advanced learners of English as a foreign language. Instructional Science, 35(6), 499–517. https://doi.org /10.1007/s11251-007-9016-7

Acha, J. (2009). The effectiveness of multimedia programs in children's vocabulary learning. *British Journal of Educational Technology*, *40*(1), 23–31. https://doi.org/10.1111/j.1467-8535.2007.00800.x

Ainsworth, S., & Loizou, A. T. (2003). The effects of self-explaining when learning with text or diagrams. Cognitive Science, 27(4), 669–681. https://doi.org/10.1207/s15516709cog2704_5

Bainbridge, W. A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face photographs. *Journal of Experimental Psychology: General*, *142*(4), 1323–1334. https://doi.org/10.1037/a0033872

Boers, F., Píriz, A. M. P., Stengers, H., & Eyckmans, J. (2009). Does pictorial elucidation foster recollection of idioms? *Language Teaching Research*, *13*(4), 367–382. https://doi.org/10.1177/1362168809341505

Brysbaert, M., Warriner, A. H., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. Journal of Educational Psychology, 98(1), 182–197. https://doi.org/10.1037/0022-0663.98.1.182

Carpenter, S. K., & Olson, K. M. (2012). Are pictures good for learning new vocabulary in a foreign language? Only if you think they are not. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(1), 92–101. https://doi.org/10.1037/a0024828

Carpenter, S. K., Pashler, H., & Cepeda, N. J. (2009). Using tests to enhance 8th grade students' retention of U.S. history facts. *Applied Cognitive Psychology, 23*(6), 760–771. https://doi.org/10.1002/acp.1507

Chen H. C. (1990). Lexical processing in a non-native language: effects of language proficiency and learning strategy. *Memory & cognition*, *18*(3), 279–288. https://doi.org/10.3758/bf03213881

Clark R. C., Mayer R. E. (2003). *e-learning and the science of instruction*. San Francisco: Jossey-Bass.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. Educational Psychology Review, 3(3), 149–210. https://doi.org/10.1007/ BF01320076

Cohen, M. T., & Johnson, H. L. (2011). Improving the acquisition of novel vocabulary through the use of imagery interventions. Early Childhood Education Journal, 38(5), 357–366. https://doi.org/10.1007/s10643-010 -0408-y

Durso, F. T., & Johnson, M. K. (1980). The effects of orienting tasks on recognition, recall, and modality confusion of pictures and words. *Journal of Verbal Learning & Verbal Behavior, 19*(4), 416–429. https://doi.org/10.1016/S0022-5371(80)90294-7

Farley, A., Pahom, O., & Ramonda, K. (2014). Is a picture worth a thou- sand words? Using images to create a concreteness effect for abstract words: Evidence from beginning l2 learners of Spanish. Hispania, 97(4), 634–650. https://doi.org/10.1353/hpn.2014.0106

Farley, A., Ramonda, K., & Liu, X. (2012). The concreteness effect and the bilingual lexicon: The impact of visual stimuli attachment on meaning recall of abstract L2 words. Language Teaching Research, 16(4), 449–466. https://doi.org/10.1177/1362168812436910

Fiore, S. M., Cuevas, H. M., & Oser, R. L. (2003). A picture is worth a thousand connections: The facilitative effects of diagrams on mental model development and task performance. Computers in Human Behavior, 19(2), 185–199. https://doi.org/10.1016/S0747-5632(02)00054-7

Hald, L. A., de Nooijer, J., van Gog, T., & Bekkering, H. (2016). Optimizing word learning via links to perceptual and motoric experience. Educational Psychology Review, 28(3), 495–522. https://doi.org/10.1007/ s10648-015-9334-2

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. Journal of Educational Psychology, 90(3), 414–434. https://doi.org/10.1037/0022-0663 .90.3.414

Hebart, M. N., Dickter, A. H., Kidder, A., Kwok, W. Y., Corriveau, A., Van Wicklin, C., & Baker, C. I. (2019). THINGS: A database of 1,854 object concepts and more than 26,000 naturalistic object images. *PLOS ONE*, *14*(10), e0223792. https://doi.org/10.1371/journal.pone.0223792

Horner, A. J., & Burgess, N. (2013). The associative structure of memory for multi-element events. *Journal of Experimental Psychology: General, 142*(4), 1370–1383. https://doi.org/10.1037/a0033626

Hulsewiesche, J. & Nieuwenstein, M. R. (2022) *The Effect of Face Memorability on Remembering Face-Name Associations: An Ethnically Inclusive Design*. [Unpublished bachelor's thesis]. University of Groningen.

Isola, P., Xiao, J., Torralba, A., & Oliva, A. (2011). What makes an image memorable? *Computer Vision and Pattern Recognition*. https://doi.org/10.1109/cvpr.2011.5995721

Kalyuga, S., & Sweller, J. (2014). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), The Cambridge handbook of multimedia learning (2nd ed., pp. 247–262). Cambridge University Press. https://doi .org/10.1017/CBO9781139547369.013

Koiter, B. & Nieuwenstein, M. R., (2022). *The Relationship Between Brand Memory, Image*

Memorability, and Fixation Durations While Viewing Picture Advertisements. [Unpublished master's thesis]. University of Groningen.

Kostova, L. & Nieuwenstein, M. R. (2021). *Can memorability influence brand memory in picture advertisements? A comparison of self-paced and short viewing times.* [Unpublished bachelor's thesis]. University of Groningen.

Kramer M., Hebart M. N., Baker C. I., & Bainbridge W. A. (2022) *The features underlying memorability of Objects.* BioRXiv. https://doi.org/10.1101/2022.04.29.490104

Lotto, L., & de Groot, A. M. B. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning*, *48*(1), 31–69. https://doi.org/10.1111/1467-9922.00032

Madan, C. R., Glaholt, M. G., & Caplan, J. B. (2010). The influence of item properties on association-memory. *Journal of Memory and Language*, *63*(1), 46–63. https://doi.org/10.1016/j.jml.2010.03.001

Mathôt, S., & March, J. (2022). Conducting linguistic experiments online with OpenSesame and OSWeb. *Language Learning*. https://doi.org/10.1111/lang.12509

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314-324. https://doi.org/10.3758/s13428-011-0168-7

Mayer, R. E. (2002). Multimedia learning. Psychology of Learning and Motivation. 41, 85-139. https://doi.org/10.1016/S0079-7421(02)80005-6

Quizlet. (2023). Quizlet. Retrieved January 22, 2023 from https://quizlet.com

Reisberg, D. (2018). *Cognition: Seventh International Student Edition*. W.W. Norton & Company. 215-219

Shahrokni, S. A. (2009). Second language incidental vocabulary learning: The effect of online textual, pictorial, and textual pictorial glosses. The Electronic Journal for English as a

Second Language, 13(3), 1–17.

van Rijn, D., van Maanen, L., & van Woudenberg, M. (2009). Passing the test: Improving Learning Gains by Balancing Spacing and Testing Effects. In A. Hoses, D. Peebles, & R. Cooper (Eds.), *Proceedings of the 9th International Conference on Cognitive Modeling, 187,* 108-114.

Sweller, J. (2006). The worked example effect and human cognition. *Learning and Instruction, 16*(2), 165–169. https://doi.org/10.1016/j.learninstruc.2006.02.005

Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251–295.

van den Broek, G. S. E., van Gog, T., Jansen, E., Pleijsant, M., & Kester, L. (2021). Multimedia effects during retrieval practice: Images that reveal the answer reduce vocabulary learning. *Journal of Educational Psychology*, *113*(8), 1587–1608.

Wilschut, T., Sense, F., Van Der Velde, M., Fountas, Z., Maaß, S., & Van Rijn, H. (2021). Benefits of Adaptive Learning Transfer From Typing-Based Learning to Speech-Based Learning. *Frontiers in artificial intelligence*, *4*. https://doi.org/10.3389/frai.2021.780131

Wit, L. & Nieuwenstein, M.. R. (2020) *Does image memorability enhance brand memory in advertisements?* [Unpublished bachelor's thesis]. University of Groningen.

WRTS. (2023). WRTS. Retrieved January 22, 2023 from https://wrts.nl

Yeh, Y., & Wang, C.-W. (2013). Effects of multimedia vocabulary annotations and learning styles on vocabulary learning. CALICO Journal, 21(1), 131–144. https://doi.org/10.1558/cj.v21i1.131-144

Yui L, Ng R, Perera-W.A. H. 2017. Concrete vs abstract words – what do you recall better? A study on dual coding theory. PeerJ Preprints 5:e2719v1

https://doi.org/10.7287/peerj.preprints.2719v1

Appendix A

Table 1

Used Words, Images and Subsequent Scores

Image Mem	Hit Rate Mem	Image Non-Mem	Hit Rate Non-Mem	Difference in Memorability
pillow_07s.jpg	0.89	pillow_02s.jpg	0.51	0.38
tomato_12s.jpg	0.92	tomato_10n.jpg	0.56	0.36
bat1_03s.jpg	0.92	bat1_15s.jpg	0.58	0.35
mattress_03s.jpg	0.92	mattress_07s.jpg	0.58	0.34
pig_06s.jpg	0.98	pig_04s.jpg	0.65	0.33
parrot_28s.jpg	0.93	parrot_11s.jpg	0.60	0.33
thorn_17s.jpg	0.88	thorn_14s.jpg	0.55	0.32
sled_01b.jpg	0.93	sled_06s.jpg	0.62	0.31
carrot_01b.jpg	0.95	carrot_06n.jpg	0.64	0.31
doll_03s.jpg	0.95	doll_01b.jpg	0.65	0.30
donkey_03s.jpg	0.88	donkey_06s.jpg	0.58	0.30
handkerchief_04s.jpg	0.92	handkerchief_14s.jpg	0.63	0.30
toad_07s.jpg	0.92	toad_13s.jpg	0.63	0.30
refrigerator_07s.jpg	0.93	refrigerator_12s.jpg	0.63	0.29
cage_04s.jpg	0.89	cage_01b.jpg	0.60	0.29

Image Mem	Hit Rate Mem	Image Non-Mem	Hit Rate Non-Mem	Difference in Memorability
apple_01b.jpg	0.95	apple_07s.jpg	0.66	0.29
tree_04s.jpg	0.84	tree_03s.jpg	0.55	0.29
ladder_04s.jpg	0.85	ladder_14s.jpg	0.56	0.29
eagle_13s.jpg	0.90	eagle_04s.jpg	0.62	0.28
bed_22s.jpg	0.91	bed_04s.jpg	0.63	0.28
basket_06s.jpg	0.88	basket_07s.jpg	0.60	0.28
grape_17s.jpg	0.90	grape_06s.jpg	0.63	0.28
cup_02s.jpg	0.85	cup_13s.jpg	0.58	0.28
ball_08s.jpg	0.93	ball_10s.jpg	0.66	0.27
sand_01b.jpg	0.90	sand_05s.jpg	0.63	0.27
clock_02n.jpg	0.80	clock_06n.jpg	0.53	0.27
tractor_11s.jpg	0.85	tractor_02n.jpg	0.60	0.25
finger_11s.jpg	0.95	finger_09s.jpg	0.70	0.25
leaf_11s.jpg	0.86	leaf_04s.jpg	0.62	0.24
snake_12s.jpg	0.92	snake_11s.jpg	0.68	0.25
jeans_05s.jpg	0.95	jeans_16s.jpg	0.71	0.24
penguin_17s.jpg	0.90	penguin_11s.jpg	0.66	0.24
lemon_01b.jpg	0.86	lemon_05s.jpg	0.63	0.24

Image Mem	Hit Rate Mem	Image Non-Mem	Hit Rate Non-Mem	Difference in Memorability
bird_20s.jpg	0.85	bird_04s.jpg	0.63	0.23
sponge_13s.jpg	0.93	sponge_10s.jpg	0.70	0.23
leopard_03s.jpg	0.90	leopard_08s.jpg	0.68	0.22
fish_08s.jpg	0,93	fish_09s.jpg	0.71	0.22
baby_19s.jpg	0.95	baby_06s.jpg	0.73	0.22
elephant_09n.jpg	0.93	elephant_10n.jpg	0.72	0.21
vase_19n.jpg	0.84	vase_06s.jpg	0.63	0.20
Mean	0.90	-	0.62	0.28

Table 2

Used Words and Subsequent Scores

Dutch Word	Finnish Word	Word Length Finnish Word	SUBTLEX	Concreteness Score Dutch Word
Kussen	Tyyny	5	581	4.47
Tomaat	Tomaatti	8	301	4.87
Vleermuis	Bat	3	1052	4.73
Matras	Patja	5	337	4.93
Varken	Sika	4	1996	4.80
Papegaai	Papukaija	9	167	4.87
Doorn	Piikki	6	260	4.80
Slee	Kelkka	6	149	4.93
Wortel	Porkkana	8	195	4.80

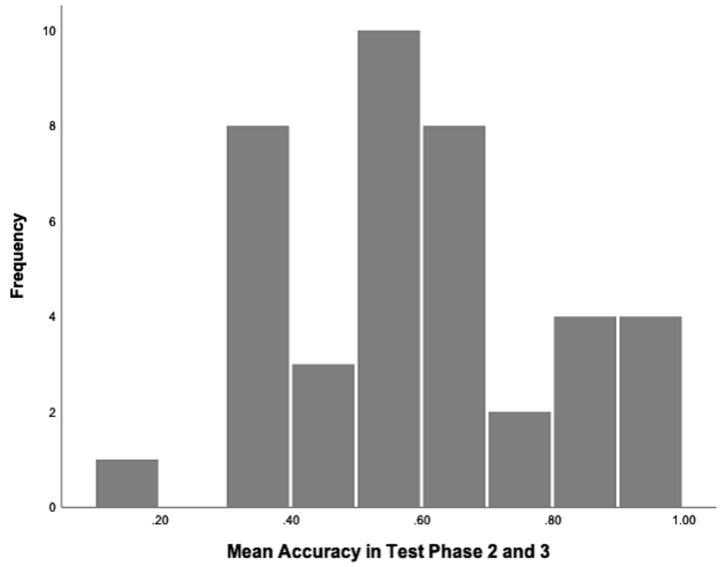
Dutch Word	Finnish Word	Word Length Finnish Word	SUBTLEX	Concreteness Score Dutch Word
Pop	Nukke	5	1263	4.27
Ezel	Aasi	4	273	5.00
Zakdoek	Nenällina	9	214	4.87
Pad	Rupikonna	9	290	4.53
Koelkast	Jääkaappi	9	427	4.87
Kooi	Häkki	5	1034	4.53
Appel	Omena	5	1207	4.67
Boom	Puu	3	3315	4.73
Ladder	Tikapuut	8	472	4.67
Adelaar	Kotka	5	586	4.93
Bed	Sänky	5	9543	4.80
Mand	Kori	4	672	4.87
Druif	Rypäleen	8	204	5.00
Kop	Kuppi	5	2634	4.60
Bal	Pallo	5	5353	5.00
Zand	Hiekka	6	1035	4.93
Klok	Kello	5	2990	4.47
Tractor	Traktori	8	190	4.93
Vinger	Sormi	5	1870	4.87
Blad	Lehti	5	265	4.87
Slang	Käärme	6	1140	4.87
Spijkerbroek	Farkut	6	337	4.73
Pinguïn	Pingviini	9	147	4.87
Citroen	Sitruuna	8	613	4.87
Vogel	Lintu	5	2318	4.87
Spons	Sieni	5	342	4.93
Luipaard	Leopardi	8	276	4.87

Dutch Word	Finnish Word	Word Length Finnish Word	SUBTLEX	Concreteness Score Dutch Word
Vis	Kala	4	4258	4.73
Baby	Vauva	5	25978	5.00
Olifant	Norsu	5	580	4.93
Vaas	Maljakko	8	196	4.93
Mean	-	6.03	1876.5	4.81

Appendix B

Figure 1

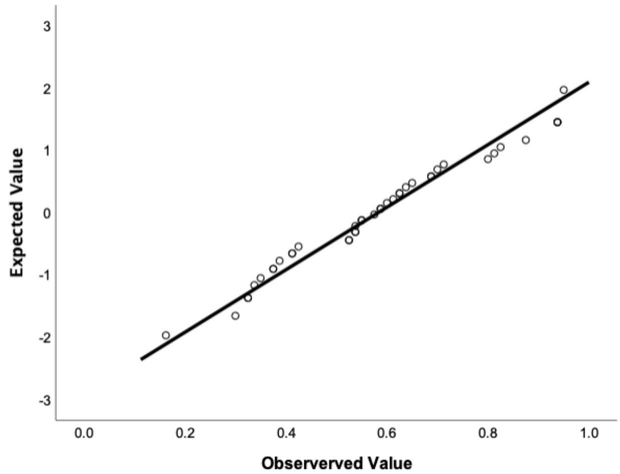
Distribution of Mean Accuracy Scores in Test Phase 2 and 3



Note. N = 40. Scores of participants in all conditions are included.

Figure 2

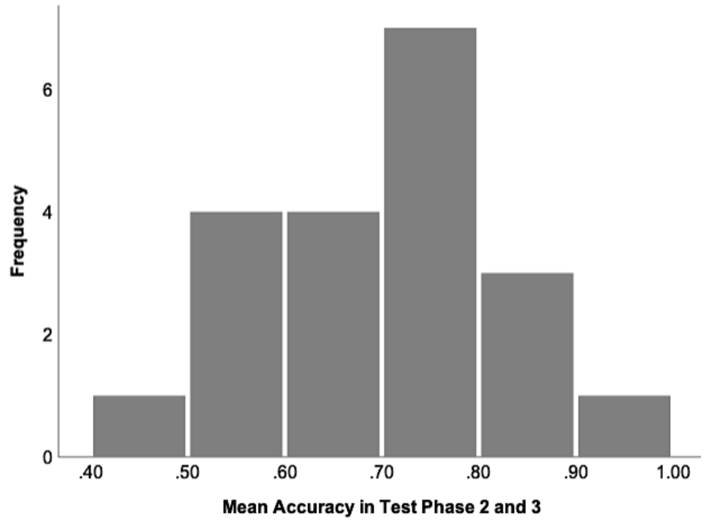
Normal Q-Q Plot of Mean Accuracy in Test Phase 2 and 3



Note. N = 40. Scores of participants in all conditions are included.

Figure 3

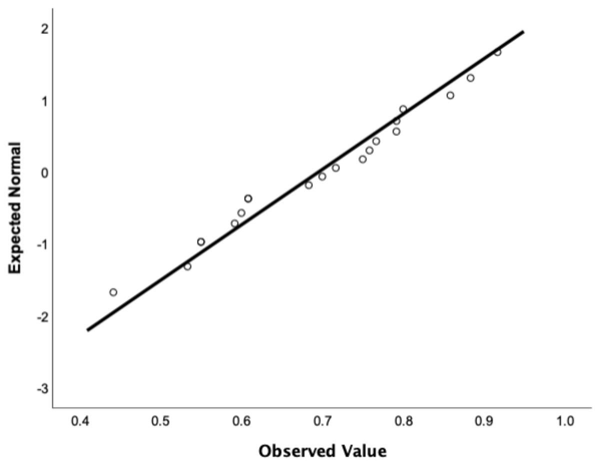
Distribution of Mean Accuracy Scores in Testphase 2 and 3



Note. Only scores of participants in condition ‘Image present’ were included with $n = 20$.

Figure 4

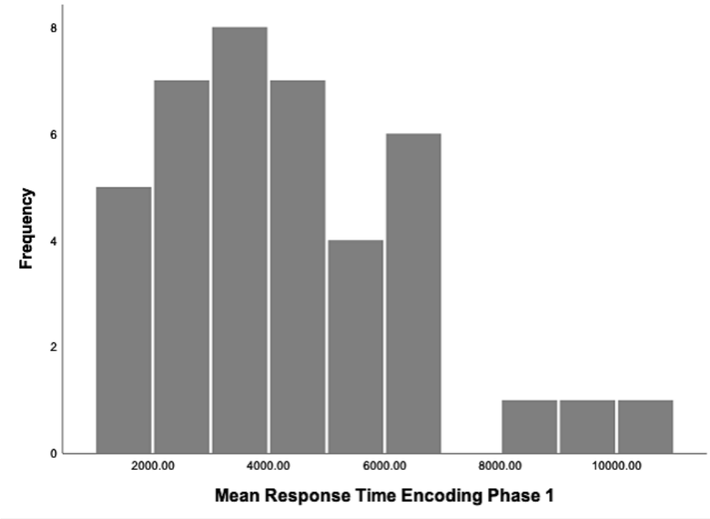
Normal Q-Q plot of Mean Accuracy in Testphase 2 and 3



Note. Only scores of participants in condition ‘Image present’ were included with $n = 20$.

Figure 5

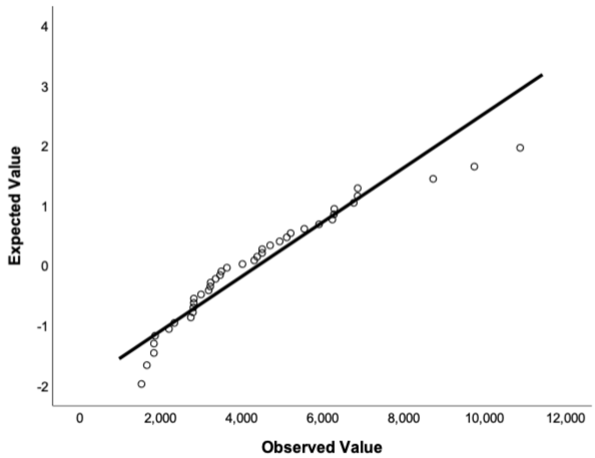
Distribution of Mean Response Time Scores in Encoding Phase 1



Note. $N = 40$. Scores of participants in all conditions are included.

Figure 6

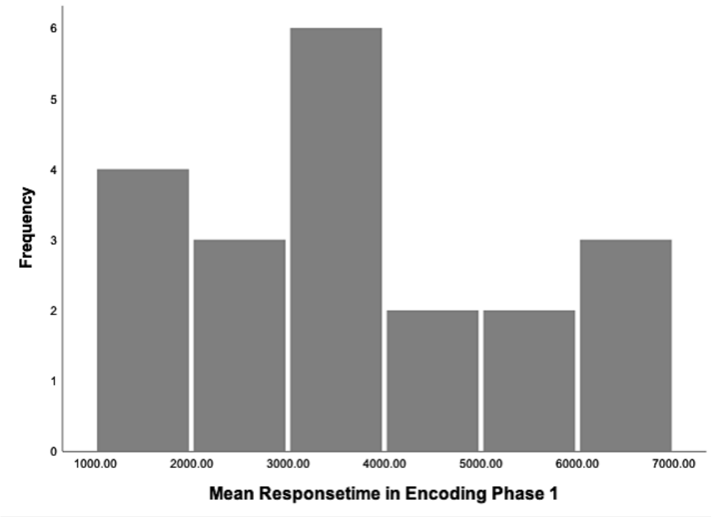
Normal Q-Q plot of Mean Response Time in Encoding Phase 1



Note. $N = 40$. Scores of participants in all conditions are included.

Figure 7

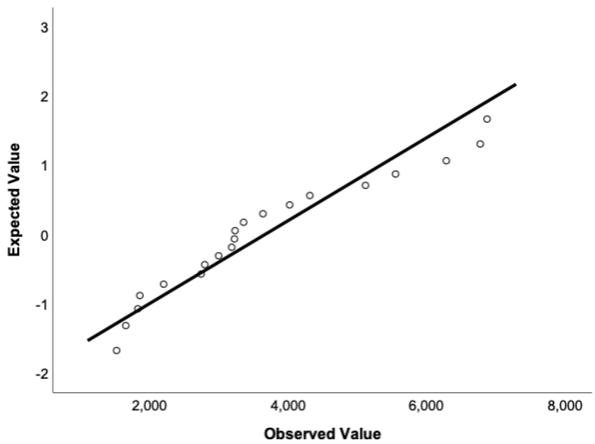
Distribution of Mean Response Time Scores in Encoding Phase 1



Note. Only scores of participants in condition 'Image present' were included with *n* = 20.

Figure 8

Normal Q-Q plot of Mean Response Time in Encoding Phase 1



Note. Only scores of participants in condition 'Image present' were included with *n* = 20