The influence of age of acquisition and frequency of used words on the performance on a visual word recognition task

Fleur Gramkow
6568564
Faculty of Social Sciences, Utrecht University
Master Social, Health, and Organizational Psychology
Track: Social Influence
Supervisor: Hans Marien
June 25th, 2023
Word count: 8595

This manuscript should be made publicly accessible

Abstract

Low-literacy poses a significant problem in the Netherlands, with an anticipated increase to 3.5 million low-literate individuals in 2024. Low-literacy individuals encounter daily challenges in areas such as healthcare, education and work. Numerous low-literate individuals remain unidentified because of intimidating testing methods and provided help by others. Incomplete insights in the actual problem makes it challenging to provide solutions for low-literate individuals. Speech-to-text software can provide solutions to investigate the actual literate levels of individuals by examining word recognition based on audio signals. This provides a less intimidating research method since it examined how well the software converts speech to text, rather than the performance of the low-literate individual. In current research, word recognition is investigated by simulating speech-to-text software, while examining the factors contributing to word recognition; age of acquisition and frequency. Current study focuses on the question to what extent the performance on a visual word recognition task is dependent on age of acquisition and frequency of used words. Results supported the influence of age of acquisition and frequency of used words on the reaction time and error rates on a visual word recognition task. Furthermore, the effect of age of acquisition and frequency of used words is not significantly greater when having to reject the words in the visual word recognition task. Current study offers a more thorough understanding of word processing and its application in speech-to-text software for low-literate individuals.

Keywords: age of acquisition, frequency, low-literacy, speech-to-text software, visual word recognition task

Introduction

Low-literacy is a significant problem in the Netherlands. In 2016, 2.5 million adults lacked functional literacy, implying that they are not able to read or write properly (National Literacy Trust (n.d.). Research from Schleicher (2019) indicates that in 2024 there will be 3.5 million low-literate individuals. Low-literacy individuals face numerous challenges on a daily basis. Research by Weiss, Hart, McGee & D'Estelle (1992) demonstrated that low-literacy is associated with health problems, which results from difficulties in understanding the wordings of healthcare systems and accessing healthcare services (Baker et al., 1996). Furthermore, low-literacy poses challenges to workforce participation and advancing education for children of low-literate individuals (Seidenberg, 2013). These challenges emphasize the importance of participating in low-literacy programs (Verkooijen & Wildenbos, 2020).

The Dutch government introduced the program called "Tel mee met Taal", to address the issue of low-literacy, and aims at reducing and preventing low-literacy (Bersee, 2019). Reaching all 2.5 million low-literate individuals in the Netherlands appears to be unfeasible (Ministerie van Algemene Zaken, 2019). The second phase of the program, from 2020 to 2024, focuses on reaching one million low-literate individuals for a tailor-made program (Ministerie van Algemene Zaken, 2019). Whether the intended target group of low-literate adults is actually reached is unknown, according to the mid-term evaluation conducted in 2022 (Ministerie van Onderwijs, Cultuur en Wetenschap, 2023).

Identification and management of this worldwide problem are complicated by the difficulty of approaching low-literacy populations. Because these individuals rely on spoken language to navigate by on a daily basis, a significant number of low-literate individuals remain unrecognized. This group might not yet be experiencing the consequences of their low-literacy, or they might be less inclined to ask for assistance (Ministerie van Algemene Zaken, 2019). Provided help of others contributes to experiencing less problems when not being able to read or write. Additionally, a portion of this unreachable population may possess poor learning skills, making identification more difficult (Ministerie van Algemene Zaken, 2019).

According to Baker et al. (1996), shame significantly contributes to the resistance to recognizing reading difficulties. Shame towards literacy is largely influenced by cultural norms and expectations. Individuals will be less inclined to identify their low-literacy and won't take part in low-literacy programs in a society with many educational opportunities. In addition, individuals with low-literacy feel more embarrassment to their low-literacy in health circumstances (Wolf, Williams, Parker, Parikh, Nowlan & Baker, 2007). Furthermore, those who have difficulty reading may feel intimidated and uncomfortable by the existing testing procedures for determining low-literacy individuals, which frequently include long and intensive reading exams (De Greef et al., 2015; Friedman & Hoffman-Goetz, 2006).

Overcoming these challenges is crucial to provide solutions for low-literacy individuals in the Netherlands. Solutions focused on spoken language could provide insights by investigating word recognition using a software. Research by Agarwal, Grover, Kumar, Puri, Singh, and Remy (2013) demonstrated the utilization of simple voice-based telecom systems instead of requiring expensive gadgets, while contemplating solutions for low-literate individuals. By utilizing software such as voice-based telecom systems or speech-to-text software, word recognition could be investigated by identifying the features that predict word recognition by analyzing spoken language.

Detection of low-literacy through spoken language

In accordance with the National Literacy Trust (n.d.), "literacy is the ability to read, write, speak, and listen in a way that lets us communicate effectively and make sense of the world." Low-literate individuals may not be able to read a book or a newspaper, may struggle to interpret price tags and road signs, may be unable to comprehend written forms and instructions, and may not be able to utilize the internet (National Literacy Trust, n.d.). Spoken language is a significant predictor for the development of literacy (Nathan, Stackhouse, Goylandris & Snowling, 2004). Children with speech impairments are more likely to encounter difficulties with reading-related activities, according to research by Nathan, Stackhouse, Goylandris, and Snowling (2004). Oral language skills in children play a crucial role in predicting their literacy development, and narrative skills provide valuable insights into these predictors (Klingner, 2004). Specifically, in various tasks, reading skills are investigated by analyzing the production of words or sentences (Tepperman et al., 2010). The existing relationship between spoken language, narrative skills and literacy, enables the detection of low-literacy based on spoken language, which can be applied in speech-to-text software.

While speech-to-text software originally were used for assisting individuals with hearing impairments, its usage has expanded across various fields (McCollum, Nation & Gunn, 2014). Speech-to-text software are already being deployed in schoolchildren who struggle with reading (McCollum, Nation, & Gunn, 2014). The software allows the translation of spoken words into written form by using a word processor to convert the audio files (McCollum, Nation, & Gunn, 2014). Machine learning can provide insights in the application of predictive models for low-literacy, to assess reading skills based on recordings of single word lists (Varol et al., 2009). The performance of the program is evaluated as a primary objective, rather than the capabilities of the person using it, by employing speech input as a source for implicit reading. This indicates that the performance of the low-literate individual is not evaluated, resulting in a less intimidating testing method.

Importance of word processing for speech-to-text software

To investigate the low-literacy on the basis of spoken language, word processing and word recognition should be investigated to understand the reading process. Fixation time is identified as a crucial aspect, which describes the length of time a reader needs to recognize a word after being fixated on it while reading (Juhasz & Rayner, 2006). Fixation time appears to be influenced by two important elements, which in turn affect word recognition. These two elements are the age at which a word is acquired and the frequency with which a word is used. According to Juhasz & Tayner (2006), both factors appear to have an impact on the recall and recognition memory of the words.

Age of acquisition (AoA) refers to the age at which individuals first encounter and use words, with words like ‘mommy’ and ‘daddy’ typically learned around the age of one. Children's language understanding significantly improves by the time they reach the age of 2.5 (Van Wiechenonderzoek - NCJ, 2023). Beyond the age of five, language comprehension shifts with the development of reading and writing skills along with longer, more complicated sentences (Zink, 2009; Schaerlaekens, 2008). Overall, early acquired words seem to be recognized better than late acquired words (Stadthagen, 2005). In both a semantic processing task (Brysbaert, Van Wijnendaele & De Deyne, 2000) and a visual word recognition task (Stadthagen-Gonzalez, Bowers, & Damian, 2004), early acquired words appear to be processed more quickly than late acquired words.

Frequency refers to the extent to which a word is used in a language. Words that are used more frequently in daily life are also more easily recognized compared to less frequently used words (Hansen, Siakaluk & Pexman, 2012). Based on research by Plourde & Besner (1997), a word's frequency of use has a significant impact on the extent to which a person performs in word recognition tests. According to Brysbaert et al. (2011), frequency is likewise the most important predictor in lexical decision tasks, indicating that words which are used more frequently are more likely to be utilized correctly.

Additional research into the association between word frequency and AoA refers to the cumulative frequency hypotheses (De Moor, Ghyselinck & Brysbaert, 2000). This hypothesis may clarify the effects of frequency and AoA by stating that early learned words with high frequency have the highest cumulative frequencies, whereas late acquired words with low frequency have the lowest cumulative frequencies. According to De Moor, Ghyselinck & Brysbaert (2000), by combining early AoA and high frequency it should be able to predict object naming speed more accurately than just taking into account AoA or frequency individually. However, discrepanties in the literature indicate questions whether which of the factors is more influenceable on word recognition. Morrison & Ellis (1995) questioned the primary influence of frequency on visual word recognition, by concluding that frequency is an artifact of AoA (Morrison & Ellis, 1995). Current study re-investigates the AoA and frequency effects of Ghyselinck, De Moor & Brysbaert (2000) to provide information about word processing in literate individuals. This knowledge is required to create speech-to-text software which facilitates learning written language based on spoken words for low-literate individuals, since it is feasible to identify features that predict word recognition by examining audio signals (Tepperman et al., 2010). Before this method is accessible, the tool needs to be developed and tested. Current research proposes manipulating the system by using a visual word recognition task, as a means to accomplish this goal.

Visual word recognition refers to the process of identifying a printed letter string as a specific word and comprehending its meaning (Rastle, 2016). When a particular representation in the orthographic lexicon reaches a particular level of activation, a word is identified. By varying AoA and frequency, previous studies (Gerhand & Barry, 1999; Morrison & Ellis, 1995) demonstrated independent effects on the visual word recognition test. A visual word recognition task enables to investigate the effects of frequency and AoA on word recognition based on spoken words, which is based on research by Tepperman et al. (2010), in which reading skills are investigated by tasks involving speech production. This information can be used to determine whether the participants recognize the words that were spoken out loud. However, as stated by Rastle (2016), results on the visual word recognition task are challenging to interpret due to the fact that AoA and frequency are highly correlated, and the AoA ratings are estimated by adults. Moreover, there is no consensus which factor has a higher influence on visual word recognition (Morrison & Ellis, 1995). Therefore, it is crucial to investigate the effects of AoA and frequency in different conditions to determine independent effects of both factors. A visual word recognition task could be used as a bogus task for imitating speech-to-text software, to assess word recognition by examining whether the software understood the word correctly, rather than the performance of the individual on the visual word recognition task.

In conclusion, current research aims to investigate the influence of AoA and frequency on the performance on a visual word recognition task. A detailed understanding of performance in such tasks is essential for gaining insights into the problem of low-literacy in the Netherlands. Furthermore, this knowledge can eventually provide solutions for these individuals for daily life problems, regarding health (Weiss, Hart, McGee & D'Estelle, 1992), education and work (Baker et al., 1996; Seidenberg, 2013). A visual word recognition task can provide more insights in the influence of AoA and frequency on word processing (Gerhand & Barry, 1999; Morrison & Ellis, 1995). If these factors and their effects are considered, speech-to-text software could be developed, which facilitates a way of investigating low-literate individuals in a non-intrusive way by investigating whether they recognize the visually presented word based on their spoken language. Current research focuses on the following question: *"To what extent is the performance on a visual word recognition task dependent on age of acquisition and frequency of used words?"* The hypotheses associated with this question are as follows: it is expected that performance on the visual word recognition task is higher for words that are acquired at a young age. In addition, it is expected that performance on the visual word recognition task is higher when the words are more frequently used. Finally, it is expected that these effects are stronger when the presented word needs to be rejected rather than accepted in the visual word recognition task. These hypotheses will be investigated by examining the performance on the visual word recognition task. In this task, participants have to indicate whether the word that appears on the computer screen is identical as the word the participant spoke out loud. The experimental design will simulate speech-to-text software evaluation, creating the perception that participants are engaging in software testing, while participants perform a visual word recognition task. Error rates and reaction time will be measured as indicators of performance on the task.

Methods

Participants

This study involved a total of 50 male and female participants with education levels varying from HBO to master's degrees. Specifically, the distribution consisted of 35 individuals with a bachelor's degree, ten individuals with a HBO degree, and five individuals with a master's degree. All participants were fluent in Dutch and were aged between 18 and 54 years old (M = 23.44, SD = 7.51). The gender distribution included 44 female participants, and 6 male participants. 46 participants were right-handed, whereas four participants were left handed. Five participants were raised bilingually, and 47 participants were native Dutch speakers.

Through a selective convenience sampling, participants were recruited. Recruitment methods included sharing an introductory text via social media, WhatsApp, and distributing flyers at the Utrecht Science Park. Additionally, students were recruited through the Social and Behavioral Sciences Research Participation System (SONA systems). Participants received an incentive of five euros or one PPU after finishing the experiment.

A sample size of at least 33 participants was determined as necessary for a medium effect size using repeated measures ANOVA, based on power analysis (Cohen, 1992). To account for possible drop outs and outliers, this research aimed to recruit 50 participants.

Materials

The program Gorilla was used for programming the experiment and developing the questionnaire. For the experiment, 40 different auditory stimuli were used, containing spoken words. The length of the stimuli was dependent on the length of each spoken word. All auditory stimuli were presented through headphones. The Gorilla program was used to present the visual stimuli on a computer screen, which consisted of 80 words. Furthermore, a microphone connected to the computer was used to record the response of the participant, which consists of the word they spoke out loud. These auditory responses were encoded by the microphone function of Gorilla.

The conditions for this experiment were based on the factors used in research of Ghyselinck, De Moor & Brysbaert (2000) regarding AoA and frequency. The conditions were *early AoA and high frequency*, *early AoA and low frequency*, *late AoA and low frequency* and *late AoA and high frequency*. The words that were chosen for these conditions were based on 1 standard deviation difference from the means of the AoA and frequency of the words. Age is used for assessing AoA, Logfreq is used for the frequency levels. For the whole list of words used in the experiment, see Appendix A.

Words that belong to the *early AoA and high frequency* condition, contained words with a mean of 9.442 - 1 SD (2.767) = 6.675 years old or earlier AoA, and a frequency of used words are selected when used more than: 2.010 + 1 SD (0.905) = 2.916 times. Words that belong to the *early AoA and low frequency* contained words that also have a mean of 6.675 or earlier age, and a frequency of 2.011- 1 SD (0.905) = 1.105 or less used. The *late AoA and low frequency* contained words with an AoA of 9.442 + 1 SD (2.767) = 12.209 years or older, and a frequency of 1.105 or less used. Finally, the *late AoA and high frequency* condition was based on words differing 0.57 SD from the mean for both AoA and frequency as there was a limited number of words within the mean ± 1 SD range for this condition. The specific values used for this condition are provided in Appendix B.

The auditory stimuli consisted of audio files containing spoken words. The words belonging to the certain conditions were pronounced by a digital text-to-speech generator (Realistic Text to Speech converter & AI Voice generator. (n.d.). The text-to-speech generator generated words containing the same settings for every auditory stimulus. The language used was Dutch, the speech speed was set to 0.8 and the selected voice was called 'Colette'.

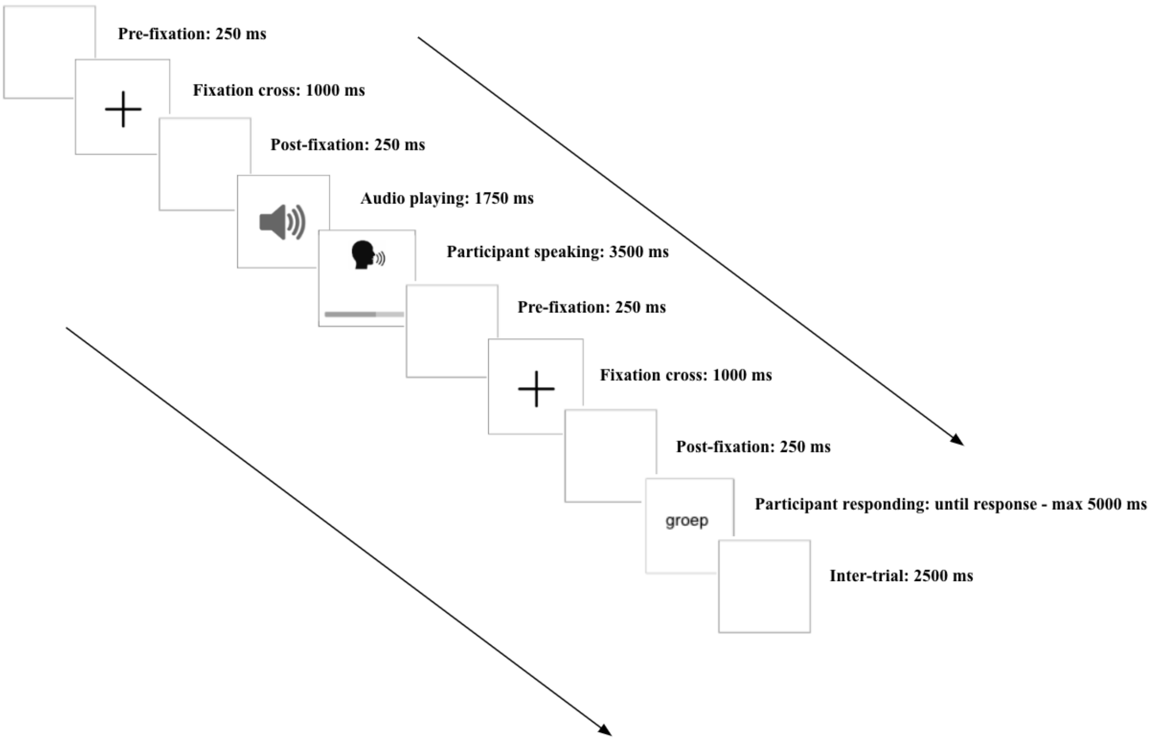
Whereas the 40 selected words from the list of Ghyselinck, De Moor & Brysbaert (2000) form the accept condition in the experiment, for each word from the word list of Ghyselinck, De Moor & Brysbaert (2000), a minimal pair was selected to develop the reject condition of the experiment. For the selection of minimal pairs, equal distribution of minimal pair positions was maintained for all conditions. Each condition contained five words with the first or last letter differing, and five words with a different letter in the middle part of the word. This control was implemented to account for potential confounding factors related to the position of letters in word recognition.

Task

Prior to conducting the experiment, ethical approval was obtained from the Ethical Commission of *Faculty of Social and Behavioural Sciences* (reference number 23-0134). This experiment involved participants engaging in a visual word recognition task, requiring them to respond as quickly and accurately as possible. Each participant was presented with auditory stimuli for 1750 ms, for which they were instructed to listen to the word and repeat the word verbally within 3500 ms. Thereafter, participants had to indicate whether the word that appeared on the screen was the same as the word they had spoken out, within 5000 ms. 1000 ms fixation screens were shown between these elements of the experiment. The specific timings of the screens are depicted in Figure 1. The participants were instructed to use the X and N keys for "same as spoke out loud" and "different than spoke out loud". The assignment of keys was counterbalanced across participants. Participants were made to believe that the computer produced their spoken words. Participants might have attributed any differences in the reject condition, where the spoken word was different from the one displayed on the computer screen, to a converting mistake of the computer. However, in reality, participants were performing a visual word recognition task, to investigate the influence of AoA and frequency on the results of the visual word recognition task via this bogus task.

Figure 1

Presented screens and timings in the visual word recognition task



Procedure

Individuals could sign up for the experiment through Sona or by sending a private message, for a specific time and day. The experiment took place in the laboratory of the Langeveld building. In one of the laboratory rooms, participants were instructed to take a seat at a table in front of a computer, and to put headphones on. The experiment started with providing information about the experiment on the computer screen. This informed participants that they would be participating in an experiment involving the recording and decoding of their spoken words to test the amounts of errors made by the speech-to-text software. This was followed by an informed consent which participants were required to actively accept before proceeding with the experiment. See Appendix C for the full informed consent and information texts. After the informed consent, participants were told that after the experiment some demographics and short questions would be conducted.

Following, the instructions of the experiment were presented. See Appendix D for the full instructions. A test phase occurred in which two words (“stoof” and “groep”) and their minimal pairs (“stoor” and “greep”) were presented in order to assess participants' comprehension of the task. These words had an average AoA and frequency; they did not fit into any of the conditions. For the two words being used in the test phase, participants were directed to evaluate the word using the X and N keys. Once the participant pressed the spacebar to turn to the next slide, the experiment began.

The task consisted of four sets of 20 words, distributed across four different conditions. Participants were given the opportunity to take a break in between sets. The assignment took approximately 15 minutes, and once the four trials were finished, the participants were asked to proceed on with additional questions. Demographic information was gathered and the comprehension of the words used in the experiment was evaluated. Questions regarding reading preference and frequency were posed. Additionally, participants were questioned regarding their familiarity with each of the words used in the experiment, and for one word in each condition, the frequency and age at which the participant learned the word were recorded. This allowed to confirm the manipulation. For a detailed description of the questions, see Appendix E. When participants finished the experiment, after approximately 30 minutes, they were instructed to leave the room. The compensation was provided at the end of the experiment.

Data analysis

Questionnaire

To determine whether the experimental manipulation was carried out correctly and to gain more insights in the participants' reading habits and preferences, a questionnaire was conducted. These insights could be used for future research and to offer alternative explanations in the occurrence of not finding any significant results.

The first question in the manipulation check asked participants to indicate whether they had correctly heard all the spoken words. Subsequently, the participants had to indicate twice for each pair of words which they used more frequently and twice for each pair of words which they knew longer. For each word, the percentage of ‘more frequently used’ and ‘known longer’ is computed. How much percentage of the manipulation check was carried out correctly was calculated for each participant individually. Subsequently, participants were given the task to report whether they were familiar with the words utilized in the experiment.

Each participant's percentage of words they recognized is calculated along with the percentage of words that were familiar to them. The demographics section of the questionnaire included questions about gender, age, education level, hand dominance, native language use of Dutch, and bilingualism. Afterwards, the preference and frequency of reading are examined by asking the participants how frequently they read books, articles, or newspapers. For each of the seven response options, the mean percentage of chosen answers was calculated. Additionally, a specific score ranging from zero to 100 was used to calculate the participants' lifetime average interest in reading. Similarly, the average reader's current level of enjoyment in reading was evaluated.

Repeated measures ANOVA

In order to interpret, present, and manipulate data, SPSS Statistics 2 was used. Gorilla was employed to collect the data, and SPSS was used to prepare it for analysis. The repeated measures ANOVA's assumptions were examined before the analysis was performed. Mauchly's test was employed to determine sphericity, examining the residuals allows for the investigation of normality. The distribution of the residuals was examined in order to assess the homogeneity of variance. Additionally, by measuring several words of various conditions within each participant, the assumption of independence was examined. Prior to performing the analysis, outliers were eliminated. Outliers were defined as trials with more than 2.5 SD difference for each of the conditions. The removal of outliers was conducted on the correct trials.

Subsequently, the repeated measures ANOVA was performed. The task included a 2x2x2 design and four trials containing 20 words. AoA, frequency, and evaluation (accept or reject), each with two levels, comprised the independent variables. The analysis is performed to investigate two main effects; to which extent the AoA affects performance on the visual word recognition task, and to which extent frequency affects performance on the visual word recognition task. Measurements of the 80 involved words were made on error rates and reaction time variations as dependent variables. Subsequently, to assess whether AoA, frequency and evaluation combined had a significant effect on the error rates and reaction times, a three-way interaction was conducted.

Results

Questionnaire

Results regarding the manipulation check indicated the following results. For 50% of the participants, all the words used in the experiment were heard loud and clear. For two words in the first condition (high frequency, early AoA), *hand* was used more (74%) than the word *stamp* (low frequency, early AoA) and *thee*[1] (high frequency, early AoA) was known for 90% longer than the word *norm* (high frequency, late AoA). For the words in the second condition (high frequency, late AoA), *omzet* was used 84% more than *puls* (low frequency, late AoA. For the third condition (low frequency, early AoA), *lego* was known 92% longer than *slenk* (low frequency, late AoA). Concluding, for 34 participants the manipulation was done 100% properly, they indicated for each word the right value of AoA and frequency. These results indicate that the manipulation is done properly, for one word of each condition is made sure that the values regarding AoA and frequency match the results from the list of Ghyselinck, De Moor & Brysbaert (2000). Subsequently, familiarity with the words used in the experiment was conducted. For the participants that were not familiar with all the words, the mean percentage of words that were not understood was 14%.

According to these results, The AoA list from Ghyselinck, De Moor & Brysbaert (2000) is an accurate representation of the age at which participants in this experiment learned the words as well as the frequency with which they used the words. The distribution of how often participants read books, articles, or newspapers consisted of the following results; 38% of the participants reads books, articles or newspapers two to three times a week, 24% of the participants reads books, articles or newspapers once a week, 24% of the participants reads books, articles or newspapers a couple times per month, 10% of the participants reads books, articles or newspapers once per month, and 4% of the participants never reads books, articles or newspapers. The mean interest in reading during the lifecourse of the participants had a score of 69.08 out of 100, the mean enjoyment of reading was 58.50 out of 100.

These findings demonstrated that the manipulation was carried out as intended; the words from the list of Ghyselinck, De Moor & Brysbaert (2000) contained the AoA and frequency values that also applied for the experiment's participants. Reading preference and frequency indicated the general level of reading of the participants, to investigate the influence of this level on the performance on the visual word recognition task in the future.

[1] For the manipulation of condition 1, 'thee' is used instead of 'twee' due to a mistake. This word has a slightly higher AoA but also a higher frequency, meaning the word almost fits in the condition

The questionnaire results revealed the participants' preferences for reading and reading frequency, which corroborate the manipulation check of reading, and indicates the high levels of both which could be influencing the performance on the visual word recognition task

Repeated measures ANOVA

A repeated measures ANOVA is conducted to investigate the effects of AoA and frequency on the performance on a visual word recognition task. Performance is investigated using error rates and reaction time. Before conducting the analysis, the assumptions of sphericity, normality, homogeneity of variance and independence of observations were examined.

The assumption of sphericity is investigated using Mauchly's test, which indicates that the assumption is met for reaction time and error rates. The assumption of normality is examined by inspecting the distribution of the residuals. The residuals looked approximately normal for the effect on reaction time, as well as for error rates. Even though the distribution was not completely normal, the repeated measures ANOVA is robust against this violation. Homogeneity of variances is examined by inspecting the distribution of residuals. This assumption is violated. However, homogeneity of variances rarely holds in a repeated measures ANOVA design (Hertzog & Rovine, 1985), which makes violation of the assumptions less important. Finally, the assumption of independence of observation is met, by measuring multiple words of different conditions within each participant.

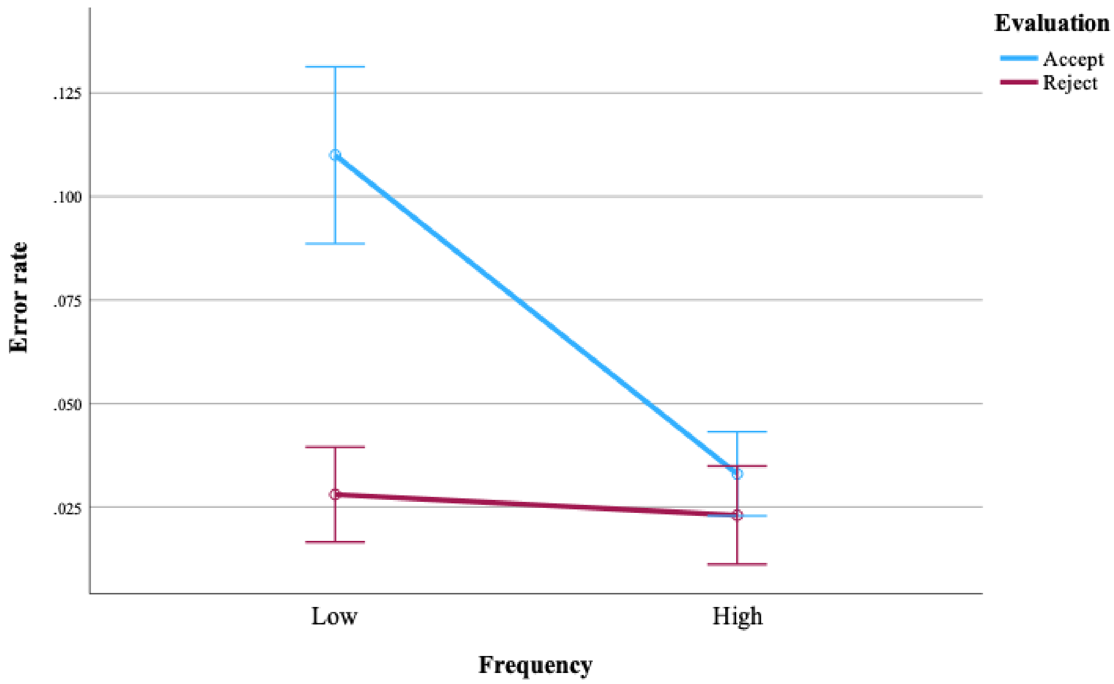
Reaction times

Results of the repeated measures ANOVA investigating the influence of frequency, AoA and evaluation are shown in Table 1 and Table 3 and indicated the following results on reaction time; a significant main effect is found for the effect of frequency on reaction time, $F(1,49) = 20.453$, $p < .001$, $\eta^2 = .294$. High frequency used words indicated a shorter reaction time ($M = 767.946$, $SD = 172.793$) than low frequently used words ($M = 804.191$, $SD = 195.825$). For the main effect of AoA, the repeated measures ANOVA showed a significant main effect of AoA on reaction time ($F(1,49) = 24.308$, $p < .001$, $\eta^2 = .332$. For early acquired words, reaction time was shorter ($M = 766.066$ $SD = 178.558$) than the reaction time for late acquired words ($M = 806.071$, $SD = 190.059$). For the main effect of evaluation, the repeated measures ANOVA showed a significant main effect of evaluation on reaction time ($F(1,49) = 55.111$, $p < .001$, $\eta^2 = .529$). In the accept condition, participants had shorter reaction times ($M = 746.640$ $SD = 191.296$) compared to reaction times in the reject condition ($M = 825.507$, $SD = 177.322$).

The repeated measures ANOVA indicated the following results regarding the interaction effect of AoA, frequency, and condition on reaction time: no significant effect is found between frequency and AoA ($F(1,49) = 3.153$, $p = .082$, $\eta^2 = .060$). No interaction effect is found between AoA and evaluation on reaction time ($F(1,49) = .329$, $p = .569$, $\eta^2 = .007$). In addition, the interaction between the frequency, AoA and evaluation had a non-significant effect on reaction time $F(1,49) = .185$, $p = .669$, $\eta^2 = .004$. However, a significant interaction is found between frequency and evaluation on reaction time ($F(1,49) = 36.383$, $p < .001$, $\eta^2 = .426$), as illustrated in Figure 2. A Bonferroni corrected pairwise comparison is used to examine the interaction effect between frequency and evaluation on reaction time. The results indicated for the reject condition that for low frequency ($M = 825.228$, $SE = 24.723$) and high frequency ($M = 825.786$, $SE = 23.366$), no significant effect is observed on reaction time ($F(1,49) = .003$, $p = .954$). For the accept condition, the results for low frequency ($M = 783.154$, $SE = 28.023$) and for high frequency ($M = 710.106$, $SE = 23.366$) indicate a significant effect on reaction time ($F(1,49) = 49.127$, $p <.001$).

Figure 2

The interaction between evaluation and frequency on error rates



Note. Error bars: +/- 2 SD

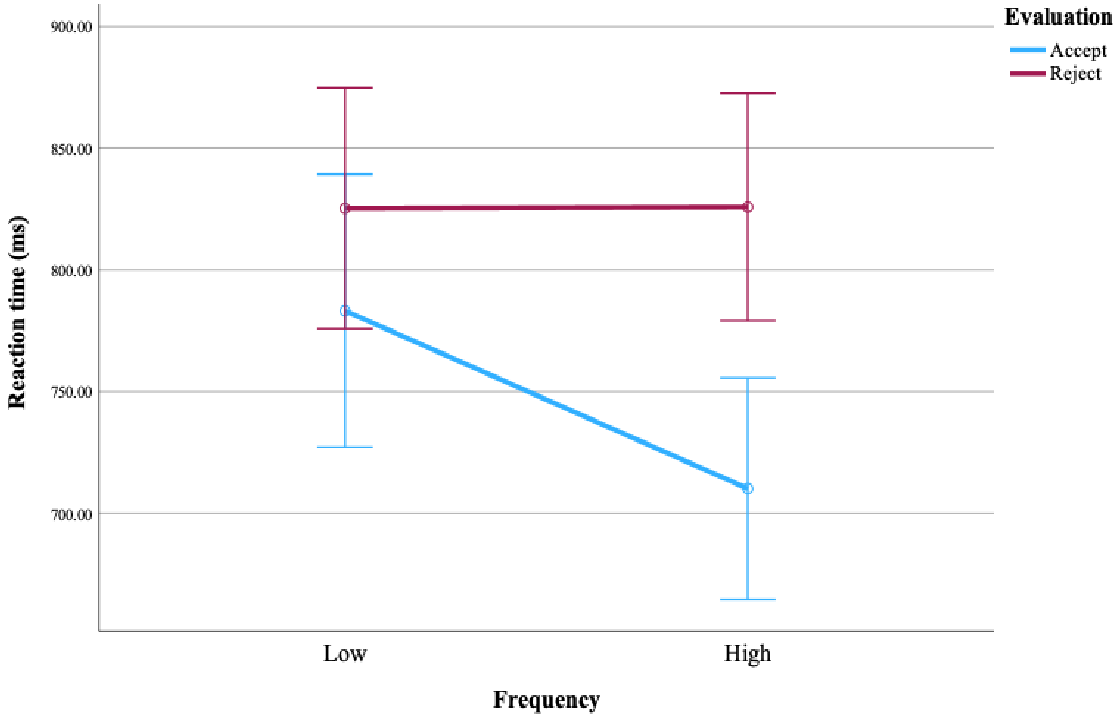
Error rates

The repeated measures ANOVA provided the following results of frequency, AoA and evaluation on error rates, and are shown in Table 2 and 3; the analysis indicated a significant main effect of frequency on error rates, $F(1,49) = 46.562$, $p < .001$, $\eta^2 = .487$. High frequency used words showed less error rates ($M = .028$ $SD = .050$) compared to error rates for low frequently used words ($M = .069$ $SD = .083$). For the main effect of AoA, the repeated measures ANOVA showed a significant main effect on error rates, ($F(1,49) = 11.671$, $p < .001$, $\eta^2 = .192$. For early acquired words, the error rate was lower ($M = 0.037$, $SD = .055$) compared to the error rates for late acquired words ($M = .060$, $SD = .068$). For the main effect of evaluation, the repeated measures ANOVA showed a significant main effect of evaluation on error rates ($F(1,49) = 65.457$, $p < .001$, $\eta^2= .572$). On the contrary, in the accept condition participants indicated more error rates ($M = .072$ $SD = .074$) than in the reject condition ($M = .026$, $SD = .068$. This discrepancy in results of evaluation on error rates compared to reaction time, indicates that low frequently used words result in the possibility of making more errors in the task.

The repeated measures ANOVA indicated the following results for the interaction effect between AoA, frequency and condition on error rates: no significant effect between frequency and AoA on error rates ($F(1,49) = 2.922$, $p = .094$, $\eta^2 = .056$). No interaction effect is found between AoA and evaluation on error rates ($F(1,49) = 2.537$, $p = .118$, $\eta^2 = .049$). In addition, the interaction between frequency, AoA and evaluation on error rates turned out to be not significant ($F(1,49) = 1.396$, $p = .243$, $\eta^2 = .028$). The interaction between frequency and evaluation on error rates turned out significant ($F(1,49) = 31.689$, $p < .001$, $\eta^2 = .393$), as illustrated in Figure 3. A Bonferroni corrected pairwise comparison is used to examine the interaction effect between frequency and evaluation on error rates. The results indicated for the reject condition that for low frequency ($M = .028$, $SE = .006$) and high frequency ($M = .023$, $SE = .006$), no significant effect is observed on error rates ($F(1,49) = 2.333$, $p = .133$). For the accept condition, the results for low frequency ($M = .110$, $SE = .011$) and for high frequency ($M = .033$, $SE = .005$) indicate a significant effect on error rates ($F(1,49) = 64.909$, $p <.001$).

Figure 3

The interaction between evaluation and frequency on reaction time (ms)



Note. Error bars: +/- 2 SD

Overall, the main effects found with the repeated measures ANOVA provide evidence for accepting the first two hypotheses, indicating that performance on the visual word recognition task is higher for words that are acquired at an early age, and that performance on the visual word recognition task are higher when they are more frequently used. Furthermore, the conducted interaction effects indicate that the third hypothesis can not be accepted. The effect of AoA and frequency are not stronger when the presented word needs to be rejected instead of accepted in the visual word recognition task. The only visible interaction effect is the interaction between frequency and evaluation. This interaction indicates that only when words have to be accepted instead of rejected, the effect on reaction time and error rates is higher for words that are more frequently used.

Table 1

Descriptive statistics of frequency, AoA and evaluation (accept/reject) on reaction time

			Age of acquisition			
			Early		Late	
		Evaluation	*M*	*SD*	*M*	*SD*
Frequency	Low	Accept	765.498	196.648	800.811	233.334
		Reject	816.463	183.396	833.994	179.920
	High	Accept	682.686	165.148	737.526	180.052
		Reject	799.618	169.041	851.954	176.931

Table 2

Descriptive statistics of frequency, AoA and evaluation (accept/reject) on error rates

			Age of Acquisition			
			Early		Late	
		Evaluation	*M*	*SD*	*M*	*SD*
Frequency	Low	Accept	.086	.086	.134	.108
		Reject	.0180	.048	.380	.057
	High	Accept	.026	.044	.040	.057
		Reject	.018	.044	.028	.054

Table 3

Results of the repeated measures ANOVA

	Reaction time				Error rates			
	df	F	p	η^2	df	F	p	η^2
Frequency	1	20.453	< .001*	.294	1	46.562	< .001*	.294
AoA	1	24.308	< .001*	.192	1	11.671	< .001*	.192
Evaluation	1	55.111	< .001*	.529	1	746.640	< .001*	.572
Frequency * AoA	1	3.153	.082	.060	1	2.922	.094	.056
Frequency * Evaluation	1	36.383	< .001*	.426	1	31.689	< .001*	.393
AoA * Evaluation	1	.329	.569	.007	1	2.537	.118	.049
Frequency * AoA * Evaluation	1	.185	.669	.004	1	1.396	.243	.028

Discussion

The findings of current research provide insights in the association between age of acquisition and frequency of used words, and their influence on error rates and reaction times on a visual word recognition task. The first significant finding demonstrates how early learned words influence the performance on a visual word recognition task. These results are consistent with earlier research from Stadthagen-Gonzalez, Bowers & Damian (2004) and Brysbaert, Van Wijnendaele & De Deyne (2000), which indicated that early acquired words appeared to be processed more quickly than late acquired words in a visual word recognition task. Current study supports these findings. Furthermore, based on research by Plourde & Besner's (1997), the impact of word frequency on the visual word identification task is investigated. According to these findings, a word's frequency of use is a significant predictor of how well a person performs in a visual word recognition task. This conclusion is also supported by the current research's findings.

Furthermore, current research addressed whether the effects of AoA and frequency on performance on a visual word recognition task are stronger when the words should be rejected rather than accepted. The effects of AoA, frequency, and evaluation on reaction times and error rates for the visual word recognition task did not interact significantly, according to current research. The variables frequency and evaluation do, however, have a significant interaction effect on both reaction times and error rates. In this instance, more often used words suggest quicker reaction times and lower error rates.

The manipulation was carried out appropriately, according to the questionnaire results regarding understandance and use of the words in the experiment. In this research, the words from the list of Ghyselinck, De Moor & Brysbaert (2000) contained the AoA and frequency values that were assigned to the words in their research. The results of the study on reading preference and frequency showed that the manipulation of reading levels was carried out correctly, and that the high level of reading preference and frequency could be employed to investigate the influence of these levels on the performance of the visual word recognition task in further research. It could suggest that a person's ability to perform well on the visual word recognition task may be influenced by their preference for reading and frequency of reading.

Current research indicated a more thorough comprehension of word processing and word recognition. The first two hypotheses could be accepted, which are in line with the presented literature (Stadthagen-Gonzalez, Bowers & Damian, 2004; Plourde & Besner, 1997), indicating that performance on the visual word recognition task is higher for words that are acquired at an early age and are used more frequently. However, the third hypothesis indicating that the effects of AoA and frequency would be stronger when presented words should be rejected instead of accepted in the visual word recognition task, could not be accepted.

Despite the absence of a significant interaction effect between frequency, AoA and evaluation, a significant interaction effect between frequency and evaluation is observed. The absence of finding a significant interaction effect between frequency, AoA and evaluation could be explained by the fact that words in the reject condition are solely visually presented to the participants. Expected was that for the reject condition of the words, the effects of frequency and AoA would be higher. However, the reject condition consists of minimal pairs of the used words from the AoA and frequency list of Ghyselinck, De Moor & Brysbaert, 2000. For the minimal pairs, AoA and frequency are not taken into account. It is feasible that the minimal pairs differ in frequency and AoA values compared to the accept condition. Frequency and AoA are manipulated for the minimal pairs in the reject condition, so no effects of frequency and AoA could be observed in this condition.

Another possible explanation of an absent interaction effect could be related to the different cognitive processes in the different conditions. In the reject condition, participants had to evaluate whether the words they spoke out differed from the word that was presented on the screen. Different cognitive processes are involved when reading and comprehending text, and distinct mental processes play a role in reading words (Weir & Khalifa, 2008). Less-skilled readers may show comprehension deficits for words because of differences in cognitive processes (Stanovich, 1982). For the reject condition, more cognitive processing is needed, to examine whether there is a difference in the auditory and visual presentation, which could result in more error rates.

However, a significant interaction is found between the frequency and accept condition, on both reaction time and error rates. The interaction between frequency and the accept condition could be explained by considering that when participants need to accept words, the measurement is based on recognizing the visually presented word. This is consistent with research by Hansen, Siakaluk & Pexman (2012), indicating that words that are more frequently used are recognized more rapidly. Moreover, Brysbaert et al. (2011) indicates that high frequently used words are more likely to be used correctly. Additionally, research by Morrison & Ellis (1995) questioned the primary influence of frequency in visual word recognition, and concluded that frequency effects are an artifact of AoA effects. Current research proves the opposite and shows evidence for the effect of frequency on evaluating visually presented words. It is important to note that in the current experiment, the accept and reject rates are manipulated with a 50/50 rate. When implementing speech-to-text software, accuracy is mostly above 50%, meaning that more trials would be within the 'accept' condition. The interaction effect between frequency and accept is therefore promising, since most of the observations in the speech-to-text system will be accepted.

Likewise, these promising effects provide opportunities for investigating word recognition in low-literate individuals due to the fact that processing of words is almost identical in all individuals (Zink, 2009; Schaerlaekens, 2008; Hansen, Siakaluk & Pexman, 2012). This indicates that the influence of high frequency on word recognition also accounts for low-literate individuals. Given the small vocabulary of low-literate individuals, high frequently used words would be recognized faster by them. Recognizing is the main goal in the accept condition, which is the objective of the speech-to-text software. Recognizing their own spoken words by using speech-to-text software, would indicate that they would recognize high frequently used words more easily in written language, compared to when they pronounce words that are not frequently used, which could result in enhanced learning of written language.

Several limitations should be considered when interpreting the findings. Even though current research contributes to the knowledge of word processing in the general population, no additional information is provided about the target group, low-literate individuals. Low-literate individuals are hard to identify, and they keep mostly under the radar because of shame and uncomfortableness (Baker et al., 1996; Friedman & Hoffman-Goetz, 2006). This not representative data results in a lack of generalizability of the information for the actual target audience (Smaling, 2016).

In addition, the used list for assessing AoA and frequency, is retrieved from research by Ghyselinck, De Moor & Brysbaert (2000), which is conducted in Belgium, while current research is conducted in the Netherlands. The languages of both countries are variants of one language, where Flemish is the Belgian variant of Dutch (Geeraerts, 2001), and all the words in the experiment are in the Dutch dictionary. However, Flemish and Dutch do have some differences. It could be the case that the levels of AoA and frequency slightly differ between the two populations, because their use of the language could be slightly different. Since all the participants in current research are speaking Dutch, this could impact the performance of the participants on the visual word recognition task, particularly for words that are more easily recognized by Flemish individuals.

Considering these limitations, future research could be focused on obtaining a more representative sample for the actual low-literacy group in the Netherlands, and a better alignment between the used wordlist and the participants. This could involve conducting the visual word recognition task in the Flemish part of Belgium. Further research could be focused on researching actual low-literate individuals. The Dutch program “Tel mee met Taal” (Bersee, 2019) aims to reach more low-literate individuals by 2024, for a tailor-made program. By cooperating with the Dutch government and using the reached low-literate individuals for research, real solutions for daily problems of low-literate individuals could be provided.

Current research addresses the influence of the age at which words are acquired and frequency of used words on the performance in a visual word recognition task. By examining these factors, this study aimed to contribute more detailed information about word processing, to incorporate these findings in developing speech-to-text software for low-literacy individuals because of the predictive relationship between spoken language, narrative skills and literacy. Addressing the increasing number of low-literacy individuals in the Netherlands (Schleicher, 2019), who face challenges regarding health, education, and work (Baker et al., 1996; Seidenberg, 2013), is crucial to discover effective solutions for this population. However, the identification of low-literacy individuals is challenging (Ministerie van Algemene Zaken, 2019), which makes speech-to-text software valuable in facilitating this process. The software investigates how well spoken language is translated into text by analyzing audio signals, which makes the process of identifying low-literacy less intimidating. However, future research is needed to reach low-literate individuals and investigate the effect of AoA and frequency in low-literate individuals. These insights would be valuable for implementing the speech-to-text software, which could be utilized to identify those individuals and provide them with a tailored program to help them overcome challenges in daily life.

Literature

Agarwal, S. K., Grover, J., Kumar, A., Puri, M., Singh, M., & Remy, C. (2013). Visual conversational interfaces to empower low-literacy users. In *Human-Computer Interaction–INTERACT 2013: 14th IFIP TC 13 International Conference, Cape Town, South Africa, September 2-6, 2013, Proceedings, Part IV 14* (pp. 729-736). Springer Berlin Heidelberg.

Baker, D. W., Parker, R. M., Williams, M. V., Pitkin, K., Parikh, N. S., Coates, W., & Imara, M. (1996). The health care experience of patients with low literacy. *Archives of family medicine, 5*(6), 329.

Bersee, T. (2019). Een nieuw plan in de eeuwigdurende strijd tegen laaggeletterdheid. *Sociaal Bestek, 81*(4), 16-18.

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect. *Experimental psychology.*

Brysbaert, M., Van Wijnendaele, I., & De Deyne, S. (2000). Age-of-acquisition effects in semantic processing tasks. *Acta psychologica, 104*(2), 215-226.

Cohen, J. (1992). Statistical power analysis. *Current directions in psychological science, 1*(3), 98-101.

De Greef, M., Segers, M., Nijhuis, J., Lam, J. F., van Groenestijn, M., van Hoek, F., ... & Tubbing, M. (2015). The development and validation of testing materials for literacy, numeracy and digital skills in a Dutch context. *International Review of Education, 61*(5), 655-671.

De Moor, W., Ghyselinck, M., & Brysbaert, M. (2000). A validation study of the age-of-acquisition norms collected by Ghyselinck, De Moor, & Brysbaert. *Psychologica Belgica, 40*(2), 99-114.

Friedman, D. B., & Hoffman-Goetz, L. (2006). A systematic review of readability and comprehension instruments used for print and web-based cancer information. *Health Education & Behavior, 33*(3), 352-373.

Geeraerts, D. (2001). Een zondagspak? Het Nederlands in Vlaanderen: gedrag, beleid, attitudes. *Ons erfdeel, 44*(3), 337-343.

Gerhand, S., & Barry, C. (1999). Age of acquisition, word frequency, and the role of phonology in the lexical decision task. *Memory & cognition, 27*(4), 592-602.

Ghyselinck, M., De Moor, W., & Brysbaert, M. (2000). Age-of-acquisition ratings for 2816 Dutch four-and five-letter nouns. *Psychologica Belgica, 40*(2), 77-98.

Hansen, D., Siakaluk, P. D., & Pexman, P. M. (2012). The influence of print exposure on the body-object interaction effect in visual word recognition. *Frontiers in Human Neuroscience, 6*, 113.

Hertzog, C., & Rovine, M. (1985). Repeated-measures analysis of variance in developmental research: Selected issues. *Child development,* 787-809.

Juhasz, B. J., & Rayner, K. (2006). The role of age of acquisition and word frequency in reading: Evidence from eye fixation durations. *Visual Cognition, 13*(7-8), 846-863.

Klingner, J. K. (2004). Assessing reading comprehension. *Assessment for effective intervention, 29*(4), 59-70.

The National Literacy Trust. (n.d.). What is literacy? Retrieved June 20, 2023, from https://literacytrust.org.uk/information/what-is-literacy/#:~:text=People%20with%20low%20literacy%20skills,medicines%20or%20use%20the%20internet.

McCollum, D., Nation, S., & Gunn, S. (2014, January). The Effects of a Speech-to-Text Software Application on Written Expression for Students with Various Disabilities. In *National Forum of Special Education Journal* (Vol. 25, No. 1).

Ministerie van Algemene Zaken. (2019, March 18). *Kamerbrief over plan van aanpak laaggeletterdheid 2020 – 2024.* Kamerstuk | Rijksoverheid.nl. https://archief28.sitearchief.nl/archives/sitearchief/20210408123129/www.rijksoverheid.nl/documenten/kamerstukken/2019/03/18/kamerbrief-over-plan-om-aanpak-laaggeletterdheid-2020-%E2%80%93-2024

Ministerie van Onderwijs, Cultuur en Wetenschap. (2023, April 18). *Tussentijdse rapportage beleidsmonitor TMMT 2020-2024.* Rapport | Tel Mee Met Taal. https://www.telmeemettaal.nl/documenten/rapporten/2023/04/18/tussentijdse-rapportage-beleidsmonitor-tmmt-2020-2024

Morrison, C. M., & Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of experimental psychology: learning, Memory, and cognition, 21*(1), 116.

Nathan, L., Stackhouse, J., Goulandris, N., & Snowling, M. J. (2004). The development of early literacy skills among children with speech difficulties. *Journal of Speech, Language, and Hearing Research, 47*(1), 33-57.

Plourde, C. E., & Besner, D. (1997). On the locus of the word frequency effect in visual word recognition. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 51*(3), 181

Rastle, K. (2016). Visual word recognition. In *Neurobiology of language* (pp. 255-264). Academic Press.

Realistic Text to Speech converter & AI Voice generator. (n.d.). SpeechGen.io. https://speechgen.io/

Schaerlaekens, A. (2008). *De taalontwikkeling van het kind.* Groningen/Houten: Wolters-Noordhoff.

Schleicher, A. (2019). *PISA 2018: Insights and Interpretations.* oecd Publishing. https://www.oecd.org/pisa/PISA%202018%20Insights%20and%20Interpretations%20FINAL%20PDF.pdf

Seidenberg, M. S. (2013). The science of reading and its educational implications. *Language learning and development, 9*(4), 331-360.

Smaling, A. (2014). Steekproeven voor generalisatie. *KWALON, 19*(1).

Stadthagen, H. (2005). *Age of Acquisition and Frequency Effects in Visual Word Recognition* (Doctoral dissertation, University of Bristol).

Stadthagen-Gonzalez, H., Bowers, J. S., & Damian, M. F. (2004). Age-of-acquisition effects in visual word recognition: Evidence from expert vocabularies. *Cognition, 93*(1), B11-B26.

Stanovich, K. E. (1982). Individual differences in the cognitive processes of reading: II. Text-level processes. *Journal of Learning Disabilities, 15*(9), 549-554.

Tepperman, J., Lee, S., Narayanan, S., & Alwan, A. (2010). A generative student model for scoring word reading skills. *IEEE Transactions on Audio, Speech, and Language Processing, 19*(2), 348-360.

Van Wiechenonderzoek - NCJ. (2023, March 3). NCJ. https://www.ncj.nl/onderwerp/van-wiechenonderzoek/

Varol, H. A., Mani, S., Compton, D. L., Fuchs, L. S., & Fuchs, D. (2009). Early prediction of reading disability using machine learning. In *AMIA annual symposium proceedings* (Vol. 2009, p. 667). American Medical Informatics Association.

Verkooijen, N, & Wildenbos, J. (2020, May 27th). *Nu is het juiste moment om laaggeletterden te helpen met hun taalvaardigheid.* Sociale Vraagstukken. https://www.socialevraagstukken.nl/nu-is-het-juiste-moment-om-laaggeletterden-te-helpen-met-hun-taalvaardigheid/

Weir, C., & Khalifa, H. (2008). A cognitive processing approach towards defining reading comprehension. *Cambridge ESOL: Research Notes, 31,* 2-10.

Weiss, B. D., Hart, G., McGee, D. L., & D'Estelle, S. (1992). Health status of illiterate adults: relation between literacy and health status among persons with low literacy skills. *The Journal of the American Board of Family Practice, 5*(3), 257-264.

Wolf, M. S., Williams, M. V., Parker, R. M., Parikh, N. S., Nowlan, A. W., & Baker, D. W. (2007). Patients' shame and attitudes toward discussing the results of literacy screening. *Journal of health communication, 12*(8), 721-732.

Zink, I. & Smessaert, H. (2009). *Taalontwikkeling Stap voor Stap.* Herentals: Vlaamse Vereniging voor Logopedie.

Appendix A

Used words and conditions in the experiment

Words from the wordlist of Ghyselinck, De Moor & Brysbaert (2000) are used for the experiment. The first ten words per condition were presented auditorily, the ten following words were the minimal pairs belonging to the previous words.

Condition 1

Ander, mens, jaar, twee, hand, kind, drie, tijd, vrouw, niets

Anker, mans, maar, thee, hard, kond, brie, rijd, brouw, fiets.

Condition 2

Rede, norm, facet, notie, leek, sekte, visie, debat, rente, omzet

Zede, vorm, lacet, motie, leuk, sekse, visje, debet, lente, omzit

Condition 3

Woef, smurf, sapje, nies, lego, zwiep, buggy, stamp, snurk, slab

Zoef, slurf, mapje, nier, logo, zwier, baggy, stomp, snerk, slak

Conditie 4

Puts, zeeg, keen, vamp, laar, puls, balg, guts, slenk, zijp

Puls, zeef, kern, ramp, raar, puts, belg, muts, slank, pijp

The conditions contained words with the following AoA and frequency distribution in the experiment:

Condition 1:

- High frequency, early AoA, accept
- High frequency, early AoA, reject

Condition 2:

- High frequency, late AoA, accept
- High frequency, late AoA, reject

Condition 3:

- Low frequency, early AoA, accept

- Low frequency, early AoA, reject

Condition 4:

- Low frequency, late AoA, accept
- Low frequency, late AoA, reject

Whereas in the experiment conditions 1 and 2, 3 and 4, 5 and 6, 7 and 8 were combined because the computer screen showed the reject condition, which are the conditions 2, 4, 6 and 8 for the words in the other conditions.

Appendix B

Calculation of values condition 2

Standard deviations used for condition 2 (late AoA, high frequency). To belong to the condition the words should have an AoA of 12,20935 or higher, and a frequency of 2,916024 or higher. Since there were no words to be found in this condition, a standard deviation of .057 is used. For these words it is shown what the values regarding AoA and/or frequency are.

Facet: frequency 2.62

Notie: frequency 2.53

Sekse: AoA 12

Visie: AoA 12

Debat: frequency 2.80

Omzet: frequency 2.49

rente: AoA 12

leek: AoA 11.9

Appendix C

Informed consent and information text

Welkom bij dit experiment, waarbij wij antwoord willen geven op de vraag;

‘In welke mate is de prestatie op een visuele woordherkenning taak afhankelijk van de leeftijd van verwerving en frequentie van gebruikte woorden?’

Hierbij wordt de reactietijd en accuratesse gemeten na het horen van de auditieve stimulus. Als participant wordt van je verwacht dat je het zojuist gehoorde woord uitspreekt, en daarna aangeeft of het uitgesproken woord hetzelfde is als dat wordt getoond op de computer. Hiermee kan de accuratesse van de spraakherkenningssoftware worden bepaald.
Als je klaar bent met het experiment, zal je de compensatie ontvangen.

Er zijn geen potentiële risico's bij deelname aan deze taak en het onderzoeksproject is goedgekeurd door de ethische commissie van de Gedragswetenschappen (referentienummer 23-0134). Dit onderzoek is een project dat wordt geadministreerd door de Universiteit Utrecht. De gegevens die in dit onderzoek worden verzameld, zijn alleen bedoeld voor wetenschappelijk onderzoek. Gegevens worden beheerd door de Universiteit Utrecht volgens de Algemene Verordening Gegevensbescherming (AVG). Anonieme gegevens van dit onderzoek kunnen voor onderzoeksdoeleinden worden gedeeld in een openbare opslagplaats en worden gepresenteerd in wetenschappelijke publicaties. Je hebt het recht om het onderzoek op elk moment te beëindigen en je kunt ons binnen 24 uur na indiening informeren om de gegevens te verwijderen.

Deelname aan het onderzoek kan op elk moment worden beëindigd, zonder enige uitleg en zonder enige negatieve gevolgen. Voor vragen of opmerkingen over de studie kun je contact opnemen met Fleur Gramkow **(f.g.d.gramkow@students.uu.nl)**. Als je een officiële klacht hebt over het onderzoek, kun je een e-mail sturen naar de klachtenfunctionaris via **klachtenfunctionaris-fetchsocwet@uu.nl**. Indien je bereid bent deel te nemen aan dit onderzoek, vragen wij voorafgaand aan het onderzoek de ‘informed consent’ op de volgende pagina te lezen, en als je het ermee eens bent, op 'ik ga akkoord' te klikken.

Toestemmingsformulier

Mijn deelname aan dit onderzoek is geheel vrijwillig. Ik weet dat ik op elk moment kan besluiten om me terug te trekken uit dit onderzoek.

De procedure omtrent geheimhouding is mij duidelijk uitgelegd. Ik begrijp dat alle gegevens anoniem zijn en dat er geen persoonlijke informatie wordt opgeslagen samen met de gegevens die in dit onderzoek zijn verzameld. Ik weet dat mijn geanonimiseerde gegevens kunnen voorkomen in mogelijke rapporten, publicaties of presentaties die voortkomen uit dit onderzoek.

____

Appendix D

Instructions and information screens during the experiment

Het experiment zal ongeveer 30 minuten duren. Je krijgt een koptelefoon voor de auditieve stimuli en de microfoon neemt de woorden op die je uitspreekt. Als het experiment voltooid is, vragen wij om nog enkele demografische gegevens in te vullen en een aantal vragen over het experiment te beantwoorden.

Het experiment bestaat in totaal uit vier sets met 20 woorden, voor elke set word je gevraagd om:

1. Naar het woord te luisteren
2. Het woord dat je hebt gehoord uit te spreken
3. **Zo snel en accuraat mogelijk** (!) aan te geven of het uitgesproken woord hetzelfde is als op het scherm wordt getoond

Als je denkt dat uitgesproken woord hetzelfde is als op het scherm, gebruik je de X-toets.
Als je denkt dat het uitgesproken woord anders is dan op het scherm, gebruik je de N-toets.

Tussen elke set van 20 woorden heb je de kans om een korte pauze te houden. Als je op 'volgende' drukt, volgt een tutorial (inclusief een audio test van het programma; de audio test is in het Engels). Daar kun je twee keer oefenen, voordat het experiment daadwerkelijk begint.

Nogmaals, als je denkt dat het gesproken woord **hetzelfde** is als op het scherm, druk dan op **X**.

Als je denkt dat het gesproken woord **verschilt** van wat er op het scherm staat, druk dan op **N.**

Hier volgt de tutorial.

Hier leer je de volgorde van het experiment en kun je oefenen met twee woorden.

Tijdens het experiment ga je naar het volgende scherm door op de spatiebalk te drukken. In sommige gevallen hoef je de spatiebalk niet te gebruiken om naar het volgende scherm te gaan, dat gebeurt dan automatisch.

Plaats de wijs- of middelvingers zoals aangegeven op de foto en houd ze zo gedurende het hele experiment.

De linker vinger plaats je op de X en de rechter vinger op de N. De spatiebalk kan met de duim worden ingedrukt.

Dus vanaf nu, gebruik je de **spatiebalk** om naar het volgende scherm te gaan.


Dit is je cue om naar het scherm te blijven kijken, kort hierna volgt een ander scherm


Op dit scherm hoor je het woord


Direct daarna zie je dit scherm en hoor je een 'piep'. Dit is je cue om het woord uit te spreken. De balk vult zich om het tijdslimiet aan te geven.

Vervolgens wordt, na een fixatiekruis, het woord getoond dat gecontroleerd moet worden, met de toetsen X en N. Antwoord **zo snel en zo nauwkeurig mogelijk** of dit het woord is dat je zojuist hoorde en hebt uitgesproken. Er is een tijdslimiet van 3 seconden.

Nu de instructies zijn gegeven, ga je dit twee keer oefenen.

Dit is het einde van de tutorial.

Als je op de spatiebalk drukt, kom je bij het beginscherm van het experiment.

Het experiment wordt gestart als je op de spatiebalk drukt.

Dit is het einde van de eerste ronde. Nog drie te gaan.
Je kunt een korte pauze houden.

Druk op de spatiebalk om het experiment te hervatten.

Houd je vingers gereed op de **X (hetzelfde)** en **N (verschillend)** toetsen, om **zo snel en accuraat mogelijk** te antwoorden.
Als je klaar bent, druk op de spatiebalk.

Dit is het einde van de tweede ronde. Nog twee te gaan.
Je kunt een korte pauze houden.

Druk op de spatiebalk om het experiment te hervatten.

Dit is het einde van de derde ronde. Nog één te gaan.
Je kunt een korte pauze houden.

Druk op de spatiebalk om het experiment te hervatten.

Je bent nu klaar met het testen van de spraak naar tekst software. Hierna volgen nog enkele vragen.

Dit is het einde van het experiment.
Heel erg bedankt voor je deelname.

De instructeur zal buiten op je wachten.
Klik op 'volgende' voordat je vertrekt. Daarna kun je de ruimte verlaten.

Appendix E

Conducted questionnaire

Heb je **alle** woorden luid en duidelijk verstaan?

Ja

Nee

Welk woord gebruik je **vaker**?

stamp

hand

Welk woord gebruik je **vaker?**

omzet

puls

Welk woord ken je **langer?**

norm

thee

Welk woord ken je **langer**?

slenk

lego

Kende je de woorden die zijn gebruikt in het experiment?

smurf - ja nee

balg - ja nee

notie - ja nee

tijd - ja nee

sapje - ja nee

jaar - ja nee

leek - ja nee

vamp - ja nee

slenk - ja nee

niets - ja nee

guts - ja nee

woef - ja nee

debat - ja nee

mens - ja nee

buggy - ja nee

slab - ja nee

facet - ja nee

rede - ja nee

stamp - ja nee

puts - ja nee

visie - ja nee

hand - ja nee

keen - ja nee

rente - ja nee

twee - ja nee

zijp - ja nee

omzet - ja nee

zeeg - ja nee

snurk - ja nee

zwiep - ja nee

drie - ja nee

kind - ja nee

puls - ja nee

lego - ja nee

vrouw - ja nee

nies - ja nee

ander - ja nee

sekte - ja nee

norm - ja nee

laar - ja nee

Hoe vaak lees je een boek, artikel of nieuwsbericht?

nooit - alleen als het moet - eenmaal per maand - een paar keer per maand - eenmaal per week - twee tot drie keer per week - dagelijks

In hoeverre ben je geïnteresseerd geweest in lezen sinds je kunt lezen?
geen interesse - veel interesse

In hoeverre geniet je van lezen op dit moment in je leven?
helemaal niet - heel erg

Beantwoord alsjeblieft enkele vragen over jezelf:

Met welk geslacht identificeer je je?
divers
mannelijk
vrouwelijk

Hoe oud ben je?

Wat is je huidige opleidingsniveau?
HBO
MBO
Bachelor
Master
andere

Wat is je dominante hand?
rechts
links

Is Nederlands je moedertaal?
ja
nee

Ben je tweetalig opgevoed?

ja

nee