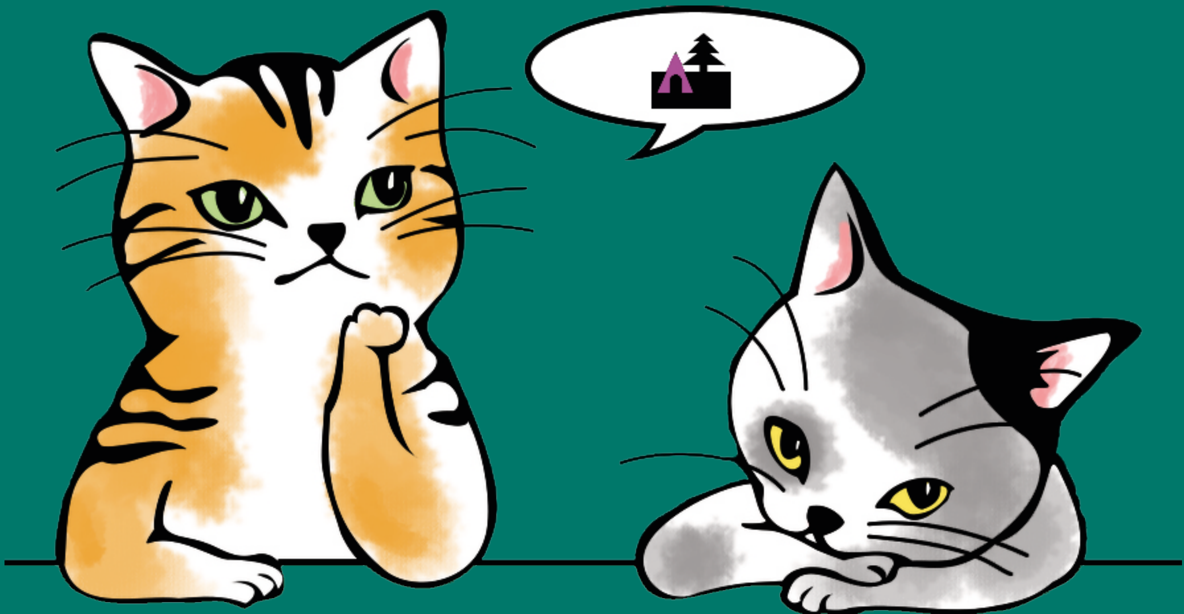
MAX PLANCK INSTITUTE
FOR PSYCHOLINGUISTICS



Spoken and written word processing:

Effects of presentation modality and individual differences in experience to written language

MEREL WOLF

Spoken and written word processing:
Effects of presentation modality and individual differences in experience to written language

Funding body
This research was funded by the Max Planck Society for the Advancement of Science (www.mpg.de/en) in the form of an IMPRS for Language Sciences PhD Fellowship (2017-2021) awarded to Merel Wolf.

International Max Planck Research School (IMPRS) for Language Sciences
The educational component of the doctoral training was provided by the International Max Planck Research School (IMPRS) for Language Sciences. The graduate school is a joint initiative between the Max Planck Institute for Psycholinguistics and two partner institutes at Radboud University – the Centre for Language Studies, and the Donders Institute for Brain, Cognition and Behaviour. The IMPRS curriculum, which is funded by the Max Planck Society for the Advancement of Science, ensures that each member receives interdisciplinary training in the language sciences and develops a well-rounded skill set in preparation for fulfilling careers in academia and beyond. More information can be found at www.mpi.nl/imprs

The MPI series in Psycholinguistics
Initiated in 1997, the MPI series in Psycholinguistics contains doctoral theses produced at the Max Planck Institute for Psycholinguistics. Since 2013, it includes theses produced by members of the IMPRS for Language Sciences. The current listing is available at www.mpi.nl/mpi-series

ISBN: 978-94-92910-31-8
Cover design by Merel Wolf, cat lineart by *Naobim* on Pixabay
Printed and bound by Ipskamp Drukkers, Enschede

 The research reported in this thesis was conducted at the Max Planck Institute for Psycholinguistics, in Nijmegen, the Netherlands

Spoken and written word processing: Effects of presentation modality and individual differences in experience to written language

Proefschrift ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college voor promoties
in het openbaar te verdedigen op

maandag 14 februari 2022
om 10.30 uur precies

door

Merel Cecile Wolf
geboren op 13 juli 1994
te Amsterdam (Nederland)

Promotoren:

Prof. dr. A.S. Meyer

Prof. dr. C.F. Rowland

Copromotoren:

Dr. A.C. Smith (Max Planck Instituut voor Psycholinguïstiek)

Dr. F. Hintz (Max Planck Instituut voor Psycholinguïstiek)

Manuscriptcommissie:

Prof. dr. P.C.J. Segers

Prof. dr. M. Brysbaert (Universiteit Gent, België)

Dr. R.M. Willems

Voor mijn familie

Contents

1 General introduction

The invention of writing 5500 years ago can be regarded as one of the most crucial inventions of the human species, allowing complex civilizations to evolve and develop. Before the existence of a writing system, human communication was more constrained, particularly by distance and time[1]. For speech, one of the primary modalities of language, to be an efficient medium of communication, the listener must be in close proximity to the speaker, as sound waves produced by human voices cannot travel long distances. Moreover, speech is bound by time, such that, once the speaker stops talking, the sound waves of the utterance cease to exist. If a listener stands in close proximity to the speaker, but an hour after the speaker has produced their message, the message will be long gone and irretrievable.

Written language, on the other hand, is not constrained by distance or time. Written information allows for communication between individuals that cannot physically meet each other. As a result, new ideas and inventions can spread easily across the globe. Moreover, by writing information down, thoughts can be immortalized. Writers can convey a message to someone that may not even be alive at the same time and ensure that things that should not be forgotten will not be lost. Thus, the invention of writing resulted in a new modality in which language could be expressed; one that was not bound to the same constraints as other, already existing language modalities, such as speech.

1.1 Processing linguistic input in different modalities

This new modality in which language could be encountered required a distinct processing pipeline. Even though the message that is conveyed may be the same across spoken and written modalities, the human brain must process the input initially through modality-specific processing systems. Indeed, models of spoken

[1]Note that with the inventions of recording technology and virtual communication systems these constraints have been reduced.

and written language comprehension assume that linguistic input is first processed acoustically (Marslen-Wilson & Welsh, 1978; McClelland & Elman, 1986; Norris, 1994) or visually (Grainger & Ferrand, 1994; Grainger & Ziegler, 2011; McClelland & Rumelhart, 1981), depending on the modality in which the input is encountered. The output from these initially modality-specific perceptual processing systems is fed forward to a processing system specific to language.

This language processing system is assumed to be abstract rather than perceptual in nature. Many psycholinguistic theories assume that words, for example, are stored in a mental lexicon (although there are alternative accounts of how language is mentally represented (see Elman, 2004)). This lexicon contains an entry for every word an individual knows. Each lexical entry within this lexicon consists of different types of representational information. These types of representations encompass semantic (meaning), phonological (sound), orthographic (spelling) and syntactic (grammatical) information. The information obtained from the acoustic or visual processing system feeds forward to, respectively, the phonological and orthographical (sublexical) representations of lexical entries stored in the mental lexicon. Eventually, through a process of lexical competition, the lexical entry that most closely matches the modality-specific perceptual input is activated. Moreover, as it is assumed that there are strong connections between the different types of representations (semantics, phonology, orthography, syntax) within a lexical entry, other representational information of the perceived word is assumed to be partially activated as well (Grainger & Ferrand, 1994; McClelland & Rumelhart, 1981; Seidenberg & McClelland, 1989).

A crucial hypothesis within the framework of the mental lexicon is that the different types of representations (semantic, phonology, orthography, syntax) within a single lexical entry can differ in their level of precision and completeness (Lexical Quality Hypothesis, Perfetti, 2007). Thus, different representations of the same word can differ in their 'lexical quality'. For example, the phonological representation of a word may be precisely defined in the mental lexicon, but the orthographic representation of the same word may be less precise or may not exist at all, if the individual cannot read. A second crucial hypothesis is that processing lexical representations that are of higher quality (i.e., more complete and precisely defined) is more efficient than processing low quality lexical representations (Perfetti, 2007). These two hypotheses give rise to certain questions: if modality-specific representations (phonological vs. orthographic) of the same lexical entry can differ in their quality, could it be possible that processing a word presented in one modality is more efficient than processing the same word in the other modality? Moreover, how and when in the process of learning a new word might differences start to occur in the quality of different types of representations of the same word?

A small number of studies have provided some evidence that suggests that familiar words are more efficiently (i.e., more accurately and faster) recognized when presented in the written than the spoken modality (Connine, Mullennix, Shernoff, & Yelen, 1990; Lopez Zunini, Baart, Samuel, & Armstrong, 2020; Turner, Valentine, & Ellis, 1998). With regard to learning new words, evidence is inconclusive, with some studies finding that words are more efficiently learned when presented in the spoken modality (Bakker, Takashima, van Hell, Janzen, & McQueen, 2014), others finding a learning benefit in the written modality (Balass, Nelson, & Perfetti, 2010; van der Ven, Takashima, Segers, & Verhoeven, 2015), and yet more finding no difference in learning efficiency between the two modalities (Dean, Yekovich, & Gray, 1988; Nelson, Balass, & Perfetti, 2005). Due to these limited and sometimes inconclusive findings, the influence of presentation modality on word processing is not yet fully understood.

1.2 Literacy influences linguistic and cognitive systems

In addition to establishing a distinct processing pipeline and representational type (orthography), the written modality brought more global changes to humans' linguistic and cognitive systems (Kolinsky, 2015; Morais & Kolinsky, 2002). There is substantial evidence that the intrinsic organization of the language faculty changes as people gain the ability to read. Alphabetic languages are characterized by strong connections between sounds (phonemes) and symbols (letters/graphemes), and it has been observed that literate individuals of alphabetic languages are better able to distinguish the different phonemes of a word than illiterate individuals (Morais, Bertelson, Cary, & Alegria, 1986; Morais, Cary, Alegria, & Bertelson, 1979). This indicates that literacy acquisition influences how language is mentally represented. Literacy acquisition also induces qualitative changes to visual processing. Scanning habits of both linguistic and non-linguistic visual stimuli are adapted to the writing direction of a language's script (Bramao et al., 2007), as are temporal (early vs. later) and spatial (small vs. large) associations (Bergen & Chan Lau, 2012; Dehaene, Bossini, & Giraux, 1993; Dobel, Enriquez-Geppert, Zwitserlood, & Bölte, 2014). Additionally, literacy acquisition in a language that uses a script with mirrored symbols (b vs. d, p vs. q) increases one's sensitivity to mirroring of non-linguistic stimuli. Illiterate individuals tend to categorize mirrored stimuli (e.g., ↙ and ↘) as "same", whereas literate individuals would place them in different categories (Kolinsky et al., 2011; Pegado et al., 2014).

In addition, there is evidence suggesting that structural changes in the human brain are induced as the result of acquiring literacy (Dehaene, Cohen, Morais, & Kolinsky, 2015; Horowitz-Kraus & Hutton, 2015). The acquisition of literacy seems to increase connectivity between brain hemispheres (Carreiras et al., 2009; Castro-Caldas et al., 1999; Petersson, Silva, Castro-Caldas, Ingvar, & Reis, 2007). Moreover, connectivity between areas *within* a hemisphere also tends to strengthen with literacy acquisition, as reflected in increased connectivity between areas known to be related to visual processing and areas related to processing speech in literates (Thiebaut de Schotten, Cohen, Amemiya, Braga, & Dehaene, 2014). This increased connectivity may be structural evidence for grapheme-phoneme conversion processes, which are crucial in reading and reading development. Not only does literacy acquisition lead to heightened connectivity within the brain, it also induces specialization in the brain. Perhaps most notably is the specialization of an area in the mid-portion of the left fusiform gyrus: this Visual Word Form Area is thought to be particularly specialized in processing scripts (L. Cohen & Dehaene, 2004; L. Cohen et al., 2000). Remarkably, this area shows an increase in response to script as people become literate (Dehaene et al., 2010).

Thus, the acquisition of literacy induces both functional and structural changes to humans' cognitive and linguistic systems, which allows them to process information presented in the written modality. It is striking that, since the invention of writing is only relatively recent within the timescale of evolution, the human brain has not had time to 'evolve' to process language in this modality. Rather, the brain seems to reshape itself during development (i.e., within the timescale of a human lifespan) in order to support written language processing. Crucially, the studies described above tend to show that the degree to which these structural and functional changes occur is related to the amount of experience individuals have had with written information, by comparing individuals that differ in their degree of literacy (early vs. late literate vs. illiterate individuals). Thus, an individual's experience with written materials influences the degree to which their linguistic and cognitive systems are functionally and structurally changed. This raises the question how experience with written materials influences the processing of linguistic information.

1.3 Individual differences in (written) language experience

In general, individuals tend to differ substantially with respect to their experience with language. Some individuals have more opportunities to immerse themselves in linguistically rich environments than others. As a result of differences in exposure to linguistic input and variation in individuals' intrinsic capacity for learning, people develop to differ substantially in their language-related skills (Dąbrowska, 2018; Kidd et al., 2020; Kidd, Donnelly, & Christiansen, 2018; Lieven, 2016). For various domains of language-related skills, evidence has been found for a performance benefit in individuals that have had more experience with language. For example, language experienced individuals show a performance benefit in the domain of word comprehension (Andringa, Olsthoorn, van Beuningen, Schoonen, & Hulstijn, 2012; Bent, Baese-Berk, Borrie, & McKee, 2016; Brysbaert, Lagrou, & Stevens, 2016; Diependaele, Lemhöfer, & Brysbaert, 2013; Federmeier, McLennan, De Ochoa, & Kutas, 2002; Janse & Jesse, 2014; Mainz, Shao, Brysbaert, & Meyer, 2017; Mani & Huettig, 2014; Yap, Balota, Sibley, & Ratcliff, 2012). Similar findings have been observed in the domains of word production (Rodriguez-Aranda & Jakobsen, 2011; Shao, Janse, Visser, & Meyer, 2014; N. Unsworth, Spillers, & Brewer, 2011; Yap et al., 2012) and vocabulary size (Hurtado, Marchman, & Fernald, 2008; Monaghan, Chang, Welbourne, & Brysbaert, 2017).

Why does language experience improve language abilities? There are two accounts, the lexical entrenchment hypothesis (Brysbaert, Lagrou, & Stevens, 2016; Diependaele et al., 2013) and lexical tuning hypothesis (Castles, 1999; Castles, Davis, Cavalot, & Forster, 2007), that suggest that language experience improves the quality of lexical representations. As mentioned before, it is thought that high quality lexical representations are processed more efficiently than low quality lexical representations (Perfetti, 2007). In a mental lexicon with high quality representations, representations will be activated faster, because there is less lexical competition of neighbouring lexical entries (Andrews, 1997; Andrews & Hersch, 2010; Perfetti, 1992), since these neighbouring lexical entries are also precisely defined. As a result of this improved processing efficiency, language experienced individuals perform better on tests of language ability.

Individuals also vary with respect to experience with *written* language. This experience with written language is also termed 'literacy' or 'reading experience' or 'print exposure' and these terms are used interchangeably in this dissertation. Some adults are very avid readers, whereas others tend to avoid reading altogether (Smith, 1996). These contrasts are also already visible in primary students (Watkins & Edwards, 1992) and secondary students (Dood, Gubbels, & Segers, 2020; Pfost, Dörfler, & Artelt, 2013). Individual differences in experience with written materials are reflected in variation in linguistic skills. Written language experienced individuals have better reading skills (Mol & Bus, 2011), more syntactic knowledge (Montag & MacDonald, 2015), improved syntactic processing skills (Favier & Huettig, 2021; Street & Dąbrowska, 2010) and are better at predicting upcoming language in sentence contexts (Favier, Meyer, & Huettig, 2020; Huettig & Pickering, 2019).

The beneficial effect of language experience on language abilities may be amplified in the written modality, because of several characteristics that distinguish written from spoken language. In general, written language is more structured, complex and more varied than spoken language[2]. Written language is known to be much more structured than spoken language, as for example subordination is 60% more common in written compared to spoken language (Kroll, 1977). Moreover, written text is characterized by more complex syntactic structures, such as passives, object clefts and participial phrases (Biber, 1991; Roland, Dick, & Elman, 2007; Scott, 2008). Finally, written language is composed of a more varied vocabulary than spoken language, as it often contains more difficult, uncommon words (Cunningham & Stanovich, 1998; Hayes & Ahrens, 1988). Thus, due to text being more structured, complex, and varied, extensive experience with written language may be particular beneficial for improving linguistic abilities, as it provides a very rich context that helps specifying lexical representations. In addition, the written modality allows for relatively more language input than spoken modality. Written input is processed faster than spoken input, since written language units are presented simultaneously whereas spoken language units unfold over time. This allows for relatively more language exposure within a given time frame in the written compared to spoken modality (Brysbaert, Stevens, Mandera, & Keuleers, 2016).

The relationship between experience with written language and language ability in several linguistic domains (reading ability, syntactic processing and lan-

[2]"Spoken language" refers here to day-to-day conversations. Some spoken language can be more structured, complex and varied, for example lectures or speeches. Moreover, sometimes text originally composed in the written modality (e.g., audio-books) is transposed to the spoken format, in which case the characteristics of the written language are retained and expressed in the spoken modality.

guage prediction) has been demonstrated, but it is unclear whether experience with written language also influences non-predictive, single word processing. The aforementioned characteristics of written language, particularly the diverse vocabulary and written language allowing for relatively more exposure within a given time frame, may provide a rich substrate to sharpen lexical representations and consequently increase word processing efficiency.

1.4 Dissertation outline

To sum up, the invention of writing created a new modality in which language could be processed. This new modality came with a distinct processing pipeline and new type of representation. Different types of representations of the same lexical entry can differ in their level of precision and completeness (i.e., lexical quality). Moreover, higher quality representations are assumed to result in more efficient processing. This raises the question: if the orthographic and phonological representation of the same word can differ in their quality, would this result in a difference in processing efficiency of the spoken and written form of the same word? In other words, are words presented in the written modality recognized faster and more accurately, or, in the case of novel words, learned more efficiently than words presented in the spoken modality? This dissertation explored this question in several chapters by examining how modality of presentation influences word processing.

Furthermore, language experience is thought to sharpen lexical representations and thereby improve language processing and linguistic ability, and this benefit may be particularly strong in the written modality. This raises the question whether individuals who are more experienced with written language show more efficient word processing than individuals who have had less experience with written language. This dissertation examined this question by looking at how individual differences in written language experience and abilities related to written language processing influence the efficiency with which words are processed.

Chapter 2 describes a re-analysis of a dataset originally obtained for the development of a receptive vocabulary test, and investigates the effects of presentation modality and experience with written language (termed "reading experience" in this chapter) on word recognition. Participants, all university students, carried out a lexical decision task where words were presented in the written, spoken or audio-visual form. Their experience with written language was approximated using a receptive vocabulary measure and a measure indicative of their knowledge of authors, which is thought to be a reliable estimation of an individual's experience with written language. The words differed in difficulty, with some words being more uncommon and some words being commonly known, to see whether the effects of presentation modality and experience with written language were more pronounced for well-known or more difficult words.

Chapter 3 presents an experimental study that used a paradigm similar to the paradigm used in Chapter 2, but with several crucial differences. Most importantly, participants that varied in their educational background were recruited with the aim of diversifying the sample with respect to experience with written language (termed "reading experience" in this chapter). The word difficulty range of the stimuli words was adapted to match this more diverse sample. Moreover, an additional measure of experience with written language was administered in the form of a reading behaviour questionnaire. Finally, the lexical decision task was timed to investigate not only how presentation modality and experience with written language influences the accuracy but also the speed with which words are recognized.

Chapter 4 describes the analysis of an open-access dataset and explored the influence of written language experience and skills obtained through extensive experience with written language (termed "literacy" in this chapter) on spoken language processing, while accounting for the influence of general cognitive skills (nonverbal intelligence and processing speed). Two spoken language processing domains were examined, namely word production and word comprehension. The publicly available dataset contained the data of participants with diverse educational backgrounds on multiple tasks measuring word production, word comprehension, literacy and processing speed, and a single measure of nonverbal intelligence (nonverbal IQ). First, the relationships between the tasks that were designed to measure the same skill was assessed. Using a latent-variable approach, a single factor score was obtained for each skill. Subsequent regression analyses examined the extent to which experience and skills related to written language explained variance in word production and word comprehension when accounting for general cognitive skills.

Chapter 5 is a systematic review that describes the current literature with regard to the relationship between written language experience (termed "literacy experience" in this chapter) and spoken and written word recognition. This relationship was explored separately for different types of lexical representations (semantic, phonology, orthography, syntax) and for different levels of representation (lexical, morpheme/syllable, sublexical). Studies conducted with adults as well as children are discussed in order to describe developmental changes in spoken and written word recognition that relate to the acquisition of literacy. The chapter offers a comprehensive summary of the current state of the literature on this topic, identifies gaps and formulates recommendations for future research. Although it may be expected to find a systematic review at the beginning of a dissertation, this review was written after the other chapters on word recognition as the COVID-19 pandemic impeded empirical work. It is therefore placed after these chapters.

Chapter 6 presents three experiments that examined the influence of presentation modality, skills related to experience with written language (receptive vocabulary, word and nonword reading) and nonverbal intelligence on novel word learning. Using a new, implicit, fast-paced word learning paradigm with a between-subjects design, participants, who were all university students, learned 24 Dutch-like pseudowords that were associated with the picture of a non-existing object. The words were either presented in the written or spoken modality. After a short period of consolidation (20 minutes), during which participants performed a non-linguistic task, their knowledge of the novel words was assessed using a matching task, in which they had to decide whether a pair of a word and picture matched according to what they had learned earlier in the training phase. As in the learning phase, the words were either presented in the written or spoken form, to investigate whether the participants did not only recognize the words in the modality in which they had been trained, but also whether this new knowledge transferred to the other modality in which they had not yet seen the words before.

Finally, Chapter 7 summarizes the findings, discusses their broader theoretical implications and provides recommendations for future endeavours within this research field.

2 The effects of input modality, word difficulty and reading experience on word recognition accuracy[1]

Abstract

Language users encounter words in at least two different modalities. Arguably, the most frequent encounters are in spoken or written form. Previous research has shown that —- compared to the spoken modality —- written language features more difficult words. An important question is whether input modality has effects on word recognition accuracy. In the present study, we investigated whether input modality (spoken, written, or bimodal) affected word recognition accuracy and whether such a modality effect interacted with word difficulty. Moreover, we tested whether the participants' reading experience interacted with word difficulty and whether this interaction was influenced by modality. We re-analysed data from 48 Dutch university students that were collected in the context of a vocabulary test development to assess in which modality test words should be presented. Participants carried out a word recognition task, where nonwords and words of varying difficulty were presented in spoken, written and audio-visual modalities. In addition, they completed a receptive vocabulary and an author recognition test to measure their exposure to literary texts. Our re-analyses showed that word difficulty interacted with reading experience in that frequent readers (i.e., with more exposure to written texts) were more accurate in recognizing difficult words than individuals who read less frequently. However, there was no evidence for an effect of input modality on word recognition accuracy, nor for interactions with word difficulty or reading experience. Thus, in our study, input modality did not influence word recognition accuracy. We discuss the implications of this finding and describe possibilities for future research involving other groups of participants and/or different languages.

[1]Adapted from Wolf, M. C., Meyer, A. S., Rowland, C. F., & Hintz, F. (2021). The effects of input modality, word difficulty and individual differences in reading experience on word recognition accuracy. *Collabra: Psychology. 7*(1): 24919. doi:10.1525/collabra.24919

2.1 Introduction

With the invention of reading and writing, humans gained the opportunity to use language in the written modality alongside, amongst others, the spoken form. This has an important consequence for the internal representational system of language: two representations (orthographic and phonological) of the same lexical item are stored. As a result of the quality and quantity of modality-specific encounters, these two representations can vary in their level of precision and completeness (i.e., lexical quality, Perfetti, 2007). Moreover, written language differs from spoken language in that written text has been shown to include a larger variety of words than speech does (Cunningham & Stanovich, 1998; Hayes & Ahrens, 1988). Consequently, the mental lexicon of frequent readers probably includes more difficult (i.e., less well-known) words than that of individuals who read less.

The fact that difficult words are encountered most often in the written modality (Cunningham & Stanovich, 1998; Hayes & Ahrens, 1988) is likely to have important consequences for the quality of their orthographic and phonological representations, which in turn may influence word recognition. That is, assuming that difficult words are more often read than heard, accessing word meaning through the written representation may be less error prone compared to hearing the same words. Our current understanding of how word recognition accuracy is affected by the modality in which words are presented is limited. Moreover, it is unclear whether any modality effects on word recognition would be moderated by the words' difficulty and/or individuals' reading experience. Demonstrating effects of input modality on word recognition would have important implications for tools measuring receptive vocabulary size through tests of word recognition. That is, if word recognition accuracy were to differ as a function of modality, researchers developing tests of word recognition would need to consider carefully in which modality to present the test words. If, on the other hand, modality did not show effects on recognition accuracy, presentation modality would only have to play a minor role when designing new tests.

In the present study, we addressed these questions by re-analyzing a dataset that was collected in the context of developing a Dutch receptive vocabulary test (capitalizing on word recognition ability). Specifically, participants in that experiment had carried out a lexical decision task. They responded to words, ranging substantially in difficulty, presented in three modalities (spoken, written, or bimodal). The goal was to assess in which modality test words in the receptive vocabulary test should be presented. Moreover, there were two groups of participants who received different instructions ("Is this an existing Dutch word?" vs. "Do you know this Dutch word?") to assess potential task instruction effects on word recognition accuracy. Finally, in addition to the main experiment, participants had completed two tests assessing their receptive vocabulary size and exposure to literary texts, respectively. Thus, given the range of word difficulty, the three modality conditions and the additional individual-differences tests, the dataset was well-suited to address the present research questions centering around modality effects on word recognition accuracy and their potential moderators.

Previous studies have reported word recognition benefits for written and bimodal (simultaneous presentation of orthographic representation and spoken production of the phonological form) modalities compared to the spoken modality using a lexical decision task (Connine et al., 1990; Lopez Zunini et al., 2020; Turner et al., 1998). Responses were found to be faster and more accurate for words presented in the written and bimodal (audio-visual) modalities compared to the spoken modality. Note that these findings do not allow for generalizations on how modality affects word recognition accuracy as lexical decision tasks typically use words with a limited difficulty range such that responses (with reaction time as the main measure of interest) are assumed to index the speed with which a lexical entry is accessed. Difficult words are rarely used in lexical decision tasks (see Goldinger, 1996, for a review).

'Megastudies' in which large numbers of participants are tested (often via the internet) are an exception and have used difficult words in their lexical decision tasks. For example, Ferrand et al. (2018) assessed how much of the variance in word recognition accuracy and lexical decision latencies for written and spoken words was explained by word difficulty, operationalized as word frequency. They reported that in the written modality, 20% of the variance in word recognition accuracy and 45% of the variance in lexical decision latencies was explained by word frequency. These estimates are in line with other reports that focused on the written modality only: studies found that word frequency explained 15% to 49% of the variance in recognition accuracy and 21% to 49% of the variance in lexical decision latencies (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Ferrand et al., 2010; Keuleers, Diependaele, & Brysbaert, 2010; Yap & Balota, 2009). Crucially, in the study by Ferrand et al. (2018), word frequency explained only a relatively small portion of variance in the spoken modality (7% and 13% of variance in recognition accuracy and lexical decision latencies, respectively). The strongest predictor of auditory lexical decision times was spoken word duration. One reason for the strong influence of word frequency on word recognition in the written but not spoken modality could be, as explained above, that written text contains more infrequent words than spoken language (Cunningham & Stanovich, 1998; Hayes & Ahrens, 1988). Thus, language users are more likely to encounter less frequent words in the written rather than the spoken modality.

Individuals differ substantially in the number and types of words they know (Mainz et al., 2017) and how often they engage in leisure reading (Gallik, 1999; Wift & Ander, 2017). It is likely that differences in receptive vocabulary size and exposure to literary texts influence the interaction between word difficulty and modality on word recognition accuracy[2]. The 'Lexical Quality Hypothesis' (LQH, Perfetti, 2007) holds that word recognition is more efficient, accurate and faster in individuals whose lexical representations are of high quality (Andrews, 2015; Elbro, 1996; Perfetti, 2007, 2011). Such high quality orthographic and phonological representations are precise, fully specified, with strong links between them, allowing for synchronous retrieval. Individuals with much reading experience are assumed to obtain high quality representations through a process called lexical tuning (Castles, 1999; Castles et al., 2007). In order to ensure accurate and fast lexical activation in an ever-expanding mental lexicon, lexical representations become more specific and precise, which improves inhibition of lexical competitors during word recognition (Andrews, 1997; Andrews & Hersch, 2010; Perfetti, 1992). Since the mental lexicon of experienced readers contains more and most likely more difficult words than that of inexperienced, infrequent readers, it is likely that their lexical mental representations are of higher quality, especially in the case of difficult words. Thus, experienced readers are likely to show better word recognition accuracy for difficult words compared to individuals with less reading experience. It is important to highlight that an individual's receptive vocabulary comprises multiple aspects, including one's ability to accurately recognize words in different modalities, as well as in-depth semantic knowledge about words. Though one would think that both are correlated (e.g., a person who recognizes many names of dog breeds might also have more in-depth knowledge about differences of dogs), they are not the same. The present work is concerned with word recognition ability.

[2]Note that, based on previous research, we assume that individuals acquire large receptive vocabularies (especially knowledge about difficult words) through reading, as difficult words appear more often in written than spoken language (Cunningham & Stanovich, 1998). We therefore see both measures (receptive vocabulary size and exposure to literary texts) as reflecting an individual's reading experience.

The present study

By conducting the present re-analysis, we aimed to complement and extend previous reports on modality effects in word recognition. Specifically, we investigated 1) whether input modality had an effect on word recognition accuracy, 2) whether such a modality effect interacted with word difficulty, 3) whether there was an interaction between the effects of word difficulty and reading experience on word recognition accuracy, and 4) whether such an interaction was influenced by input modality.

The present dataset was in many respects similar to previous studies that had investigated modality effects on spoken word recognition: in a within-participants design, Dutch university students were presented with words and nonwords in three modalities (spoken, written, and audio-visual) and were asked to carry out a binary decision task (e.g., lexical decision). However, there were also important methodological differences: as pointed out above, the words participants responded to varied substantially in word difficulty, which led to many more no-responses than in a typical lexical decision experiment. In a typical lexical decision experiment, researchers are predominantly interested in reaction times for words that are recognized correctly (yes-responses), and errors (i.e., no-responses for existing words) are attributed to momentary lapses of attention rather than lack of knowledge of the words. Thus, words are selected from a limited difficulty range to avoid data loss. The present dataset focused on recognition accuracy rather than speed, and, more importantly, modality effects on accuracy, which required the difficulty range to be much larger than in typical lexical decision tasks. That is, participants were presented with words they knew, but also words they did not know or knew less well to avoid ceiling effects.

Relatedly, in contrast to previous studies, word difficulty was approximated using prevalence norms rather than word frequency values. Prevalence norms reflect the degree to which a word is known by the population: the word "apple" is most likely known by 99% of the population, whereas the proportion of people knowing the word "phoneme" is substantially lower. According to Keuleers, Stevens, Mandera, and Brysbaert (2015), prevalence norms provide a more realistic picture of a word’s difficulty than frequency does. This is especially true for low-frequency words. For example, while the word "academia" is probably recognized by the majority of English language users in the US, it rarely occurs in language corpora (i.e., with a frequency of one occurrence per one million words (Brysbaert, New, & Keuleers, 2012)). Keuleers et al. (2015) reported a medium-sized correlation (r = .35) between prevalence and word frequency (based on data from the Dutch Lexicon Project, Keuleers, Diependaele, & Brysbaert, 2010).

A final methodological difference to earlier studies was that half of the participants had received the standard instruction for a lexical decision task ("Indicate whether this is an existing Dutch word"), and the other half were instructed to "Indicate whether you know the word", with the latter being a slightly more intuitive task and drawing less on meta-linguistic reasoning. This manipulation (as part of the efforts to develop the receptive vocabulary test) was implemented to test whether word recognition accuracy would vary as a function of task instruction.

In addition to the word recognition task, the participants had completed a receptive vocabulary test (Dutch version of the Peabody Picture Vocabulary Test Dunn & Dunn, 1997; Schlichting, 2005) and the Dutch version of the Author Recognition Test (Brysbaert, Sui, Dirix, & Hintz, 2020) to assess exposure to literary texts. It is worth pointing out that even though the participants were university students, one may still expect substantial variation in how frequently individuals engage in literary reading in their leisure time (Acheson, Wells, & MacDonald, 2008). That is, while course reading may contribute to how often students read and to the nature of the texts read, it is by no means the case that all students exhibit the same reading frequency. It was therefore important to include tests that gauge individuals' reading experience.

To re-cap, the present re-analysis investigated modality effects on word recognition accuracy and their potential moderators. Specifically, the first goal was to investigate whether the written and audio-visual word recognition benefit reported in previous studies would hold when extending the difficulty range of stimulus words. The second goal was to test whether modality interacts with word difficulty such that, as words become more difficult, recognition accuracy is higher in the written or audio-visual compared to the spoken modality. This hypothesis was based on the observation that written text contains more difficult words than speech. The third goal was to test the hypothesis that individuals with larger receptive vocabularies and more exposure to literary texts, show better recognition accuracy of difficult words compared to individuals with less reading experience. Furthermore, as difficult words are more often encountered in written form, individuals with extensive reading experience and larger vocabularies may have a particular advantage when recognizing difficult words in the written and audio-visual compared to the spoken modality. Thus, the fourth goal of the study was to test whether individuals with more reading experience, reflected in larger receptive vocabularies and more exposure to literary texts, show higher recognition accuracy than individuals with less experience, especially when these words are presented in the written and audio-visual modality.

2.2 Methods

Participants

Forty-eight participants (M age: 22.38 years old, SD = 1.78, 39 female) had contributed to the present dataset. All participants were students at the Radboud University in Nijmegen and were native speakers of Dutch. They had normal or corrected-to-normal vision and hearing, and gave written informed consent prior to testing. Participants were paid for their participation. Half of the participants took part in Experiment 1a, the other half in Experiment 1b. Ethical approval to conduct the study was provided by the ethics committee of the Faculty of Social Sciences at Radboud University.

In addition to the three tests (word recognition test, Peabody Picture Vocabulary Test, Dutch Author Recognition Test) described here, all participants had also completed two spoken processing speed tests (Hintz, Jongman, et al., 2020) and Raven's Advanced Progressive Matrices test (Raven, Court, & Raven, 1998) in the context of the receptive vocabulary test development.

Test materials and procedure

Word recognition test. On each trial of the word recognition test, participants responded to a target word that was presented either visually, auditorily or bimodally (audio-visual). In Experiment 1a, participants were instructed to decide whether the word was an existing Dutch word or not. In Experiment 1b, participants were instructed to indicate whether they knew the presented target word. Participants were told that "knowing a word" meant that they had previously encountered the word and had a vague idea of its meaning. In both sub-experiments, participants were informed that some of the presented targets were made-up nonwords.

The selection of words was based on the prevalence database provided by Keuleers et al. (2015). This database contains prevalence measures for approximately 54,000 Dutch words, approximating to what extent each of these words is known to the whole population (i.e., ranging from $< 5\%$ to $> 99\%$). Keuleers and colleagues established the prevalence values in a large-scale online study involving more than 360,000 unique participants. The participants performed an untimed lexical decision task on a randomly selected set of 100 words. The words were presented visually. The authors established item difficulty (i.e., prevalence) by applying item-response theory (i.e., fitting a Rasch model, Doran, Bates, Bliese, & Dowling, 2007). Using these prevalence values, we selected 240 target words from the database by Keuleers et al. (2015). The mean prevalence for these words was 0.75 (SD = 0.09, range = 0.60 – 0.91). See Appendix I for a full list of word stimuli.

The words for the present study were selected to have similar prevalence values across males and females and different age groups (younger adults, middle-aged individuals, older citizens). Plural forms, past tense forms of verbs, first person singular forms of verbs, and loanwords were not selected. The 240 words were divided evenly into three groups in a way that mean prevalence and range were matched precisely across groups (M = 0.75, range = 0.60 – 0.90). Furthermore, we selected 48 nonwords, which were generated in Wuggy, a multilingual pseudoword generator (Keuleers & Brysbaert, 2010) and used in the mega-study by Keuleers et al. (2015). All of these nonwords had an average accuracy (i.e., correct rejection rate) of at least 90%. See Appendix II for a full list of nonword stimuli.

As for the words, we divided the selected nonwords into three equal groups. Each group of 80 words was complemented with 16 nonwords. The 96 targets in each group were rotated across the three modalities such that each participant was presented with each target only once. Trial presentation was blocked by modality. The order of word and nonword trials within each block was pseudo-randomized prior to the experiment. We counterbalanced the order of blocks across participants. Rotating each target across the three modalities and counterbalancing the order of modality blocks resulted in six experimental lists. Participants were randomly assigned to one list; each participant was presented with all 288 targets (240 words, 48 nonwords, 96 per modality) on a given list.

Each trial started with a centered fixation cross. Participants advanced by pressing a button. Following their button press, they either saw a visually presented target, heard an auditorily presented target or, on audio-visual trials, saw and heard a target (visual and auditory presentation coincided). To parallel the written trials, participants could listen to targets on auditory and audio-visual trials as often as they wanted, just as they could look at the written target for as long as they wanted. They used the right control button on the keyboard to provide a "this is a Dutch word/I know this word" response and the left control button to give a "nonword/ I don’t know this word" response. The task was untimed and participants could take short pauses between the modality blocks.

The dependent variable was word recognition accuracy (1 vs. 0). Our analyses, based on participants’ average word recognition accuracy, showed that the data were neither skewed (-0.29) nor kurtotic (-0.59).

Peabody Picture Vocabulary Test (PPVT). Participants’ receptive vocabulary size was assessed using a digitized version of the Dutch PPVT (Dunn & Dunn, 1997; Schlichting, 2005, for the Dutch translation). On each trial, participants first previewed four numbered line drawings on their screen. When they were ready, they pressed the Return key on their keyboard to hear the probe. They had to indicate which of the pictures best corresponded to the meaning of the spoken word by typing the corresponding number (1, 2, 3, or 4). Following the standard protocol for the test, items were presented in blocks of twelve items, with blocks increasing in difficulty. The starting level was 13, and the best level participants could attain was 17. The test ended when a participant made nine or more errors within one block. Participants took, on average, twelve minutes to complete the test (range: 8 – 15 minutes). The participants’ score was their raw score, that is, the serial number of their last item minus the number of errors made during the test. The maximum score was 204. Analyses, including participants from both sub-experiments, showed that the distribution of scores was neither skewed (-0.23), nor kurtotic (-0.11).

Dutch Author Recognition Test (DART). We used a pen-and-paper version of the Dutch Author Recognition Test, developed by Brysbaert et al. (2020), to measure reading frequency. The Author Recognition Test is a validated, recognized proxy measure of reading frequency (Acheson et al., 2008; Dąbrowska, 2018; James, Fraundorf, Lee, & Watson, 2018; Mar & Rain, 2015; Payne, Gao, Noh, Anderson, & Stine-Morrow, 2012; Stanovich & West, 1989). The underlying assumption is that the awareness level of authors’ names increases as individuals read more often. In the test, participants were provided with a list of 132 names, divided into three columns of 44 names each. The 132 names were 90 names of Dutch and international fiction authors and 42 foils (names of non-authors). Brysbaert et al. (2020) had established the suitability of the material in multiple pre-tests, starting from a list of almost 15.000 fiction (book) authors. The final selection of 90 author names covers the whole difficulty spectrum, ranging from authors that are likely to be known by a large proportion of individuals to authors that are likely to be known only by frequent readers of fiction. The order of author and foil names was random and was the same for each participant. Participants' task was to indicate which of the listed names were authors. Participants' score was the proportion of correctly identified author names minus the proportion of incorrectly selected foils. The maximum score was 1. Analyses, including participants from both sub-experiments, showed that the distribution of DART scores was moderately skewed (1.16) and kurtotic (1.28). Overall, the scores were on the lower end of the performance spectrum suggesting that the test was fairly difficult.

2.3 Results

Table 2.1 summarizes participants' scores on the PPVT and DART. Means, standard deviations (SDs) and ranges were very similar in Experiment 1a and 1b. Importantly, SDs and ranges suggested quite some variability across participants. PPVT and DART were moderately correlated (r = .56) such that participants with larger receptive vocabularies were also frequent readers (i.e., knew more authors).

Table 2.1: Participants' test results on PPVT and DART in Experiment 1a and Experiment 1b.

	Experiment 1a					Experiment 1b				
Task	Mean	SD	Range	Skewness	Kurtosis	Mean	SD	Range	Skewness	Kurtosis
PPVT	178.54	9.49	160-198	0.19	-0.60	178.13	8.82	155-191	-0.76	0.14
DART	26.79	14.19	10-63	1.04	0.38	21.91	9.50	7-41	0.53	-0.78

Note. Range is rounded up for brevity. PPVT = Peabody Picture Vocabulary Test, DART = Dutch Author Recognition Test.

Word recognition test

False alarm rate (i.e., the proportion of 'Yes-responses' to nonwords) was, on average, 8% (SD = 16%, range = 2% - 100%; M Experiment 1a = 5%, M Experiment 1b = 11%). One participant from Experiment 1b was excluded from all analyses because they had a false alarm rate of 100%, which means they responded "Yes, I know this word" to all nonwords. This suggested that they did not take the test seriously or had not understood the task. With the removal of that participant, the false alarm rate dropped to 5% (SD = 4%, range = 2% -

19%). Overall, participants found it easy to recognize the non-existing words (high correct rejection and low false alarm rates). This was the case for all three modality conditions (see also Figure 2.1).

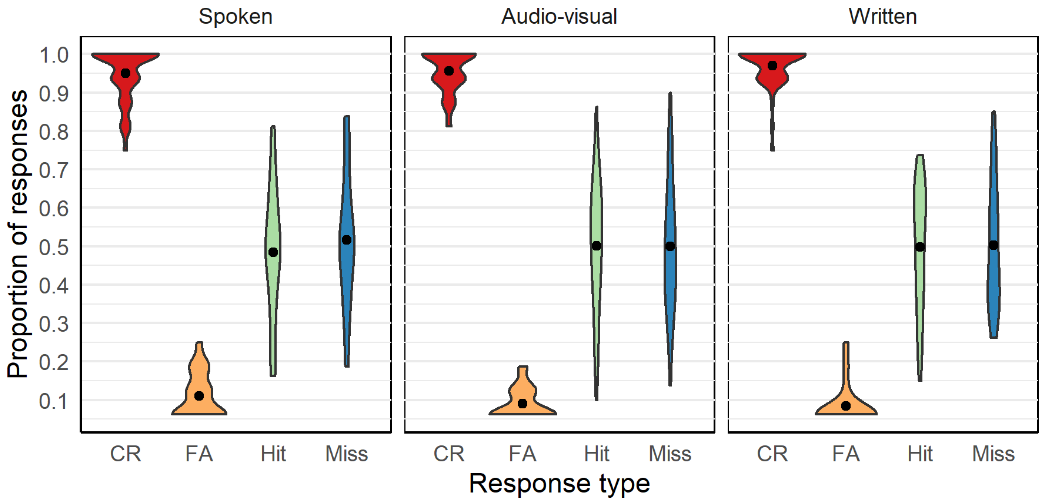


Figure 2.1: Proportion of correct rejections and false alarms for nonwords, and hits and misses for words by modality. CR = correct rejection, FA = false alarm.

Table 2.2 depicts the mean word recognition accuracy by modality condition. Overall word recognition accuracy was 49%. The means suggest there was little difference between spoken, written and audio-visual modalities. Participants in Experiment 1b were numerically slightly less accurate than participants in Experiment 1a. Figure 2.2 plots word recognition accuracy as a function of word difficulty. It is important to highlight that the prevalence scores denoted on the x-axis of Figure 2.2 are not equivalent to 'word recognition accuracy'. Instead, these values were obtained by Keuleers et al. (2015) by applying item-response theory (i.e., a Rasch model). Though the recognition scores were overall lower than expected on the basis of the norming data, the figures shows that there was a strong relationship between the experimental and norming datasets. The correlations of the recognition scores in the three modality with the prevalence values were r = .66 (audio-visual), r = .62 (spoken) and r = .69 (written), respectively.

Table 2.2: Word recognition accuracy by experiment and modality.

	Modality			
	Overall	Spoken	Audio-visual	Written
Experiment 1a	.52 (.17)	.52 (.17)	.53 (.17)	.52 (.17)
Experiment 1b	.46 (.17)	45 (.16)	.47 (.16)	.48 (.16)
Average	.49 (.16)	.48 (.17)	.50 .(17)	.50 (.16)

Note. Standard deviations are displayed in brackets.

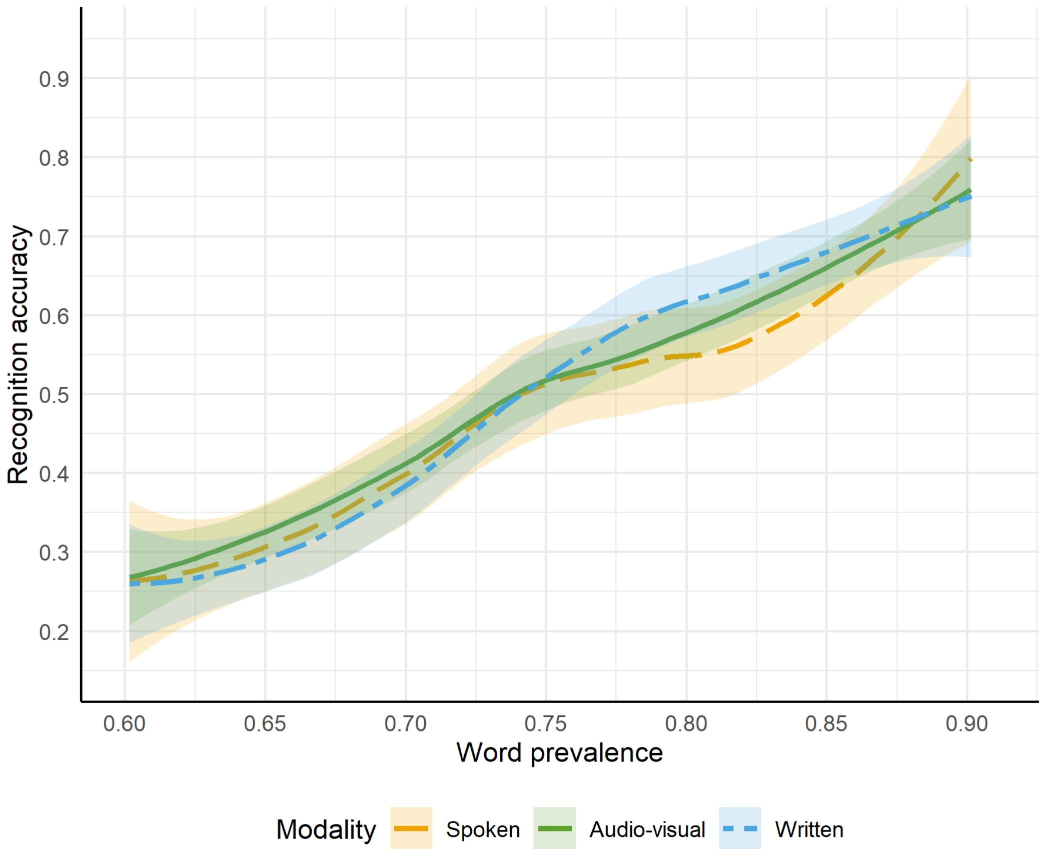


Figure 2.2: Smoothed (using loess regression) word recognition accuracy split out by modality and word difficulty. Bands indicate 95% confidence intervals around the predicted values of the loess regression.

Recognition accuracy was analysed using Bayesian logistic mixed-effects modelling in R (R Development Core Team, 2008), using the brms package (Bürkner, 2017). Analyses were conducted on responses to words. Bayesian analyses are concerned with the likely magnitude of effects rather than statistical significance. Effects were considered meaningful when the 95% Credible Intervals (CI) did not contain zero, which indicates that the parameter has a non-zero effect with high certainty. Moreover, effects were considered meaningful if the point estimate was about twice the size of its error, indicating that the estimated effect is large compared to the uncertainty around it. The posterior probability is reported for these effects, which indicates the proportion of samples with a value equal to or more extreme than the estimate. In addition, Bayes Factors (BF10, BF01) were calculated for all effects, which give an indication of the relative evidence for the alternative hypothesis (H1) compared to the null hypothesis (H0) or vice versa. Our interpretation of the Bayes Factors followed the guidelines by Jeffreys (1961), where a BF of 1 – 3 can be interpreted as anecdotal evidence, a BF of 3 – 10 as substantial evidence and a BF of > 10 of strong evidence for or against the null/alternative hypothesis. Note that BF10 indicates a Bayes factor that favors H1 over H0, and BF01 indicates a Bayes factor in favor of H0 over H1. The model had four chains of 8000 iterations each, with the first half representing a warm-up period. A weak prior (Cauchy distribution with center 0 and scale 2.5 using a sampling algorithm) was used, as is appropriate for non-hierarchical logistic regression models (Gelman, Jakulin, Pittau, & Su, 2008). Models were run until the $\hat{R}$ value for each parameter was 1.00, indicating full convergence. Modality was contrast-coded based on simple contrasts, with the spoken modality being the reference level in the first model, and the audio-visual modality being the reference in the second model. With simple contrast coding, the reference level is always coded as -1/3, and the level that it is compared to is coded as 2/3. This way of coding is similar to treatment contrast coding, but has the advantage that the intercept corresponds to the grand mean instead of corresponding to the mean of the reference level. Moreover, factors outside of interactions can be interpreted as main effects.

The models contained Modality (spoken vs. written vs. audio-visual) as a fixed factor. Word Difficulty was scaled and centered and added to the model as continuous predictor. Participants' PPVT and DART scores were centered and scaled and added to the model as continuous predictors. Because both sub-experiments differed in task, as we included a manipulation of task version ("Is this an existing word?" in Experiment 1a versus "Do you know this word?" in Experiment 1b), Task Version was added as a fixed factor to model the difference in task instruction between the participants. Based on our hypotheses, interactions between Modality, Word Difficulty and PPVT/DART were added to the model. Furthermore, we added interactions between Task Version, Modality and Word Difficulty to test whether Task Version affected the modality effect or the interaction effect between input modality and word difficulty. The random effect structure included random intercepts by word and participant and random slopes for modality by word and participant. The model formula was thus: brm(Correct ∼ (Modality * cWord_Difficulty) * (Task_Version + cPPVT + cDART) + (1 + modality | PP_nr) + (1 + modality | Word), family = bernoulli, data = all_Data, chains = 4, cores = 2, iter = 8000, prior = Pr1).

The full model output for the model with the spoken modality as the reference level is displayed in Table 2.3a and the model output for the model with the audio-visual modality as the reference level is displayed in Table 2.3b. As to be expected, we observed strong evidence for a main effect of Word Difficulty with easier (i.e., more prevalent) words leading to more correct responses than difficult words. We observed no evidence for a main effect of Modality. In fact, the Bayes factors suggested substantial evidence in favor of the null hypothesis (BF01 > 10). Similarly, we did not observe main effects of Task Version (BF01 = 6.67 – 7.69). None of the interactions involving modality showed a significant effect: all showed strong evidence in favor of the null hypothesis. Furthermore, Word Difficulty interacted with Task Version, PPVT, and DART. However, the Bayes factors showed that there was substantial evidence only for the last mentioned interaction. It suggests that frequent readers performed better than less frequent ones in particular for difficult words (Figure 2.3).

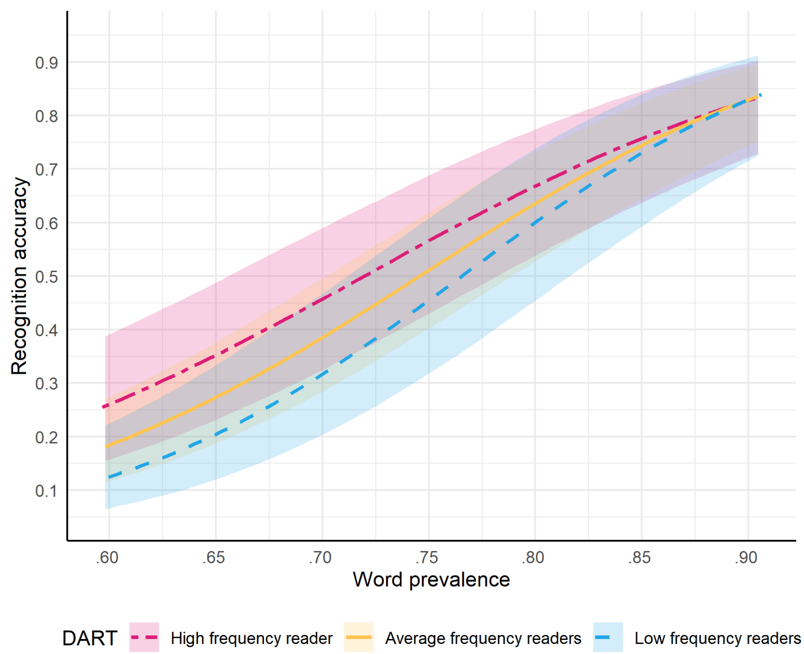


Figure 2.3: Predicted effect of word difficulty (prevalence) and DART scores on recognition accuracy. The shaded areas represent the 95% credible intervals.

Table 2.3a: Full model output for the model with the spoken modality as reference level.

Predictor	Estimate	SE	95% credible interval	BF10	BF01
(Intercept)	0.08	0.20	-0.30, 0.47	0.02	50
Modality: AV	0.07	0.12	-0.16, 0.30	0.04	25
Modality: W	-0.01	0.13	-0.28, 0.24	0.04	25
Word Difficulty	**0.92**	**0.06**	**0.80, 1.03**	**1.53e+13**	**6.54e-09**
Task Version	-0.25	0.27	-0.78, 0.29	0.15	6.67
PPVT	0.00	0.16	-0.32, 0.32	0.05	20
DART	0.20	0.16	-0.13, 0.52	0.11	9.09
Modality: AV * Word Difficulty	-0.02	0.09	-0.20, 0.16	0.03	33.33
Modality: W * Word Difficulty	0.02	0.09	-0.16, 0.21	0.03	33.33
Modality: AV * Task Version	0.05	0.16	-0.27, 0.36	0.05	20
Modality: W * Task Version	0.19	0.19	-0.17, 0.56	0.10	10
Modality: AV * PPVT	0.11	0.09	-0.08, 0.29	0.06	16.67
Modality: W * PPVT	0.00	0.11	-0.23, 0.21	0.03	33.33
Modality: AV * DART	-0.01	0.10	-0.20, 0.18	0.03	33.33
Modality: W * DART	-0.01	0.11	-0.23, 0.21	0.04	25
Word Difficulty * Task Version	**-0.12**	**0.05**	**-0.21, -0.02**	**0.22**	**4.55**
Word Difficulty * PPVT	**0.09**	**0.03**	**0.03, 0.15**	**1.08**	**0.93**
Word Difficulty * DART	**-0.11**	**0.03**	**-0.17, -0.05**	**3.78**	**0.26**
Modality: AV * Word Difficulty * Task version	-0.06	0.12	-0.29, 0.18	0.04	25
Modality: W * Word Difficulty * Task version	0.00	0.12	-0.24, 0.23	0.04	25
Modality: AV * Word Difficulty * PPVT	0.07	0.07	-0.07, 0.21	0.04	25
Modality: W * Word Difficulty * PPVT	0.07	0.07	-0.07, 0.21	0.04	25
Modality: AV * Word Difficulty * DART	-0.02	0.07	-0.16, 0.13	0.02	50
Modality: W * Word Difficulty * DART	0.10	0.07	-0.04, 0.24	0.06	16.67

Note. Meaningful effects are displayed in bold. AV = audio-visual, W = written, PPVT = Peabody Picture Vocabulary Test, DART = Dutch Author Recognition Test.

Table 2.3b: Full model output for the model with the audio-visual modality as reference level.

Predictor	Estimate	SE	95% credible interval	BF10	BF01
(Intercept)	0.08	0.19	-0.29, 0.45	0.02	50
Modality: S	-0.07	0.11	-0.29, 0.16	0.05	20
Modality: W	-0.09	0.13	-0.34, 0.17	0.05	20
Word Difficulty	**0.92**	**0.06**	**0.80, 1.03**	**4.00e+14**	**2.50e-15**
Task Version	-0.25	0.27	-0.79, 0.28	0.13	7.69
PPVT	0.00	0.16	-0.31, 0.31	0.06	16.67
DART	0.20	0.16	-0.12, 0.52	0.12	8.33
Modality: S * Word Difficulty	0.03	0.09	-0.15, 0.21	0.03	33.33
Modality: W * Word Difficulty	0.05	0.09	-0.12, 0.21	0.03	33.33
Modality: S * Task Version	-0.06	0.15	-0.35, 0.25	0.05	20
Modality: W * Task Version	0.14	0.19	-0.23, 0.51	0.08	12.5
Modality: S * PPVT	-0.11	0.09	-0.29, 0.08	0.06	16.67
Modality: W * PPVT	-0.11	0.11	-0.34, 0,12	0.06	16.67
Modality: S * DART	0.02	0.09	-0.17, 0.20	0.03	33.33
Modality: W * DART	0.00	0.11	-0.22, 0.23	0.03	33.33
Word Difficulty * Task Version	**-0.12**	**0.05**	**-0.21, -0.02**	**0.25**	**4**
Word Difficulty * PPVT	**0.09**	**0.03**	**0.04, 0.15**	**1.68**	**0.60**
Word Difficulty * DART	**-0.11**	**0.03**	**-0.17, -0.05**	**3.75**	**0.27**
Modality: S * Word Difficulty * Task version	0.05	0.12	-0.18, 0.28	0.04	25
Modality: W * Word Difficulty * Task version	0.06	0.12	-0.17, 0.29	0.04	25
Modality: S * Word Difficulty * PPVT	-0.07	0.07	-0.21, 0.07	0.04	25
Modality: W * Word Difficulty * PPVT	0.00	0.07	-0.14, 0.14	0.04	25
Modality: S * Word Difficulty * DART	0.01	0.07	-0.13, 0.16	0.02	50
Modality: W * Word Difficulty * DART	0.11	0.07	-0.03, 0.25	0.08	12.50

Note. Meaningful effects are displayed in bold. S = spoken, W = written, PPVT = Peabody Picture Vocabulary Test, DART = Dutch Author Recognition Test.

2.4 Discussion

The present study investigated whether input modality had an effect on word recognition accuracy, whether this modality effect interacted with word difficulty, whether there was an interaction between word difficulty and reading experience on word recognition accuracy, and whether these interactions were influenced by input modality. To address these questions, we re-analysed a dataset collected in the context of the development of a vocabulary test.

Our first goal was to examine how word recognition accuracy would be affected by the modality of word presentation. We hypothesized, in line with previous literature on modality effects in word recognition (Connine et al., 1990; Lopez Zunini et al., 2020; Turner et al., 1998), that word recognition accuracy would be higher when words are presented in the written or audio-visual compared to the spoken modality. Our Bayesian analyses did not confirm this hypothesis. An explanation for this may lie in methodological differences between the present dataset and previous experiments. For example, in order to avoid ceiling effects in accuracy, the words in the present dataset varied much more in word difficulty than the stimulus words selected for standard lexical decision tasks. Extremely difficult words are typically avoided to reduce loss of data due to high error rates. Consequently, errors in traditional word recognition paradigms mostly indicate momentary failures of attention when participants respond to words that they are expected to know. By contrast, in our study, errors most likely indicated that the participants did not know the word. Moreover, unlike in standard lexical decision tasks, responses in the present study were untimed. Time-pressure might be crucial for seeing modality effects. In the written modality, the entire word is immediately available to the cognitive processing systems (see Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), whereas in the spoken modality the same information becomes available in an incremental fashion (see Marslen-Wilson & Tyler, 1980; McClelland & Elman, 1986; Norris, McQueen, & Cutler, 2000). The fact that a word's constituents are available all at once in the visual modality might have led to modality effects in traditional, timed lexical decision tasks where participants respond as quickly as possible. Our results suggest that modality is of less importance in untimed lexical decision tasks, where participants are instructed to consider carefully whether they know the target word or not. Our study was conducted in Dutch, and some of the results discussed here may be language-specific. However, this conclusion – that timed responses might be more sensitive to modality effects than untimed ones – should hold for other languages as well.

Our second goal was to investigate the interaction of modality and word difficulty on word recognition accuracy. We hypothesized that, as words became more difficult, recognition accuracy would be increasingly higher in the visual and audio-visual modality than the spoken modality. Arguably, difficult words are more often encountered in the written form and consequently orthographic representations were predicted to be of higher quality than the phonological representations of the same words. However, this hypothesis was not supported by our findings: There was no significant interaction between difficulty and modality. This may indicate that, even though difficult words are most likely to be encountered in the written modality, their phonological representations might be just as precise and complete as those of easier words. Theories of reading aloud (Coltheart et al., 2001) and reading acquisition (Ehri, 1995; Shankweiler, 1999; Share, 1995) propose a mechanism that describes how phonological representations are created from written input. During recoding, readers mentally recode the graphemes into phonemes upon a written encounter with a novel word, thereby creating both an orthographic and phonological representation of the novel word. Such a mechanism might work specifically well in transparent languages, such as Dutch where graphemes generally map one-to-one onto a phonemes (Seymour, Aro, & Erskine, 2003). It is conceivable that recoding is less efficient in opaque languages, such as English, where grapheme-phoneme correspondences are more unreliable. Moreover, this explanation may also especially apply to the sample tested in the present study. Our participants were university students with no deficiencies in the linguistic domain. Our findings may not generalize to individuals with language or reading disabilities, or individuals with weak grapheme-phoneme correspondences. For these groups, one might find a general advantage of spoken or audio-visual over written presentation or a specific modality advantage for harder words.

The third goal of the study was to investigate the interaction between word difficulty and individual differences in receptive vocabulary size and exposure to literacy texts on word recognition accuracy. We expected, and found, that the indicators of vocabulary size, the PPVT score, and of reading experience, the DART score, were correlated ($r = .56$). This correlation most likely arose because written texts are likely to use a varied vocabulary, including low-prevalence words. Thus, frequent reading enriches a person's vocabulary. We predicted that both variables, PPVT and DART would predict word recognition scores, especially for low-prevalence words. This is because the high-prevalence words should be included in most individuals' vocabularies, whereas the low-prevalence words should be more likely to be included in the vocabularies of individuals with larger receptive vocabularies and more exposure to literary texts.

With respect to the PPVT scores, our prediction that individuals with high PPVT scores would show an accuracy advantage for difficult words over individuals with low PPVT scores was not borne out. The models revealed a statistically significant interaction between PPVT and word difficulty (participants with larger PPVT scores recognized easier words more accurately than participants with lower scores), however, the Bayes factors suggested that there was at best anecdotal evidence for this effect. Given that the Dutch version of the PPVT has been shown to predict adults' word recognition performance in other studies (e.g., Hintz, Jongman, et al., 2020), this result is unexpected and so far unexplained.

For the DART scores, we obtained evidence for the expected interaction. We indeed observed that participants who read frequently (i.e., knew more authors) recognized more difficult words than participants who read less often. This finding corroborates the idea that increased exposure to novel words fine-tunes lexical representations (Castles, 1999; Castles et al., 2007) and that these high quality representations improve word recognition (Perfetti, 2007) by increasing the speed and accuracy of word recognition (Andrews, 1997; Andrews & Hersch, 2010; Perfetti, 1992).

The fourth goal of the present study was to investigate whether the interaction between word difficulty and reading experience on word recognition accuracy was influenced by modality. We predicted that experienced readers, compared to individuals, who are less experienced, would show increased word recognition accuracy of difficult words, especially when these words are presented in the visual and audio-visual modality, as difficult words are most often encountered in the written form. Our results did not provide any evidence for this three-way interaction. A possible explanation may be that, as discussed above, it is possible to create phonological representations of difficult words that are sufficiently precise and accurate to recognize this word in its spoken form efficiently, regardless of reading experience. This explanation, however, may only pertain to transparent languages, such as Dutch, and populations similar to the sample in the present study, which consisted of highly literate university students without any language or reading disabilities. Investigating the modality effect in other languages and in samples with a larger range of language and reading abilities may be important avenues for future research.

A final goal of the study was to explore the effects of different instructions on the participants' word recognition scores. We found that asking participants "Is this an existing word?" versus "Do you know this word?" had no significant influence on their word recognition accuracy.

Though the primary goals of the study concerned the effects of presentation modality, it also offers the opportunity to explore the merit of using prevalence, rather than frequency, as an indicator of word difficulty. We opted for varying prevalence because recent studies had shown that prevalence explained about 7% of additional variance on top of the variance explained by frequency in word recognition tasks. Moreover, criticism has been expressed about the validity of frequency norms for difficult words (Brysbaert, Stevens, et al., 2016; Keuleers et al., 2015). That is, some words with a low frequency of occurrence may not be difficult to recognize, as they are known to a large part of the population. Our data confirmed that prevalence indeed predicted word recognition accuracy, especially for low prevalence words. Therefore, the present study may also be seen as a small-scale validation of the prevalence norms, as it demonstrated the predictive value of the norms in different modalities.

An obvious question is whether prevalence was a better predictor of word recognition accuracy than word frequency. It is important to stress that our study was not designed with this question in mind. Nonetheless, we performed several complementary analyses to explore this issue. We used Google Books to establish the word frequencies for our materials. Search options were set to occurrences in the Dutch language, within Dutch internet pages and restricted to a time window of January 1,1995 to January 1, 2020. The words had a mean frequency of 1468 raw occurrences in Google Books ($SD = 1836$, range = 4 – 12700 occurrences). Frequencies were log-transformed, and correlated with the prevalence values. We found no significant correlation between prevalence and Google frequency (Pearson's correlation: $n = 240$, $r = 0.06$, $p = .36$, Spearman rank correlation: $n = 240$, $r = 0.07$, $p = .30$). This is unexpected as Keuleers et al. (2015) reported a medium-sized correlation ($n \sim 14{,}000$, $r = .35$, based on data from the Dutch lexicon project (Keuleers, Diependaele, & Brysbaert, 2010) of prevalence and frequency. Note, however, that their correlation was based on a different prevalence database than the one we used for the present study. Moreover, we only used a small subset ($n = 240$) of the 54,000 words listed in Keuleers et al. (2015). More importantly, recognition accuracy did not correlate with Google frequency ($r = 0.06$, $p = .35$). This contrasts with the strong correlation between recognition accuracy and prevalence ($r = 0.73$, $p < .001$). We re-ran the Bayesian models described above (Table 2.3a, Table 2.3b), replacing prevalence with Google frequency. There was no evidence for a main effect of Google frequency (estimate = 0.12, SE = 0.08, 95% CI = [-0.03, 0.28]), nor any interaction effects with the other predictors, except anecdotal evidence for an interaction with Task Version. A comparison of the models (using the WA and LOO Information Criteria) showed that replacing prevalence with frequency decreased model fit as reflected in larger LOOIC and WAIC values (model with prevalence predictor: LOOIC = 12410.34, WAIC = 12408.22; model with frequency predictor: LOOIC = 12452.45, WAIC = 12449.98).

Thus, in our study word recognition accuracy was predicted by prevalence, but not by Google frequency. To reiterate, our study was not designed to assess the effects of word frequency and we do not wish to claim that frequency can never have an impact on word recognition. There is, of course, a large body of work clearly demonstrating the influence of word frequency on the speed and accuracy of lexical access in word comprehension tasks (see Brysbaert & Stevens, 2018, for a review). However, it is not known how influential prevalence would be in the same tasks. Important goals for further research would be to develop prevalence norms for other languages than Dutch and to explore and contrast the impact of prevalence and frequency in different linguistic tasks (see Brysbaert, Mandera, McCormick, & Keuleers, 2019, for prevalence norms for 62.000 English words). Frequency and prevalence norms provide complementary information, one telling us how well represented words are in a corpus, the other telling us how well they are represented in the minds of a panel of speakers of the language. High prevalence words are probably recognized by many because they appear often in written and spoken language. Low prevalence words (often technical, political terms), on the other hand, are more likely to be acquired through reading. Each way of garnering information, from corpora or via metalinguistic judgements, has advantages and disadvantages, and consequently the usefulness of the information will depend on the investigator's research goals.

In sum, we found no evidence that the modality of input affected word recognition in Dutch. This held regardless of word difficulty and participants' reading experience. This lack of a modality effect suggests that word knowledge, more specifically individuals' ability to recognize words, can be assessed equally well in the written and spoken modality. However, we wish to stress again that we tested speakers of an orthographically highly transparent language, and that the participants were university students. We cannot rule out that input modality matters for assessments of word recognition ability in less transparent languages and, perhaps more importantly, for assessments of participants with overall lower levels of reading experience or skills.

2.5 Appendix I - Words and their prevalence (prev)

Word	Prev	Word	Prev	Word	Prev	Word	Prev
optornen	0.59	rederijker	0.67	ooglijk	0.75	viltig	0.83
resumptie	0.60	beknorren	0.67	heterofilie	0.75	exotisme	0.83
perfusie	0.60	loosheid	0.67	quotatie	0.75	polytheen	0.83
biocide	0.60	cholerisch	0.67	welvoeglijk	0.75	doffig	0.83
causerie	0.60	polemologie	0.67	deugddoend	0.76	harpenist	0.84
bekijven	0.60	totalitarisme	0.67	routeren	0.76	inlassing	0.84
gepresseerd	0.60	biduur	0.67	postuleren	0.76	wijduit	0.84
pikkel	0.60	vlieden	0.68	primeren	0.76	loffelijk	0.84
situatief	0.60	verzoeten	0.68	xenofilie	0.76	aftroggelarij	0.84
oncogeen	0.60	nestvlieder	0.68	omdoping	0.77	ritualist	0.84
purgeren	0.61	convertor	0.68	zinsbedrog	0.77	fenomenologie	0.84
lorrig	0.61	stribbeling	0.69	rotskunst	0.77	geruim	0.84
diachronisch	0.61	peigeren	0.69	abdominaal	0.77	condoleantie	0.84
overhoeks	0.61	reconversie	0.69	spiritist	0.77	verwijven	0.85
juvenaat	0.61	endemie	0.69	schalm	0.77	onbemerkt	0.85
epiek	0.61	decagram	0.69	ordinaat	0.77	onberecht	0.85
afpaling	0.61	demissie	0.69	eedbreuk	0.78	afratelen	0.85
wierig	0.61	lobberig	0.69	homeostase	0.78	autogram	0.85
adventief	0.61	erving	0.69	endotherm	0.78	replicator	0.85
performant	0.61	cytologie	0.69	carbolineum	0.78	afstuiven	0.86
papist	0.62	collectioneren	0.69	legalisme	0.78	interferentie	0.86
negatie	0.62	serafine	0.70	paleografie	0.78	lofprijzend	0.86
marmiet	0.62	wiegelen	0.70	exclusie	0.78	hallucinant	0.86
stilet	0.62	scharren	0.70	afkukelen	0.78	lymf	0.86
andragoog	0.62	sloffig	0.70	versmachten	0.79	zinnelijk	0.87
lijdelijk	0.62	agronomie	0.70	weeklacht	0.79	cyclisme	0.87
debiliseren	0.62	relevatie	0.70	rasperig	0.79	admissie	0.87
triangulatie	0.63	ressorteren	0.71	spitsig	0.79	reductionisme	0.87
dras	0.63	bezemklas	0.71	patroneren	0.79	onduldbaar	0.87
spijzigen	0.63	zwemblaas	0.71	schrijnwerker	0.79	brosheid	0.87
indolentie	0.63	geschulpt	0.71	oplaaiing	0.79	plichtig	0.87
bescheid	0.63	utilitarisme	0.71	historiek	0.80	curatief	0.87
walen	0.63	secretarie	0.71	verificateur	0.80	cilindrisch	0.87
tabuleren	0.63	nomadisme	0.71	biogeen	0.80	solutie	0.87
ganselijk	0.63	bijtreden	0.71	kwetsing	0.80	welgeaard	0.87
rijmelen	0.64	sculpturaal	0.72	signalisatie	0.80	smakker	0.88
flinterig	0.64	gewemel	0.72	radiatie	0.80	veeweide	0.88
silicose	0.64	heliocentrisch	0.72	vermaking	0.80	linkerrij	0.88
raagbol	0.64	insolide	0.72	pylon	0.80	vlotweg	0.88
vezelen	0.64	beroemen	0.72	schuieren	0.80	thuisloos	0.88
waadbaar	0.64	havist	0.73	promoting	0.80	opraapsel	0.88
schrokkig	0.65	assertie	0.73	fundatie	0.81	eruitzien	0.88
bedaagd	0.65	fantasme	0.73	meerderwaardig	0.81	mediaan	0.89
suppositie	0.65	verevenen	0.73	poolshoogte	0.81	tinkelen	0.89
hagelsnoer	0.65	reticulair	0.73	reformisme	0.81	propeller	0.89
gruizelen	0.65	knoeper	0.73	willigen	0.81	verzeilen	0.89
supprimeren	0.66	verderfenis	0.74	scenarist	0.81	uitschateren	0.89
syndicaal	0.66	kroezelig	0.74	lamelle	0.81	onromantisch	0.89
ontologie	0.66	plagiator	0.74	obsederen	0.81	kletserig	0.90
netelen	0.66	uitloven	0.74	sjokkerig	0.82	spijten	0.90
smoezig	0.66	picturaal	0.74	modereren	0.82	baatzucht	0.90
ingeboren	0.66	wanbesef	0.74	fluoresceren	0.82	bezaaiing	0.90
futselaar	0.67	traverseren	0.74	homologie	0.82	wilsgebrek	0.90
oculeren	0.67	roezemoezig	0.74	alreeds	0.82	erkentenis	0.90
morrig	0.67	deputatie	0.74	empirist	0.82	onheilig	0.90
afbietsen	0.67	temporeel	0.74	stranding	0.82	kleerhaak	0.91
beredderen	0.67	taxonoom	0.75	keutelig	0.82	kwadrateren	0.91
ordinantie	0.67	kosterij	0.75	knapperen	0.83	vermissen	0.91
schoeiing	0.67	exciteren	0.75	druksel	0.83	duellist	0.91
satineren	0.67	postaal	0.75	nijverig	0.83	regularisatie	0.91

2.6 Appendix II - Nonwords

garidijn
reurkop
strensen
omkaven
spietsel
flingen
violaan
keelnaald
kokasme
ocuur
eldernek
aankoven
blautenie
apamoleek
artolen
spaaien
bekielen
tarateel
schrallen
lording
zeelap
avatoir
doxerperen
lommiek
artissor
spanbos
pontijn
puivel
smoemen
afzeinsen
hosteren
hekeren
infossen
grosteren
fonzig
akopeel
onloeien
stuim
atrugie
gravieur
blenkelen
verratsen
luiring
roezing
hatselen
afhoezen
karteloen
bleemheid

3 Recognizing words varying in their difficulty in a diverse sample: effects of presentation modality and reading experience

Abstract

There are large differences between individuals regarding the amount of language they encounter in the written modality. Written and spoken language tend to differ, in particular with respect to how many and what kind of words are encountered. Written language allows for more language intake, as reading is faster than speech, and more encounters of difficult words. In the present study, we examined how presentation modality and word difficulty affect word recognition, and how this interaction is affected by individual's experience with written language. We recruited participants (n = 156) from both the university and vocational student population, who completed an online word recognition test with words varying in difficulty, presented in the spoken, written and audio-visual modality. Reading experience was assessed using objective (receptive vocabulary size, author knowledge) and subjective (reading frequency questionnaire) measures. For word recognition accuracy, no modality effects were observed, but experienced readers displayed an advantage over less experienced readers at recognizing high to medium difficulty words. Analyses of reaction times suggested that the word difficulty effect (faster responses to easy words) was larger in the written compared to spoken modality, and larger for experienced readers than less experienced readers. Thus, our study provided evidence that recognition of words varying in difficulty is influenced by reading experience and presentation modality.

3.1 Introduction

Through reading, literate individuals are able to store two representations (orthographic and phonological) of the same lexical item. As a result of the quality and quantity of modality-specific encounters, these two representations may vary in their level of precision and completeness (i.e., lexical quality Perfetti, 2007). The quantity and quality of modality-specific encounters may vary substantially among individuals, particularly as a result of individuals differing in how often they read. Reading is a fast and, for many people, fairly automatized process, which offers opportunities to engage with language and encounter many words. Extensive reading provides opportunities to extend one's lexicon and improve orthographical representations of words encountered in the written modality. This may be particularly the case for more difficult words, as written language is characterized by a larger variation of words than those that we tend to use in everyday speech (Cunningham & Stanovich, 1998; Hayes & Ahrens, 1988). People who do not read frequently may therefore be more likely to have smaller vocabularies and less precise orthographical representations. Hence, reading experience may exert a large influence on the efficiency of people's word recognition: due to differences in quality of orthographical representation as a result of differences in reading experiences, the ease of word recognition may differ for words encountered in the written vs. spoken modality.

In the previous study (Chapter 2), which explored the relationship between reading experience, presentation modality and word difficulty, variation in reading experience was found to be related to individual differences in word recognition efficiency. No effects of input modality were observed. The previous study, however, was administered to a homogeneous group of university students, who arguably are highly experienced readers. To increase the variability among participants, the present study recruited its participants from a broader population. This more diverse sample may provide opportunities to observe evidence of the effect of reading experience on written and spoken word recognition.

The literature suggests that there is a word recognition benefit for words presented in the visual and bimodal (audio-visual) compared to auditory modality: written and audio-visually presented words have been reported to be recognized more accurately and faster than spoken words (Connine et al., 1990; Lopez Zunini et al., 2020; Turner et al., 1998). An explanation for this phenomenon may be that extensive reading may improve the quality of orthographic representations. Reading is also faster than listening to speech, and therefore allows for relatively more language exposure within a given time frame (Brysbaert, Stevens, et al., 2016). Extensive exposure to language is thought to improve the quality of lexical representations (cf. lexical tuning hypothesis, Castles, 1999; Castles et al., 2007), and high quality representations are assumed to lead to more efficient processing compared to low quality representations (Perfetti, 2007). Extensive reading may strengthen orthographic representations in particular and subsequently improve recognition of written words, giving rise to the observed recognition benefit in the written and audio-visual modality.

This recognition benefit may express itself particularly strongly in two scenarios: for more difficult, less commonly known words, and in more experienced readers. That is, written texts often contain more difficult, uncommon words than spoken language (Cunningham & Stanovich, 1998; Hayes & Ahrens, 1988). Therefore, the orthographical representations of these uncommon words may become more complete and precise than their phonological counterparts. This may facilitate the recognition of difficult words presented in the written compared to spoken modality. According to this line of reasoning, accuracy and speed differences between recognizing easy and more difficult words should be smaller for words presented in the written modality, but larger for words presented in the spoken modality.

In contrast to this prediction, evidence has been observed for word difficulty explaining more variance in word recognition speed and accuracy for words presented in the written (semi-partial $R^2 = 0.36$ and 0.24 respectively) than spoken modality (semi-partial $R^2 = 0.05$ and 0.07 respectively) (Ferrand et al., 2018). This suggests that word difficulty, operationalized as word frequency, had a larger effect in written compared to spoken word recognition. Ferrand et al. (2018) did not report whether there was an effect of word difficulty on word recognition accuracy as well. Their results do suggest that the RT difference between recognizing easy and more difficult words is larger in the written than spoken modality. However, when interpreting results comparing these two modalities, one must keep in mind that each modality has different processing demands. Due to the processing demands of the spoken modality, the strongest predictor of auditory lexical decision times was spoken word duration (semi-partial $R^2 = 0.24$). This predictor had no effect on written word recognition (semi-partial $R^2 = 0.00$), because efficient processing of the written form does not rely on the acoustic duration of the perceived word. Thus, the finding that word difficulty explained little variance in the spoken modality may have been a result of a fundamental difference in processing between the spoken and written modality. The temporal nature of the spoken modality may have lowered the impact of word-level characteristics such as word difficulty. It is therefore important to control for spoken word duration when analysing word recognition latencies across modalities.

The recognition benefit in the written and audio-visual compared to spoken modality may also vary as a function of an individual's reading experience. Experienced readers, compared to less experienced readers, have had many opportunities to encounter words in their written forms, thereby allowing their orthographical representations to become of high quality. Since written text often contains more difficult words than everyday speech, experienced readers' orthographical representations of difficult words may be of high quality compared to people who are less experienced readers. Their recognition of difficult words may thus be enhanced in the written compared to spoken modality. On the other hand, literature as described above (Ferrand et al., 2018) found word difficulty effects to be increased in the written compared to spoken modality. If this increased word difficulty effect is the result of the extensive exposure to language that the written but not spoken modality provides, we may expect that the pattern of findings we observe with regard to the interaction between modality and word difficulty may be even stronger for individuals who read often.

The previous study (Chapter 2) investigated the effects of presentation modality, word difficulty and reading experience on word recognition accuracy in a homogeneous sample of university students and found that word difficulty and reading experience influenced word recognition accuracy. The word recognition test used in this previous study consisted of words varying in the degree to which they are known to the general population (i.e., "prevalence") and nonwords. Words were presented in the written, spoken or audio-visual modality and participants had to decide whether they knew them or not. Participants' reading experience was assessed using the Dutch version of the Author Recognition Task (DART), which is a validated measure of print exposure, and a measure of receptive vocabulary (Peabody Picture Vocabulary Test). We found that experienced readers, as measured with the DART, recognized difficult words more accurately than participants who were less experienced readers.

The results from the previous study indicate that even in a homogeneous sample of university students, who in general terms are highly experienced readers, variation in reading experience affects word recognition. There was no evidence for a modality effect or interactions with modality in this previous study. The existence of individual differences in reading experience among university students and their modulating effect on overall word recognition accuracy suggests that interaction effects between modality and reading experience are more likely to surface in a group of participants that is more varied with regard to their reading experience. One way to obtain a more diverse sample is to recruit not only university students, but also age-matched vocational education students. Previous studies showed that a sample of students from various educational backgrounds exhibit larger variation with respect to language experience than homogeneous samples of university students (Mainz, 2018; Mainz et al., 2017). Moreover, this way of diversifying a sample has been proven effective previously, as studies that recruited both vocational and university students demonstrated that individual differences in language experience and cognitive abilities experience were related to variation in language skills (Hintz, Jongman, et al., 2020; Jongman, Khoe, & Hintz, 2020; Mainz, 2018).

Present study

The present study aimed to continue investigating 1) how presentation modality affected recognition of words varying in difficulty and 2) how the relationship between presentation modality and word difficulty is moderated by reading experience. As in the previous study, participants carried out a lexical decision task, where nonwords and words varying in difficulty were presented in the written, spoken or audio-visual modality. In addition, they completed measures of receptive vocabulary (Peabody Picture Vocabulary Test) and print exposure (Dutch Author Recognition Test) to approximate their reading experience. The present study differed from the previous study (Chapter 2) in five aspects.

First and most importantly, participants were sampled from a more diverse population in terms of educational background. Participants were not only sampled from the university student population, traditionally tested in psychological and linguistic research (Andringa & Godfroid, 2020; Arnett, 2016), but also from the vocational student population. Research has shown that, on average, Dutch students in preparatory vocational secondary education read less than students in preparatory scientific secondary education (Dood et al., 2020). It is likely that this trend continues in tertiary education and as such, it is likely that vocational students are less experienced readers than university students. Using a heterogeneous sample with respect to reading experience may increase the sensitivity to detect effects of reading experience and its interaction with presentation modality and word difficulty.

A second change was that we adjusted the difficulty of the words in the word recognition test such that the selected words were on average slightly easier than the words used in the previous study. This was done to ensure that our less experienced readers, who likely had a smaller vocabulary than experienced readers (Mainz et al., 2017; Mol & Bus, 2011; Sullivan & Brown, 2015; West, Stanovich, & Mitchell, 1993), were able to perform the task. The adjustment of word difficulty is explained in more detail in the Methods section.

A third key difference was that the present study was administered as an online experiment, whereas the previous study was administered in the lab. Due to the COVID-19 pandemic, in-person testing was not possible. A downside of online testing is that environmental influences (e.g., noise, presence of other people) are not controllable and that motivation of participants may be lower than during in-lab test situations. To address this, we provided the participants with clear test instructions and applied very strict exclusion criteria when pre-processing the data, which are described in more detail in the Methods and Results section.

A fourth difference was that, in the present study, response speed during the word recognition test was recorded alongside accuracy. Response latencies are a complementary measure of word recognition to recognition accuracy, as they measure the ease (or difficulty) with which words are recognized, rather than whether the recognition process is successful or not. Thus, by measuring not only the success but also the ease with which words are recognized, we hoped to get a broader understanding of how reading experience influences both the accuracy as well as ease with which words are recognized.

A final change was that we used a latent variable approach to approximate reading experience. Reading experience is a multi-facetted construct and therefore difficult to approximate using a single measure. Therefore, we calculated a reading experience factor score for each individual separately, based on their scores on the DART (an objective assessment of reading behaviour), the PPVT (an objective measure of language experience), and a reading behaviour questionnaire (a subjective measure of reading behaviour). By using three measures that are assumed to tap different aspects of reading experience and differ in their level of subjectivity, we aimed to capture people's interaction with written text more exhaustively.

Hypotheses

With regard to word recognition accuracy, we predicted an interaction effect between presentation modality and word difficulty: we expected to see an increasing accuracy benefit for the written and audio-visual compared to the spoken modality as words become more difficult. Since written text contains more difficult words, we expected that orthographic representations of difficult words would be of higher quality than their phonological counterparts, resulting in more accurate recognition of a difficult word's orthographic form in the written and audio-visual modality compared to phonological form in the spoken modality. In addition, we predicted a three-way interaction effect between presentation modality, word difficulty and reading experience in that we expected the written and audio-visual presentation modality accuracy benefit for difficult words to be larger for people who have had extensive reading experience. Experienced readers will have had more opportunities to encounter difficult words in the written modality as a result of their increased reading frequency, thereby sharpening the orthographical representations of these difficult words and showing a recognition benefit in the written and audio-visual over spoken modality.

With regard to response speed, directly comparing the different presentation modalities is difficult for two reasons. First, stimulus presentation in online experiments is subject to jitter due to technical characteristics of the participant's machine and/or internet connection. Auditory presentation in particular, but not visual presentation, is affected by this uncontrollable jitter. This results in presentation delays of 40 – 60 ms for auditory stimuli, as observed through extensive testing of our in-house software. To control for this, we calculated a baseline reaction times variable by averaging reaction times for each individual in each modality separately and added this variable as a predictor to our reaction time models. This baseline reaction time variable will explain the variance in reaction times related to presentation delays that varied between participants and presentation modality.

A second difficulty in comparing reaction times in different modalities is that the processing speed of written and spoken materials differ. People are faster readers than listeners, because written information is available all at once, whereas spoken information unfolds incrementally. Thus, we cannot distinguish whether a main effect of modality in reaction times is the result of this inherent difference in processing speed, or whether it is the result of difference in word recognition speed. Therefore, we do not make predictions regarding main effects of modality on reaction times, but instead predict that the speed difference between recognizing difficult and easier words differs as a function of presentation modality. For all modalities, difficult words are expected to be responded to more slowly than easy words. We formulated two contrasting hypotheses regarding the effect of presentation modality. First, RT differences between difficult and easy words may be smaller in the written and audio-visual modality than spoken modality.

This hypothesis was guided by the idea that written language is more diverse in terms of word difficulty than spoken language, thereby strengthening orthographic representations of difficult words, which in turn would facilitate written and audio-visual recognition of difficult words compared to spoken recognition. The second hypothesis predicts the contrary: that RT differences between difficult and easy words would be larger in the written and audio-visual compared to spoken modality. This prediction is guided by previous results (Ferrand et al., 2018; Hasenäcker, Verra, & Schroeder, 2019), which reported RT differences between difficult and easy words to be larger for words presented in the written than spoken modality. Our final prediction is that, regardless of direction, the effect is expected to be larger in experienced readers, as they — compared to less experienced readers — expose themselves more often to written language, thereby magnifying the modality effect.

3.2 Methods

Participants

We conducted a power analysis using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) to calculate the required number of participants to find a medium sized three-way interaction effect of presentation modality, word difficulty and reading experience. In G*Power, settings were set to the F-test family, "linear multiple regression test, fixed model, R^2 increase", effect size of $f^2 = .15$ (medium), $\alpha = 0.05$, power = .95. The number of tested predictors was set to 4 (two two-way interactions presentation modality * word difficulty + two three-way interactions presentation modality * word difficulty * reading experience; since modality had three levels, each interaction was modelled twice to be able to compare each level to each other). Total number of predictors was set to 11 (presentation modality (3 levels), word difficulty and reading experience, and all their interactions). The total sample size as calculated by G*Power was 129. This calculation guided our decision to aim for a minimum sample size of 140 participants: 70 university students and 70 vocational education students.

In total, 156 participants (M age: 20.97 years old, SD = 1.95, 109 female) took part in the present study. All except for one participant stated that Dutch was (one of) their native language(s). Most participants (n = 143) stated they were monolingual Dutch by birth, and thirteen participants were brought up bilingually with Turkish (n = 4), Papiamento (n = 2), Arabic (n = 2), English (n = 2), German, Greek and Frisian (all n = 1) as a second native language.

Seventy-six of the 156 participants were university students and recruited through the participant database of the Max Planck Institute for Psycholinguistics. Fourteen were enrolled in a bachelor's programme at an applied university, 28 in a bachelor's programme at a university, 33 in a master's programme at university, and one participant was in their final year of pre-university secondary education. The other 80 participants were recruited at several vocational education schools. Seventy-four of these participants were enrolled in a program on the fourth and highest level of vocational education, four of them on the third level, and two on the first and lowest level of vocational education.

All participants had normal or corrected-to-normal (n = 30) vision and hearing, except for one participant who stated they had uncorrected eyesight. Most participants (n = 139) were right-handed, thirteen participants were left-handed and four participants were ambidextrous. Twenty participants, all of which enrolled in vocational education, were diagnosed with dyslexia. One participant had a speaking disorder (stuttering) and nine participants stated they had a diagnosis of attentional deficits. Participants gave online informed consent prior to testing and were paid € 10,- for their participation. Ethical approval to conduct the study was provided by the ethics committee of the Faculty of Social Sciences at Radboud University.

Test materials

Word recognition test. The word recognition test used in the present study was the same as in Chapter 2 with a few adjustments. Participants were presented with a target word and were asked to decide whether they knew the word or not. Target words were presented in their written, spoken or audio-visual form.

Target words difficulty was operationalized using prevalence norms rather than word frequency norms. Prevalence refers to the degree to which a word is known in the population and are a more reliable predictor of word knowledge than frequency norms, particularly for low-frequency words. Prevalence and frequency norms tend to correlate to a medium degree (r = .35) (Keuleers et al., 2015).

Compared to the previous study, in which only university students participated, target words in the present study were selected to be slightly easier (i.e., of higher prevalence). In the previous study, the average prevalence of the target words, as obtained from the prevalence database provided by Keuleers et al.

(2015), was 75%, with a range of 60% – 90%. For the present study, 175 words from the 240 target words used in the previous study were selected. These 175 words had a prevalence of 67% – 90%. In addition, 65 words with a prevalence ranging from 91% – 98% were selected from the prevalence database. This resulted in a final selection of 240 target words with a mean prevalence of 82% and a prevalence range of 67% – 98%. In addition, the 46 nonwords from the previous study, created using the multilingual pseudoword generator Wuggy (Keuleers, Diependaele, & Brysbaert, 2010) were added to the selection of words. The selected target words are provided in Appendix I, with the new target words in bold.

To investigate the association between the prevalence and frequency of our selected words, frequency norms for the words used in the present study were established using Google Books. Search options were set to occurrences in the Dutch language, within Dutch internet pages and restricted to a time window of January 1 1995 to January 1 2020. The words had a mean frequency of 2585 occurrences in Google Books (SD = 4132, range = 8 – 31200 occurrences). The correlation between log-transformed Google frequencies and prevalence was medium, ($r = .27$, $p < .001$), similar to reports in previous studies.

As in the previous study, the 240 words and 46 nonwords were divided into three groups of 96 words. The 96 targets in each group were rotated across the three modalities such that each participant was presented with each target only once. Trial presentation was blocked by modality. The order of word and nonword trials within each block was pseudo-randomized prior to the experiment. The order of blocks was counterbalanced across participants, resulting into six experimental lists. Participants were randomly assigned to a list. During the task, participants were presented with a written, spoken or audio-visually presented word (presentation of written and spoken form coincided), and were instructed to decide as quickly as possible whether they knew the word or not. Both accuracy and reaction times (the difference between reaction time and stimulus presentation onset) were recorded.

In contrast to the experiment described in Chapter 2, written words were shown for the acoustic duration of their corresponding audio recording, so that sensory stimulation was approximately equal across conditions. Mean stimulus duration was 1009 ms (SD = 801 ms). Target words were only presented once. To avoid the possibility participants responded "I don’t know this word" in cases where they were unable to process the word, we included an "I did not hear/see the word"-button, which appeared after the written word had disappeared and/or the recording of the spoken word had ended. Another change to the test was the introduction of a practice phase, consisting of four real words and two nonwords (equally distributed across modalities) to make participants familiar with the task and the button-response association (M-button: "I know this word", Z-button: "I don't know this word"). During the practice trials, the correct/incorrect labels were presented on the computer screen in red and green, respectively.

Peabody Picture Vocabulary Test (PPVT). Participants' receptive vocabulary size was assessed using the Dutch PPVT (Dunn & Dunn, 1997; Schlichting, 2005, for the Dutch translation). On each trial, participants saw four line drawings on their screen and heard a recording of the target word. Participants had to indicate which picture corresponded to the meaning of the target word by clicking on the correct picture. Following the standard protocol for the test, items were presented in blocks of twelve items, with blocks increasing in difficulty. The test ended when a participant made nine or more errors within one block. The participants' score was their raw score, that is, the serial number of their last item minus the number of errors made during the test.

Dutch Author Recognition Test (DART). The Dutch Author Recognition Test (Brysbaert et al., 2020) was used to measure print exposure. The Author Recognition Test is a validated proxy of print exposure, based on the assumption that the awareness level of authors' names increases as individuals read more often. Participants saw a list of 132 names, consisting of 90 Dutch and international authors and 42 foil names. The list was presented in three columns of 44 names each. The order of author and foil names was random and was the same for each participant. Participants were instructed to indicate which authors they recognized by ticking the boxes behind these names. Participants were informed that some of the names were foils to ensure that they did not mark all names as authors. Participants' scores were the percentage of correctly identified author names minus the percentage of incorrectly selected foils.

Reading behaviour questionnaire. The questionnaire measuring reading behaviour was adapted from the questionnaires used in the Programme for International Student Assessments (PISA) of 2009 and 2018. Dutch questions were retrieved from of Dutch outcomes reports (Gubbels, van Langen, Maassen, & Meelissen, 2019; Kordes, Feenstra, Partchev, Feskens, & de Graaf, 2012). The questionnaire, translated into English, can be found in Appendix II. The questionnaire consisted of three domains, namely reading frequency, reading enjoyment and attitudes towards one's own reading ability. There were 24 questions in total. All questions were answered on a 5-point Likert scale. Scores on negative-worded questions were transposed, so that high scores indicated positive reading behaviours (high reading frequency, high reading enjoyment and favourable attitudes towards one's reading ability). For each participant, a sum score was calculated.

Procedure

Participants were sent the link to the online experiment via email. They were instructed to take part in the experiment in one sitting, to use headphones and to sit in a quiet room. First, participants gave informed consent. Then personal information were obtained, including several background questions regarding their native language(s), handedness and problems with sight, hearing, speech or reading. Next, participants filled out the reading questionnaire. Participant's audio settings were checked to ensure they heard the spoken stimuli before they received instructions for the word recognition test. They first completed the practice trials and then performed the recognition task. Then, the PPVT and the DART were administered. At the end of the experiment, participants were debriefed. After each task, participants were informed about their progress. The test session took between 45 and 50 minutes.

3.3 Results

Data pre-processing

Online data collection naturally allows for less environmental control than in-lab studies. Therefore, data were checked based on several predetermined exclusion criteria. If one exclusion criterion was met, all data from this participant were excluded from further analyses.

Background questions. Participants were excluded if they provided nonsensical answers, if they indicated that they had uncorrected vision or hearing problems or if they did not state that Dutch was one of their native languages. One participant was excluded due to not stating Dutch as one of their native languages.

Reading behaviour questionnaire. Participants were excluded if 100% of their responses belonged to a single category on the 5-point response scale, including the questions that were reversely coded and should have been answered in the opposite direction. No participants were excluded based on these criteria

PPVT. Participants were excluded if they had a score of 0, indicating that they made more than five errors out of the twelve items in each difficulty block. One participant was excluded due to having a PPVT score of 0, indicating that they did not perform the task in a serious manner or that they did not understand the task. This was the same participant that also indicated that they had uncorrected eyesight problems. In addition, participants were excluded if 100% of their responses belonged to only one of the four quadrants in which the pictures appeared during the test. None of the participants met this criterion.

DART. Participants were excluded if they selected none of the authors (miss rate of 100%) or all foil authors (false alarm rate of 100%). None of the participants met these criteria. However, one participant was excluded from the analyses due to having a score of 94 on the DART (maximum = 100). Further inspection revealed that it took them 24.05 minutes to complete the DART, whereas the mean completion duration was 3.18 (SD = 2.92) minutes. This suggests that this person looked up the names of the authors during the test. For two participants, DART scores were calculated manually due to the server not storing their scores during the online test session. These participants sent photographs of their screens during the test session, which showed which authors they ticked on the DART.

Word recognition test. Regarding the word recognition test, participants were excluded if they responded on more than half of the trials that they "did not hear or see the word". None of the participants met this criterion. In addition, participants were excluded when they responded "I know this word" to 75% or more of the foil words (false alarm rate) or when they responded "I don't know this word" to 75% of the target words (miss rate). Two participants were excluded due to having a false alarm rate of 85% and 75% respectively. Furthermore, 0.5% of the data were excluded as a result of participants clicking the button "I did not hear/see the word" and 0.2% of the data were excluded due to a response logging error in the Frinex system (i.e., negative RTs). Finally, trials with RTs below 300 ms and above 5 seconds were removed, resulting in data loss of 0.88% and 0.48% respectively. In total, 98.61% of the data of the word recognition test was used in the analyses.

Descriptive analyses

Reading behaviour questionnaire. Figure 3.1 displays the distribution of the sum scores for the questionnaire as a whole, as well as split by domain. Table 1a summarizes participants' scores on the reading behaviour questionnaire. There was variability across participants as shown by the SDs and ranges. Moreover, as can be seen in Table 1b, university students scored higher on the questionnaire than vocational students, indicating that they read more frequently. Figures 3.2a and 3.2b corroborate this pattern, suggesting that education level influenced reading behaviour. Differences between university and vocational students' responses on separate questions of the reading behaviour questionnaire are provided in Appendix III.

PPVT and DART. There was variability in participants' PPVT and DART scores as shown by the SDs and ranges in Table 3.1a. Table 3.1b and Figure 3.2a and 3.2b suggest that university students had a larger receptive vocabulary (top middle panel) and knew more authors (top right panel) than participants enrolled in vocational education.

Correlations. Correlations between the reading behaviour sum score, scores on the PPVT and ART and average accuracy on the word recognition test are displayed in Table 3.2. The reading behaviour sum scores correlated moderately with PPVT and DART scores in the expected direction. PPVT and DART were also moderately correlated.

Factor analysis reading experience factor. Given these moderate correlations between reading measures, for our modelling analyses, we condensed the three variables into one factor score. To that end, an unrotated, scaled principal component analysis (PCA) was performed in R version 4.0.3 (R Development Core Team, 2008). Prior to the PCA, assumptions were checked. Variables were measured at the continuous level and were linearly related. Sample adequacy was good (KMO statistic = .72, KMO statistics for each variable = .71 – .73) (Kaiser, 1974) and a significant Bartlett's Test of Sphericity indicated that the data was suitable for data reduction. This principal component analyses resulted in a

3-factor solution. The first factor explained 73.37% of the variance. Each individual's score on the first component was used as a reading experience factor score in further analyses. As expected, there was considerable variation between participants indicated by the SDs and range of this factor score (Table 3.1a). University students scored higher on the reading experience factor score than vocational education students (Table 3.1b, Figure 3.2a and 3.2b, bottom right).

Table 3.1a: Participants' scores on the variables of the present study.

	Mean	SD	Range	Skewness	Kurtosis
Accuracy	0.62	0.15	0.27, 0.94	-0.17	-0.59
RT words	1207.63	545.52	301, 4946	2.06	7.16
RT nonwords	1170.18	488.88	205, 4926	2.37	9.70
PPVT	167.84	14.01	136, 194	-0.23	-0.94
DART	20.77	13.76	-1.27, 65.40	0.88	0.42
Reading behaviour sum score	79.16	12.05	51, 108	-0.24	-0.50
Reading experience factor	0	1.48	-3.28, 3.46	0.12	-0.77

Note. Accuracy is only based on responses to words, nonwords were excluded. Reaction times are based on correct responses to words and nonwords.

Table 3.1b: Participants' scores on the variables of the present study, split by education level.

	Vocational education			University		
	Mean	SD	Range	Mean	SD	Range
Accuracy	0.6	0.16	0.27, 0.93	0.64	0.14	0.28, 0.94
RT words	1242.87	577.06	301, 4946	1175.72	512.22	345, 4904
RT nonwords	1231.62	540.1	205, 4926	1117.87	433.87	462, 4699
PPVT	159.65	12.81	136, 188	175.92	9.87	155, 194
DART	12.29	7.87	-1.27, 44.29	29.13	13.22	4.13, 65.40
Reading behaviour sum score	73.41	11.1	51, 99	84.83	10.16	53, 108
Reading experience factor	-0.97	1.06	-3.28, 1.66	0.96	1.19	-2.15, 3.46

Note. Accuracy is only based on responses to words, nonwords were excluded. Reaction times are based on correct responses to words and nonwords.

Table 3.2: Correlations between average accuracy on the word recognition test, DART, PPVT, reading behaviour factor score and reading experience factor score.

	Accuracy	PPVT	DART	Reading behaviour sum score
Accuracy	1			
PPVT	.39**	1		
DART	.29**	.61**	1	
Reading behaviour sum score	.37**	.60**	.59**	1
Reading experience factor	.41**	.86**	.86**	.85**

Note. * $p < .05$, ** $p < .001$.

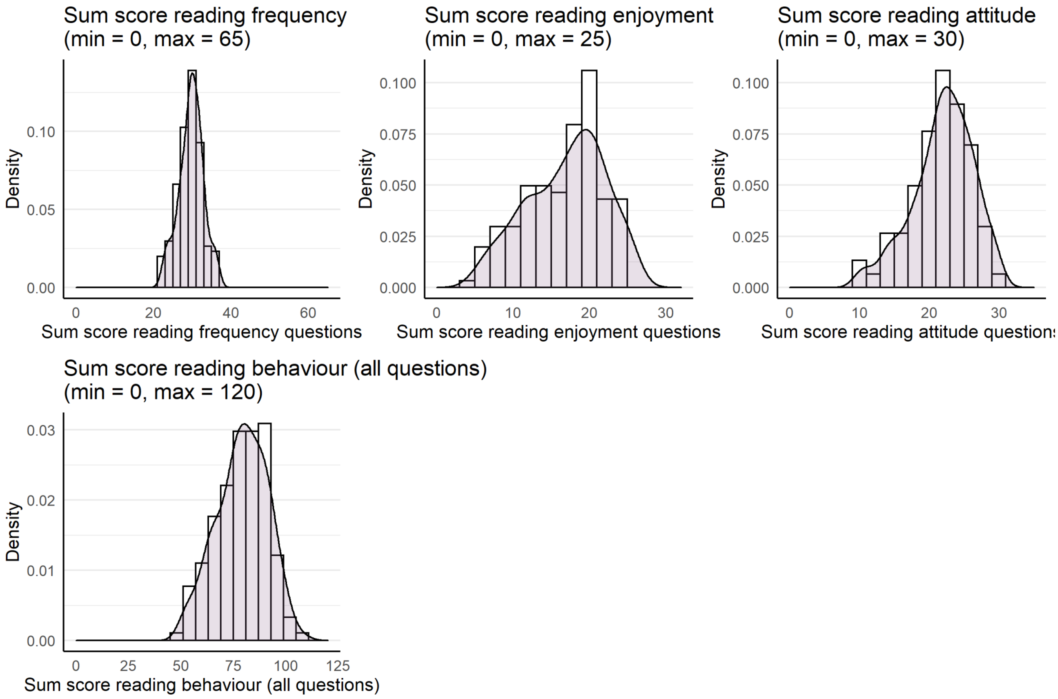


Figure 3.1: Histograms and density plots of the sum scores of the reading behaviour questionnaire. Top row: sum scores split by different domains of the questionnaire. Score limits are provided for each domain and the overall sum score.

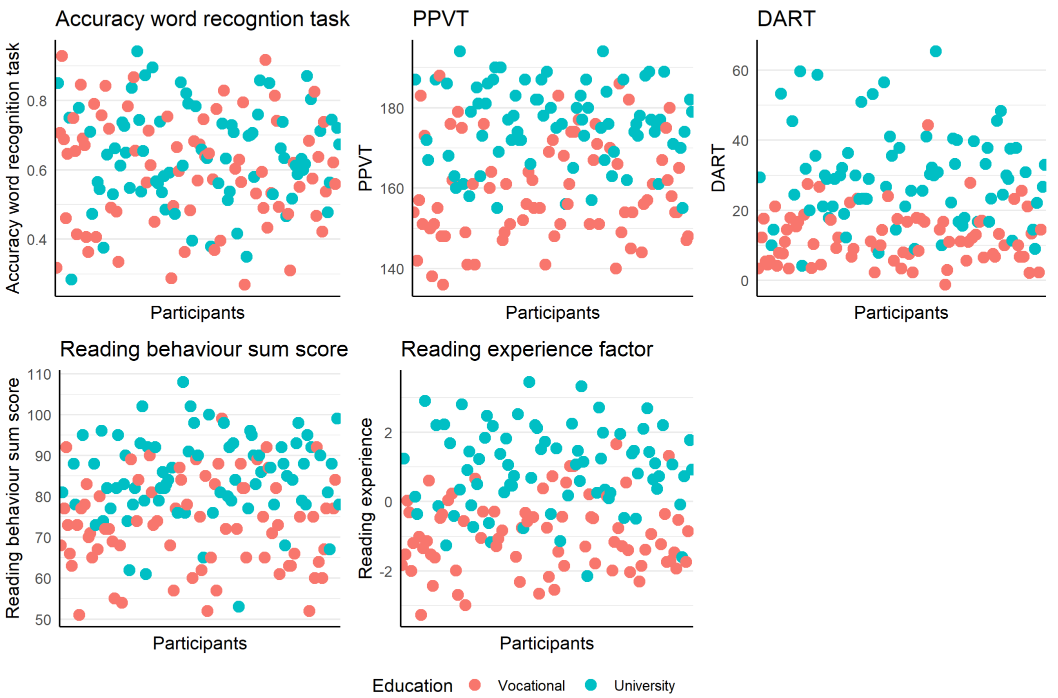


Figure 3.2a: Scatterplots of participant's scores on the variables of the present study, split by education level.

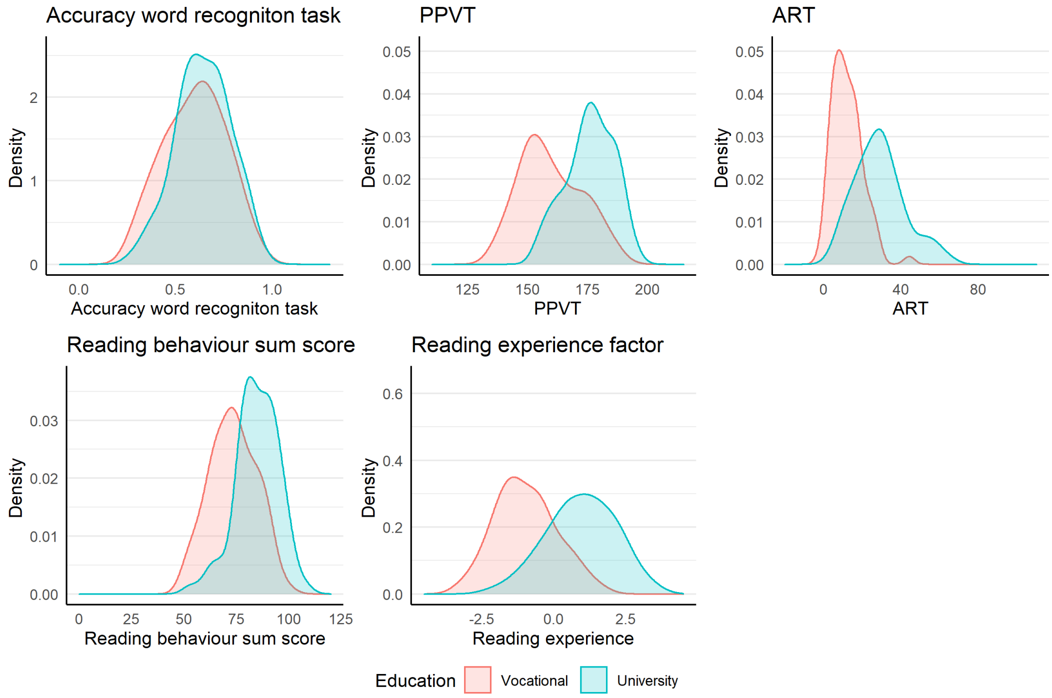


Figure 3.2b: Histograms of participant's scores on the variables of the present study, split by education level.

Word recognition test.

Accuracy. Average accuracy for word decisions was .62 (SD = 0.15), with little differences between the three modalities (Table 3.3a) or education levels (Table 3.3b, Figure 3.3a and 3.3b, top left panel). Participants provided more accurate responses as word difficulty decreased (Figure 3.3a), and this pattern was visible across modalities and education levels (Figure 3.3b). Accuracy for difficult words was lower in the spoken compared to written and audio-visual modality, particularly for vocational students (Figure 3.3b). False alarm rate (i.e., the proportion of "I know this word" to nonwords) was, on average, 15% (SD = 15%, range = 0% – 70%). Miss rate (i.e., the proportion "I don't know this word" to existing Dutch words) was 38% (SD = 15%, range = 0.0% – 73%). In general, participants found it easier to recognize the nonwords (high correct rejection and low false alarm rates) than to recognize the real Dutch words (Figure 3.4a). This pattern was visible for all three modalities and did not differ for participants enrolled in university or vocational education (Figure 3.4b), which motivated our decision to perform accuracy analyses combining the two groups of participants rather than performing separate analyses for each group.

Table 3.3a: Participants' average accuracy (proportion) on the word recognition test for all trials combined, and split by modality.

	Mean	SD	Range	Skewness	Kurtosis
Overall	0.62	0.15	0.27, 0.94	-0.17	-0.59
Spoken	0.60	0.16	0.20, 0.94	-0.16	-0.56
Audio-visual	0.62	0.16	0.24, 0.98	-0.13	-0.55
Written	0.64	0.16	0.22, 0.95	-0.41	-0.39

Note. Accuracy is only based on responses to words, nonwords were excluded.

Table 3.3b: Participants' average accuracy (proportion) on the word recognition test for all trials combined, and split by modality and education level.

	Vocational education			University		
	Mean	SD	Range	Mean	SD	Range
Overall	0.6	0.16	0.27, 0.93	0.64	0.14	0.28, 0.94
Spoken	0.58	0.18	0.20, 0.94	0.62	0.14	0.26, 0.91
Audio-visual	0.61	0.16	0.24, 0.96	0.64	0.16	0.29, 0.98
Written	0.61	0.17	0.24, 0.91	0.67	0.15	0.22, 0.95

Note. Accuracy is only based on responses to words, nonwords were excluded.

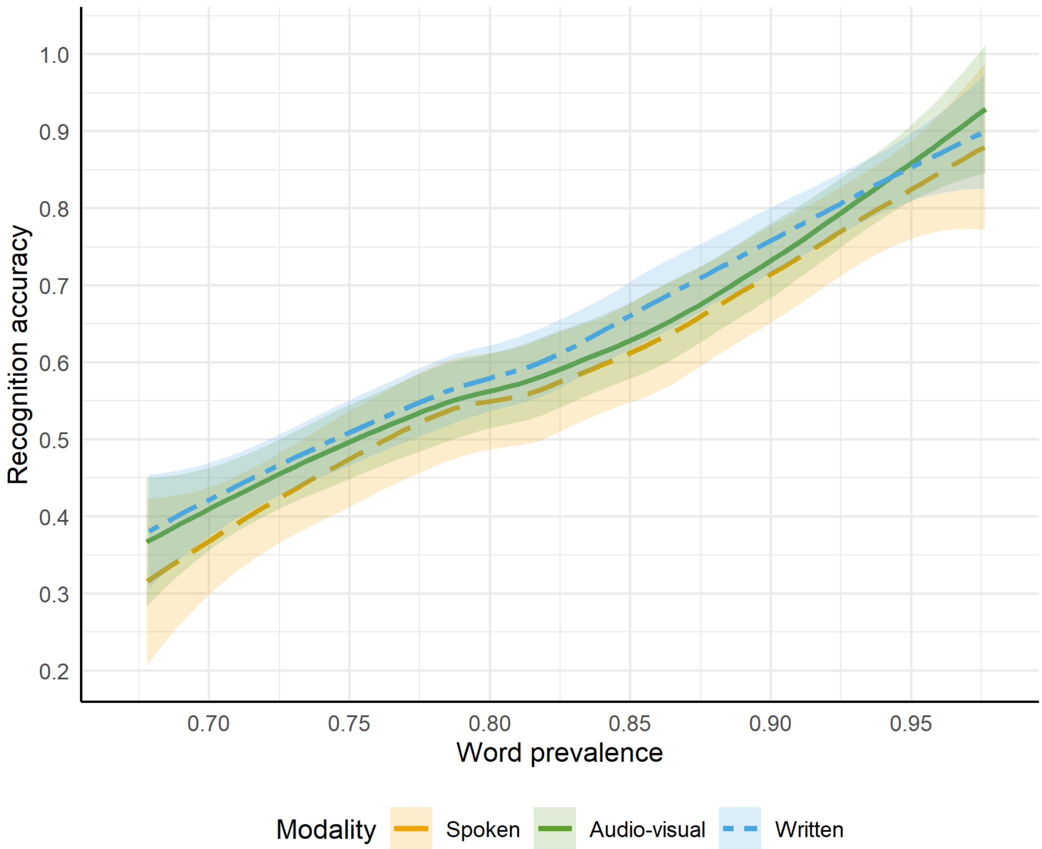


Figure 3.3a: Smoothed (using loess regression) word recognition accuracy by word difficulty (each step is a 2% increase of word prevalence), split by presentation modality. Bands indicate 95% confidence intervals around the predicted values of the loess regression.

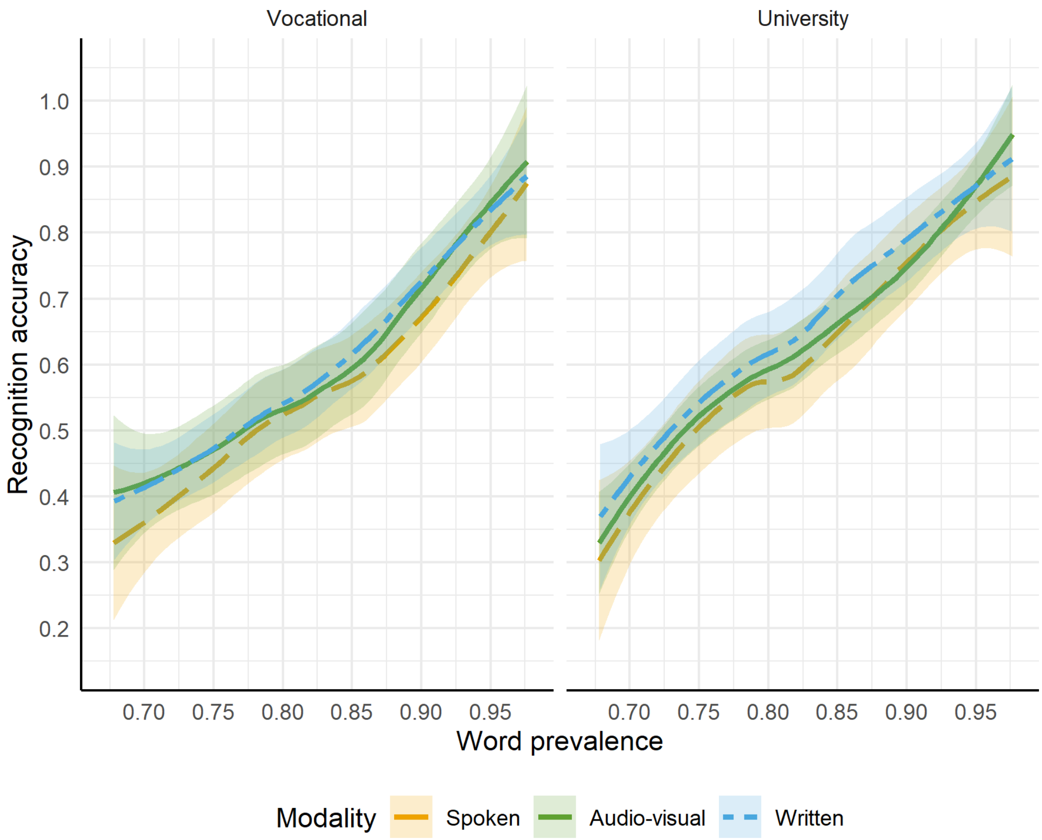


Figure 3.3b: Smoothed (using loess regression) word recognition accuracy by word difficulty (each step is a 2% increase of word prevalence), split by education level and presentation modality. Bands indicate 95% confidence intervals around the predicted values of the loess regression.

Reaction times. RT analyses were split into responses to words (83.37% of the RT data) and nonwords (16.63% of the RT data). Furthermore, only reaction times to correct responses were analysed (62.07% of the total responses to words and 85.05% of the total responses to nonwords). Participants responded on average at 1208 ms (Table 3.4a) to words and 1170 ms to nonwords (Table 3.4b). For both words and nonwords, participants responded faster in the written (words: 1039 ms, nonwords: 1013 ms) and audio-visual modality (words: 1087 ms, nonwords: 1046 ms) than the spoken modality (words: 1509 ms, nonwords: 1448 ms). This pattern is displayed in Figure 3.5a. The black line represents the mean offset time of the (visually and/or auditorily presented) stimulus. The pattern suggests that in the spoken modality, participants only started responding after the spoken stimulus had ended, but that in the written and audio-visual modalities, participants already respond before the stimulus

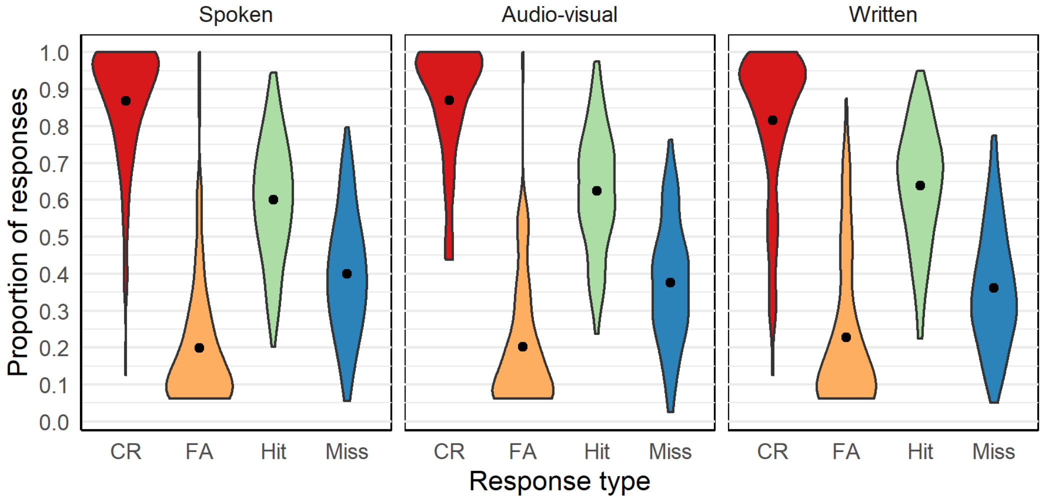


Figure 3.4a: Proportion of correct rejections and false alarms for nonwords, and hits and misses for words by modality. CR = correct rejection, FA = false alarm.

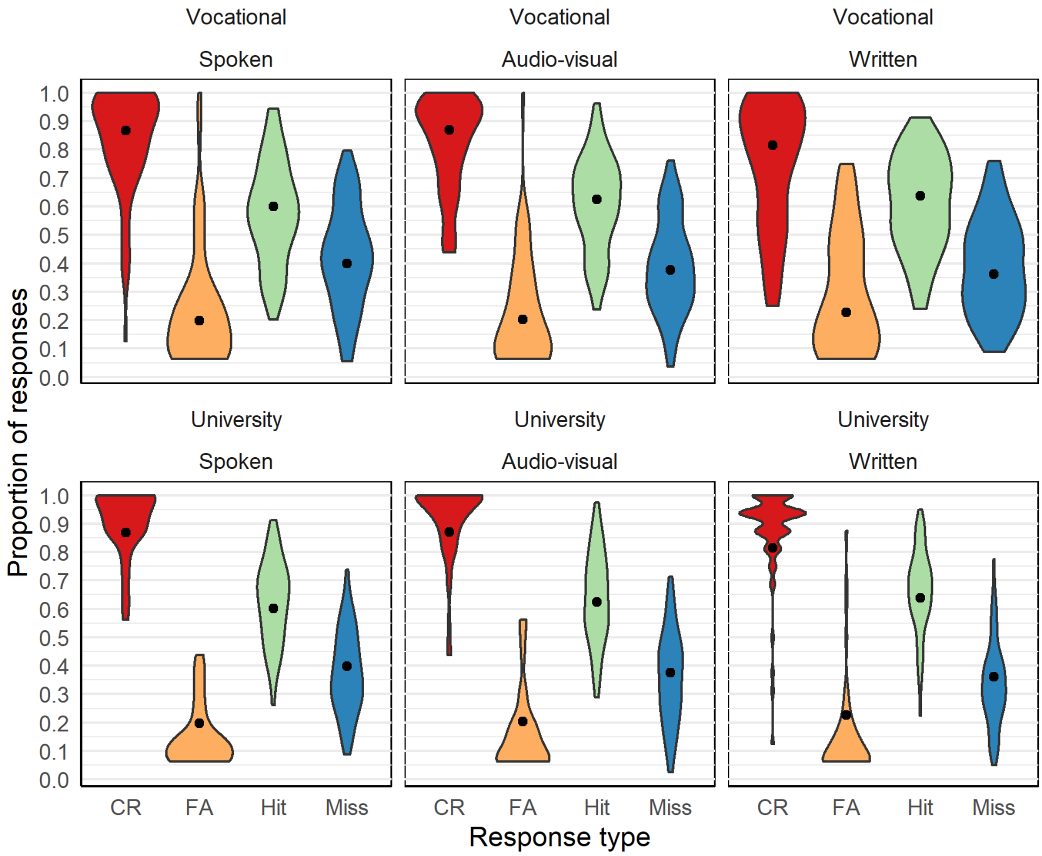


Figure 3.4b: Proportion of correct rejections and false alarms for nonwords, and hits and misses for words by modality, split by education level. CR = correct rejection, FA = false alarm.

disappeared. The RT patterns was the same for participants enrolled in university and in vocational education (Table 3.4c for words, Table 3.4d for nonwords, Figure 3.5b). This led us to combine the two groups of participants in one analysis.

Table 3.4a: Participants' average RTs (ms) of correct responses to words on the word recognition test, split by modality.

	Mean	SD	Range	Skewness	Kurtosis
Overall	1207.63	545.52	301, 4946	2.06	7.16
Spoken	1508.85	524.15	313, 4946	2.37	9.48
Audio-visual	1087.43	466.43	205, 4901	2.61	11.17
Written	1039.19	519.34	301, 4890	2.52	9.56

Table 3.4b: Participants' average RTs (ms) of correct responses to nonwords on the word recognition test, split by modality.

	Mean	SD	Range	Skewness	Kurtosis
Overall	1170.18	488.88	305, 4926	2.37	9.70
Spoken	1446.70	488.38	348, 4926	2.69	10.89
Audio-visual	1045.80	401.52	305, 4588	3.04	15.66
Written	1012.89	448.60	356, 4890	2.86	13.05

Table 3.4c: Participants' average RTs (ms) of correct responses to words on the word recognition test, split by modality and education level.

	Vocational education			University		
	Mean	SD	Range	Mean	SD	Range
Overall	1242.87	577.06	301, 4946	1175.72	513.22	345, 4904
Spoken	1515.81	562.44	313, 4946	1502.54	486.92	729, 4904
Audio-visual	1135.26	500.30	305, 4901	1043.21	428.11	371, 4894
Written	1088.78	570.89	301, 4890	955.10	464.40	345, 4818

Table 3.4d: Participants' average RTs (ms) of correct responses to nonwords on the word recognition test, split by modality and education level.

	Vocational education			University		
	Mean	SD	Range	Mean	SD	Range
Overall	1231.62	540.1	305, 4926	1117,86	433.87	462, 4699
Spoken	1471.37	545.53	348, 4926	1424.98	430.93	807, 4699
Audio-visual	1127.09	461.30	305, 4588	976.06	326.45	462, 3609
Written	1084.16	525.16	356, 4890	954.78	364.88	487, 4154

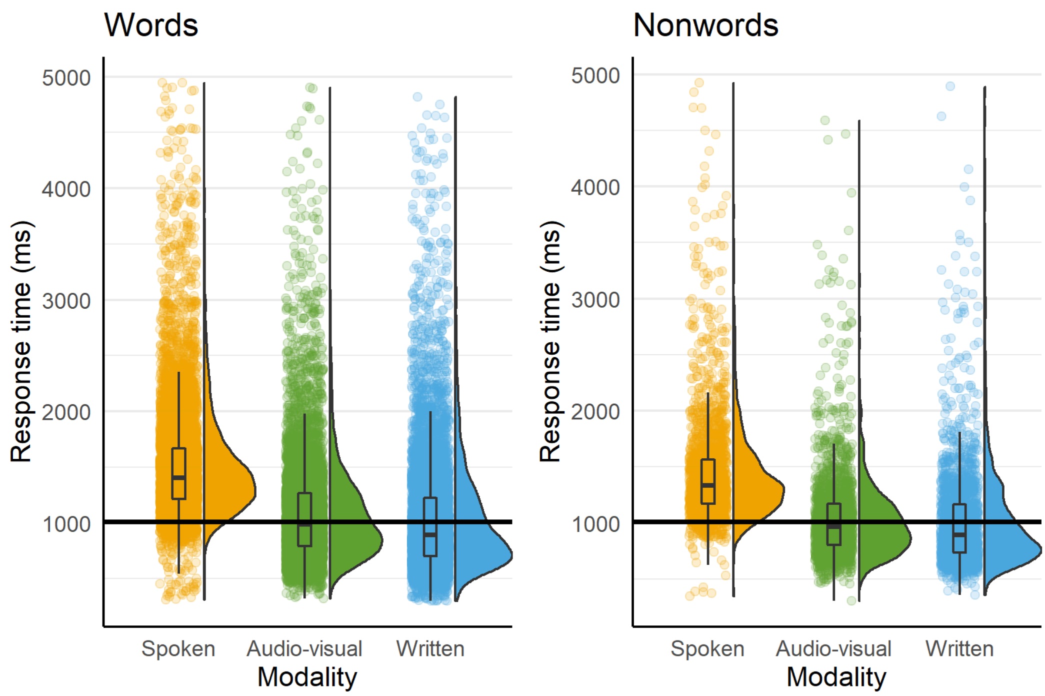


Figure 3.5a: RT distribution for correctly recognized words and nonwords, split by presentation modality. The black vertical line indicates mean offset of stimulus.

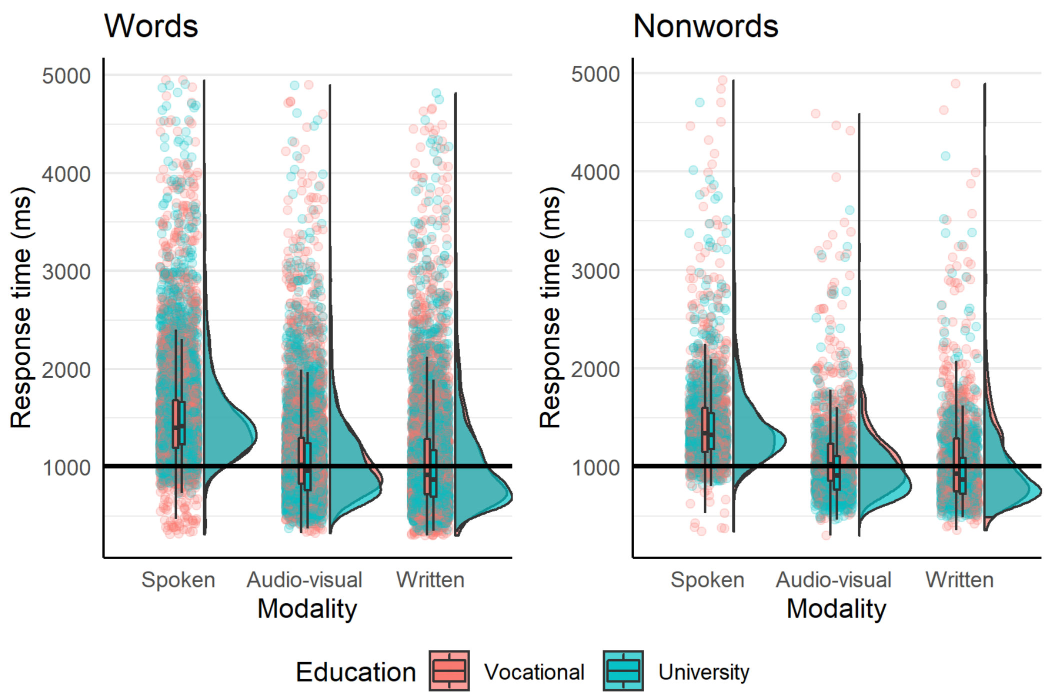


Figure 3.5b: RT distribution for correctly recognized words and nonwords, split by presentation modality and education level. The black vertical line indicates mean offset of stimulus.

Modelling

Data were analysed with Bayesian logistic or linear mixed-effects models in R (R Development Core Team, 2008, version 3.6.2), using the brms package (Bürkner, 2017). Bayesian analyses consider the likely magnitude of an effect rather than statistical significance. We considered effects to be meaningful when the 95% CIs (CI) did not contain zero. This indicates that there is high certainty that the parameter has a non-zero effect. To give an indication of the relative evidence for the alternative hypothesis (H1) compared to the null hypothesis (H0) or vice versa, Bayes Factors (BF10, BF01) were calculated for all effects. BF10 indicates a Bayes factor that favours H1 over H0, and BF01 indicates a Bayes factor in favour of H0 over H1. To interpret Bayes Factors, we followed the guidelines by Jeffreys (1961). A BF of 1 – 3 can be interpreted as anecdotal evidence, a BF of 3 – 10 as substantial evidence and a BF of > 10 of strong evidence for or against the null/alternative hypothesis.

The models had four chains of 8000 iterations each, with the first half being a warm-up period. As is appropriate for non-hierarchical logistic regression models, a weak prior (Cauchy distribution with center 0 and scale 2.5 using a sampling algorithm) was used (Gelman et al., 2008). Models were run until the $\hat{R}$ value for each parameter was 1.00, indicating full convergence.

Accuracy. To analyse the accuracy data, Bayesian mixed-effects models in R were performed. To enable full-factorial comparisons, two models with different intercepts were run. In both models, only responses to words were analysed. Modality was contrast-coded based on simple contrasts. Simple contrast coding is similar to treatment coding, but the intercepts correspond to the grand mean rather than the mean of the reference level. The spoken modality was the reference level in the first model and the audio-visual modality in the second model. The models included Modality as a fixed factor. The predictor Word Difficulty was scaled and centered. The Reading Experience factor was also added as a predictor, as were interactions between these three variables. The random effect structure included random intercepts by word and participant and random slopes for modality by word and participant. The final model was thus: brm(Correct ~ (modality * cWord_Difficulty * Reading_Experience_F) + (1 + modality | Uuid) + (1 + modality | word), family = bernoulli, data = all_Data, chains = 4, cores = 2, iter = 8000, prior = Pr1).

Table 3.5a and 3.5b display the model output of the model with the spoken and audiovisual modality as reference levels respectively. We observed evidence for a main effect of Modality: recognition accuracy was lower in the spoken than audio-visual and written modalities, but accuracy in the audio-visual modality did not differ from the written modality. However, the Bayes factors suggested the evidence for the recognition benefit of the audio-visual compared to spoken modality was almost non-existent (BF10 = 0.26 – 0.31). Instead, there was substantial evidence for the absence of an accuracy difference between the audio-visual and spoken modality (BF01 = 3.23 – 3.85). With regard to the accuracy difference between the spoken and written modality, evidence was inconclusive, with anecdotal evidence for (BF10 = 1.04) and almost anecdotal evidence against (BF01 = 0.96) the existence of such a difference.

Extremely strong evidence was found for a main effect of Word Difficulty, indicating that as words became more prevalent, they were recognized more accurately. Extremely strong evidence was also found for an effect of Reading Experience and suggested that participants with more reading experience were more accurately recognizing words. Finally, very strong evidence was found for an interaction between Word Difficulty and Reading Experience. Figure 3.6 suggests that the accuracy difference between experienced readers and experienced readers is larger for words of high to medium difficulty, with a 20% accuracy difference for words with a prevalence of .75, than for easy words, with a 10% accuracy difference for words with a prevalence of .95.

Table 3.5a: Full model output for the model with the spoken modality as reference level.

Predictor	Estimate	SE	95% CI	BF10	BF01
(Intercept)	**0.76**	**0.09**	**0.58, 0.93**	**1.51e+06**	**6.62e-07**
Modality: AV	***0.14***	***0.06***	***0.02, 0.25***	***0.31***	***3.23***
Modality: W	***0.19***	***0.07***	***0.06, 0.33***	***1.04***	***0.96***
Word Difficulty	**1**	**0.06**	**0.88, 1.11**	**2.27e+20**	**4.40e-21**
Reading Experience factor	**0.27**	**0.05**	**0.18. 0.37**	**6997.81**	**0**
Modality: AV * Word Difficulty	0.02	0.05	-0.08, 0.12	0.02	55.56
Modality: W * Word Difficulty	-0.05	0.06	-0.16, 0.07	0.03	38.46
Modality: AV * Reading Experience factor	-0.02	0.03	-0.08, 0.04	0.01	83.33
Modality: W * Reading Experience factor	0.01	0.03	-0.06, 0.08	0.01	83.33
Word Difficulty * Reading Experience factor	**0.05**	**0.01**	**0.03, 0.07**	**188.86**	**0.01**
Modality: AV * Word Difficulty * Reading Experience factor	0.03	0.02	-0.01, 0.08	0.02	50
Modality: W * Word Difficulty * Reading Experience factor	0.01	0.02	-0.04, 0.06	0.01	125

Note. Meaningful effects are displayed in bold. Effects that seem to be present according to the 95% CI, but for which evidence is non-existent or inconclusive as indicated by BFs are italicized. AV = audio-visual, W = written.

Table 3.5b: Full model output for the model with the audio-visual modality as reference level.

Predictor	Estimate	SE	95% CI	BF10	BF01
(Intercept)	**0.75**	**0.09**	**0.58, 0.94**	**5.38e+06**	**1.86e-07**
Modality: S	***-0.14***	***0.06***	***-0.25, -0.02***	***0.26***	***3.85***
Modality: W	0.06	0.05	-0.05, 0.16	0.03	33.33
Word Difficulty	**1**	**0.06**	**0.89, 1.11**	**2.59e+16**	**3.86e-17**
Reading Experience factor	**0.27**	**0.05**	**0.18, 0.36**	**4283**	**0.0002**
Modality: S * Word Difficulty	-0.02	0.05	-0.12, 0.09	0.02	50
Modality: W * Word Difficulty	-0.06	0.05	-0.15, 0.03	0.04	25
Modality: S * Reading Experience factor	0.02	0.03	-0.04, 0.08	0.01	100
Modality: W * Reading Experience factor	0.03	0.03	-0.03, 0.10	0.02	50
Word Difficulty * Reading Experience factor	**0.05**	**0.01**	**0.03, 0.07**	**566.94**	**0.002**
Modality: S * Word Difficulty * Reading Experience factor	-0.03	0.02	-0.08, 0.02	0.02	50
Modality: W * Word Difficulty * Reading Experience factor	-0.02	0.02	-0.07, 0.02	0.01	100

Note. Meaningful effects are displayed in bold. Effects that seem to be present according to the 95% CI, but for which evidence is non-existent or inconclusive as indicated by BFs are italicized. S = spoken, W = written.

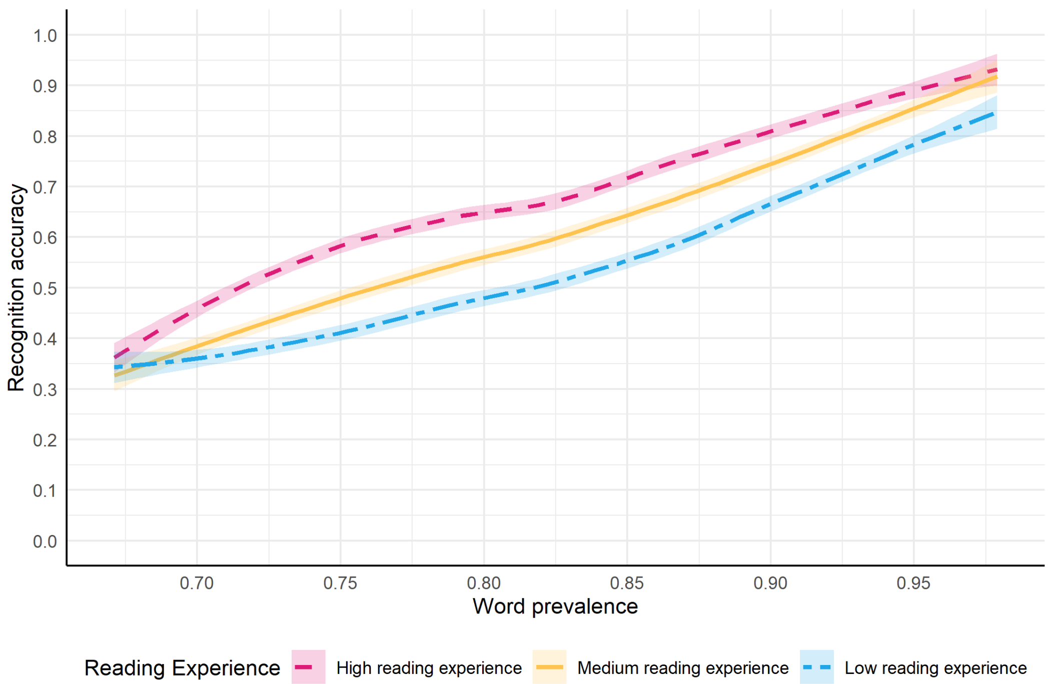


Figure 3.6: Smoothed (using loess regression) accuracy by word difficulty averaged over modality, split by reading experience. Bands indicate 95% confidence intervals around the predicted values of the loess regression.

Reaction times. In order to investigate the RT data, two Bayesian linear mixed-effects models in R were performed on the RTs of correct responses to words (62.0% of the word trials). For the first model, the spoken modality was used as the reference level, and for the second model the audio-visual modality was the reference model. In addition, RTs of correct responses to nonwords (85.05% of the nonword trials) were modelled to examine word-level decisions in the absence of lexical activation. RTs were inverse-transformed prior to analysis to decrease the noise associated with skewed RT data (Brysbaert & Stevens, 2018). The models included scaled and centered Word Difficulty and the Reading Experience factor. The models also included the continuous, centred and scaled variable Stimulus Duration to account for the effect of presentation duration on RTs. Finally, the models included a baseline measure of RT that would account for the differences in presentation delays between the three modalities. In online testing environments, particularly auditory presentation is delayed with a variable amount of time (40 – 60 ms) due to technical characteristics of the participant's machine and/or internet connection. For the models of RTs to words, we calculated for each participant individually their average inverse-transformed RTs to nonwords presented in the written, spoken and audio-visual modality separately. For the models of RTs to nonwords, we calculated these baselines for each individual based on their inverse-transformed RTs to words, for each modality separately. In the models, this individual, modality-specific RT baseline would capture the variance in RTs associated with presentation delays, that varied between participants and presentation modality. The random effect structure included random intercepts by word and participant. The model formula was: brm (invRT_words ~ (cWord_ Difficulty * Reading_Experience_F * cStimulus_Duration + Baseline_RTs_nonwords) + (1 | Uuid) + (1 | word), data=all_Data_words, chains = 4, cores = 2, iter = 8000, prior = Pr1). The models for nonwords were the same to the models for words, except for the removal of the Word Difficulty variable.

RTs of correct responses to words Output of the two models is displayed in Table 3.6a and Table 3.6b. Very strong evidence was observed for a main effect of Modality: RTs were lower in the spoken compared to audio-visual modality and in the spoken compared to written modality (M and SD for each modality separate are displayed in Table 3.4a). Although 95% CI's suggest a difference in RTs between the written and audio-visual modality, the evidence was non-existent (BF10 = 0.3). Strong evidence was observed for a main effect of Word Difficulty in that RTs were lower as words became more prevalent. Strong evidence for an interaction between Word Difficulty and the written modality was found, suggesting that the word difficulty effect (lower RTs as words became more prevalent) was particularly large in the written compared to spoken modality (Figure 3.7). Although 95% CI's suggested that this interaction was also present for the contrast audio-visual vs. spoken modality, and the contrast audiovisual vs. written modality, evidence for these interactions was non-existent (BF10 = 0.04 – 0.65 and 0.21 respectively). Finally, there was anecdotal (model with spoken modality as reference level) and substantial (model with audio-visual modality as reference level) evidence for an interaction between Word Difficulty and Reading Experience (BF10 = 1.72 – 5.69). This effect indicated that the Word Difficulty effect (lower RTs as words became more prevalent) was larger for individuals who were experienced readers (Figure 3.8).

Table 3.6a: Full model output for the RT model of correct responses to words, with the spoken modality as the reference level.

Predictor	Estimate	SE	95% CI	BF10	BF01
(Intercept)	**-0.22**	**0.04**	**-0.29, -0.15**	**556.02**	**0**
Modality: AV	**-0.09**	**0.02**	**-0.12, -0.06**	**194.29**	**0.01**
Modality: W	**-0.14**	**0.02**	**-0.17, -0.10**	**1487.34**	**0**
Word Difficulty	**-0.04**	**0.01**	**-0.05, -0.03**	**315.71**	**0**
Reading Experience factor	0	0.01	-0.02, 0.01	0	500
Stimulus duration	0.01	0.01	0.00, 0.03	0.01	142.86
Baseline RTs	**0.73**	**0.03**	**0.67, 0.80**	**6.09e+22**	**1.64e-23**
Modality: AV * Word Difficulty	***-0.02***	***0.01***	***-0.03, -0.01***	***0.04***	***24.39***
Modality: W * Word Difficulty	**-0.04**	**0.01**	**-0.06, -0.03**	**297.36**	**0**
Modality: AV * Reading Experience factor	-0.01	0.01	-0.02, 0.01	0	333.33
Modality: W * Reading Experience factor	-0.01	0.01	-0.03, 0.01	0	250
Word Difficulty * Reading Experience factor	**-0.01**	**0**	**-0.01, -0.00**	**1.72**	**0.58**
Modality: AV * Stimulus duration	0	0.01	-0.02, 0.02	0	333.33
Modality: W * Stimulus duration	-0.02	0.01	-0.04, 0.00	0.02	66.67
Word Difficulty * Stimulus duration	0	0.02	-0.03, 0.03	0.01	200
Reading Experience factor * Stimulus duration	0	0	0.00, 0.00	6.92e-04	1444.88
Modality: AV * Word Difficulty * Reading Experience factor	0	0	-0.01, 0.01	8.88e-04	1126.63
Modality: W * Word Difficulty * Reading Experience factor	-0.01	0	-0.01, 0.00	0.01	125
Modality: AV * Word Difficulty * Stimulus duration	-0.03	0.02	-0.07, 0.01	0.02	55.56
Modality: W Word Difficulty * Stimulus duration	-0.03	0.02	-0.08, 0.01	0.02	52.63
Modality: AV * Reading Experience factor * Stimulus duration	0	0.01	-0.01, 0.01	0	500
Modality: W * Reading Experience factor * Stimulus duration	0	0	-0.01, 0.01	0	500
Word difficulty * Reading Experience factor * Stimulus duration	0	0	0.00, 0.01	0	333.33
Modality: AV * Word Difficulty * Reading Experience factor * Stimulus duration	0.02	0.01	0.00, 0.04	0.02	50
Modality: W * Word Difficulty * Reading Experience factor * Stimulus duration	0.01	0.01	-0.01, 0.03	0	250

Note. Meaningful effects are displayed in bold. Effects that seem to be present according to the 95% CI, but for which evidence is non-existent or inconclusive as indicated by BFs are italicized. AV = audio-visual, W = Written.

Table 3.6b: Full model output for the RT model of correct responses to words, with the audio-visual modality as the reference level.

Predictor	Estimate	SE	95% CI	BF10	BF01
(Intercept)	**-0.22**	**0.04**	**-0.29, -0.15**	**1837.74**	**0**
Modality: S	**0.09**	**0.02**	**0.06, 0.12**	**407.86**	**0**
Modality: W	***-0.04***	***0.01***	***-0.07, -0.01***	***0.3***	***3.30***
Word Difficulty	**-0.04**	**0.01**	**-0.05, -0.03**	**616.38**	**0**
Reading Experience factor	0	0.01	-0.02, 0.01	0	333.33
Stimulus duration	0.01	0.01	0.00, 0.03	0.01	142.86
Baseline RTs	**0.73**	**0.03**	**0.67, 0.80**	**6.54e+22**	**1.53e-23**
Modality: S Word Difficulty	***0.02***	***0.01***	***0.01, 0.03***	***0.65***	***1.54***
Modality: W * Word Difficulty	***-0.02***	***0.01***	***-0.03, -0.01***	***0.21***	***4.72***
Modality: S * Reading Experience factor	0.01	0.01	-0.01, 0.02	0	333.33
Modality: W * Reading Experience factor	0	0.01	-0.02, 0.02	0	333.33
Word Difficulty * Reading Experience factor	**-0.01**	**0**	**-0.01, -0.00**	**5.69**	**0.18**
Modality: S * Stimulus duration	0	0.01	-0.02, 0.02	0	250
Modality: W * Stimulus duration	-0.02	0.01	-0.04, 0.00	0.37	2.70
Word Difficulty * Stimulus duration	0	0.02	-0.03, 0.03	0.01	200
Reading Experience factor * Stimulus duration	0	0	0.00, 0.00	7.23e-04	1383.13
Modality: S * Word Difficulty * Reading Experience factor	0	0	-0.01, 0.01	0	1000
Modality: W * Word Difficulty * Reading Experience factor	-0.01	0	-0.01, 0.00	0.01	83.33
Modality: S * Word Difficulty * Stimulus duration	0.03	0.02	-0.01, 0.07	0.02	52.63
Modality: W * Word Difficulty * Stimulus duration	0	0.02	-0.05, 0.04	0.01	142.86
Modality: S * Reading Experience factor * Stimulus duration	0	0.01	-0.01, 0.01	0	500
Modality: W * Reading Experience factor * Stimulus duration	0	0	-0.01, 0.01	0	500
Word difficulty * Reading Experience factor * Stimulus duration	0	0	0.00, 0.01	0	333.33
Modality: S * Word Difficulty * Reading Experience factor * Stimulus duration	-0.02	0.01	-0.04, 0.00	0.02	55.56
Modality: W * Word Difficulty * Reading Experience factor * Stimulus duration	-0.01	0.01	-0.03, 0.01	0.01	166.67

Note. Meaningful effects are displayed in bold. Effects that seem to be present according to the 95% CI, but for which evidence is non-existent or inconclusive as indicated by BFs are italicized. S = spoken, W = Written.

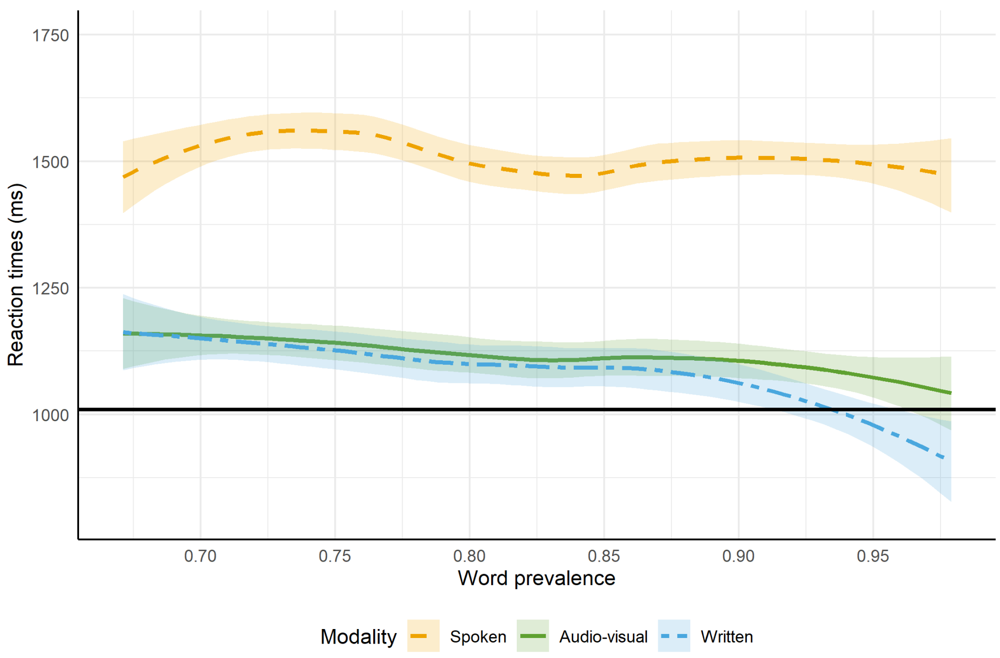


Figure 3.7: Smoothed (using loess regression) RTs for correctly recognized words by word difficulty, split by presentation modality. Bands indicate 95% confidence intervals around the predicted values of the loess regression. The black horizontal line indicates mean offset of stimulus.

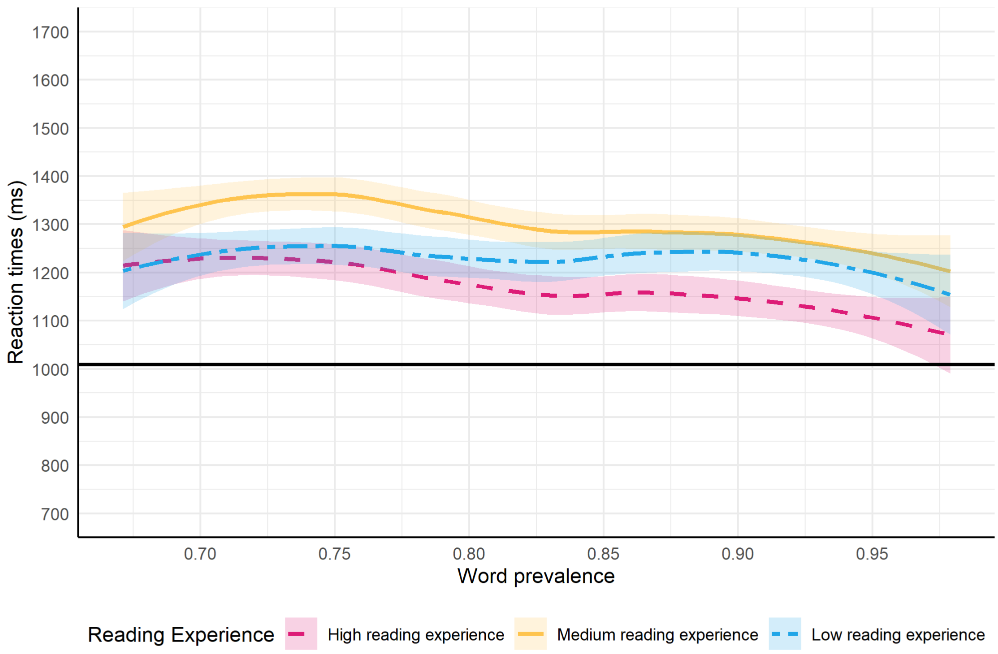


Figure 3.8: Smoothed (using loess regression) RTs for correctly recognized words by word difficulty, split by reading experience. Bands indicate 95% confidence intervals around the predicted values of the loess regression. The black horizontal line indicates mean offset of stimulus.

RTs of correct responses to nonwords The output of the two models on RTs to nonword data is displayed in Table 3.7a and Table 3.7b. Evidence was found for a main effect of Modality: compared to the spoken modality, responses to nonwords were faster in the audio-visual modality (M and SD for each modality are displayed in Table 3.4b). The 95% CI indicated that the effect was also present for the contrast spoken vs. written modality, but evidence was anecdotal (BF10 = 1.16). The presence of an interaction effect between Modality and Reading Experience was detected according to the 95% CI's, but the evidence for this effect was non-existent (BF10 = 0.13 – 0.15). The 95% CI's also indicated that there was an interaction between Modality and Stimulus Duration, but again the evidence for this effect was non-existent (BF10 = 0.56).

Table 3.7a: Full model output for the RT model of correct responses to words, with the spoken modality as the reference level.

Predictor	Estimate	SE	95% CI	BF10	BF01
(Intercept)	**-0.29**	**0.03**	**-0.36, -0.23**	**1.07e+08**	**9.34e-09**
Modality: AV	**-0.07**	**0.02**	**-0.1, -0.04**	**53.3**	**0.02**
Modality: W	***-0.07***	***0.02***	***-0.11, -0.03***	***1.16***	***0.86***
Reading Experience factor	0	0.01	-0.01, 0.01	0	500
Stimulus duration	0.02	0.01	0.00, 0.00	0.04	24.39
Baseline RTs	**0.71**	**0.03**	**0.65, 0.78**	**3.84e+20**	**2.60e-21**
Modality: AV * Reading Experience factor	***-0.02***	***0.01***	***-0.03, -0.01***	***0.15***	***6.85***
Modality: W * Reading Experience factor	-0.01	0.01	-0.02, 0.01	0	333.33
Modality: AV * Stimulus duration	-0.02	0.01	-0.04, 0.00	0.05	19.61
Modality: W * Stimulus duration	***-0.04***	***0.01***	***-0.07, -0.02***	***0.56***	***1.78***
Reading Experience factor * Stimulus duration	0	0	-0.01, 0.00	0	1000
Modality: AV * Reading Experience factor * Stimulus duration	0.01	0	0.00, 0.01	0	333.33
Modality: W * Reading Experience factor * Stimulus duration	0	0	-0.01, 0.01	0	500

Note. Meaningful effects are displayed in bold. Effects that seem to be present according to the 95% CI, but for which evidence is non-existent or inconclusive as indicated by BFs are italicized. AV = audio-visual, W = Written.

Table 3.7b: Full model output for the RT model of correct responses to words, with the audio-visual modality as the reference level.

Predictor	Estimate	SE	95% CI	BF10	BF01
(Intercept)	**-0.29**	**0.03**	**-0.36, -0.23**	**1.56e+09**	**6.42e-10**
Modality: S	**0.07**	**0.02**	**0.04, 0.10**	**78.8**	**0.01**
Modality: W	0	0.02	-0.03, 0.04	0.01	200
Reading Experience factor	0	0.01	-0.01, 0.01	0	500
Stimulus duration	0.02	0.01	0.00, 0.04	0.03	29.41
Baseline RTs	**0.71**	**0.03**	**0.65, 0.78**	**1.46e+22**	**6.86e-23**
Modality: S * Reading Experience factor	***0.02***	***0.01***	***0.01, 0.03***	***0.13***	***7.77***
Modality: W * Reading Experience factor	0.01	0.01	0.00, 0.03	0.01	76.92
Modality: S * Stimulus duration	0.02	0.01	0.00, 0.04	0.05	21.28
Modality: W * Stimulus duration	-0.02	0.01	-0.05, 0.01	0.01	71.43
Reading Experience factor * Stimulus duration	0	0	-0.01, 0.00	0	500
Modality: S * Reading Experience factor * Stimulus duration	-0.01	0	-0.01, 0.00	0	333.33
Modality: W * Reading Experience factor * Stimulus duration	0	0	-0.01, 0.00	0	500

Note. Meaningful effects are displayed in bold. Effects that seem to be present according to the 95% CI, but for which evidence is non-existent or inconclusive as indicated by BFs are italicized. S = spoken, W = Written.

3.4 Discussion

The present study investigated the relationships between presentation modality, word difficulty and reading experience in a sample featuring variation in reading experience. Participants with diverse educational backgrounds carried out a lexical decision task where nonwords and words varying in difficulty were presented in the written, spoken or audio-visual modality. Their reading experience was approximated using an objective measure of receptive vocabulary (PPVT) and an objective and subjective measure of reading behaviour (DART and reading behaviour questionnaire respectively).

The first aim of this study was to examine how presentation modality affected recognition of words varying in difficulty. With regard to word recognition accuracy, we predicted an interaction effect between presentation modality and word difficulty in that for more difficult words, accuracy would be higher in the written and audio-visual compared to the spoken modality. Written text contains more complex language than speech, which would increase the quality of orthographic representations and facilitate recognition in the written and audio-visual compared to spoken modality. No evidence was found for this interaction between modality and word difficulty, nor conclusive evidence for a main effect of presentation modality.

With regard to reaction times, we formulated two contrasting hypotheses. We assumed an effect of word difficulty on RTs, such that RTs were slower for difficult than easy words. We either expected this RT difference to be smaller in the written and audio-visual than spoken modality, or larger. Written text contains more difficult words, which would improve orthographic representations of these difficult words and enhance recognition of their written forms. This would decrease the RT difference between easy and difficult words in the written and audio-visual modality, but not in the spoken modality. In contrast, studies (Ferrand et al., 2018; Hasenäcker et al., 2019) reported RT differences between difficult and easy words to be larger for words presented in the written than spoken modality. Our findings suggested that the effect of word difficulty was particularly present in the written modality, corroborating these earlier findings. Thus, it seems that, particularly in the written compared to spoken modality, easy words are recognized faster.

The second aim of the present study was to investigate the relationship between presentation modality and word difficulty in a diverse sample to see how the relationship between the aforementioned variables would be influenced by reading experience. With regard to accuracy, we predicted that particularly experienced readers would show a recognition benefit for difficult words in the written and audio-visual compared to spoken modality. Experienced readers will have encountered more difficult words in the written modality as a result of their extensive exposure to text, which would strengthen their orthographical representations of difficult words and would facilitate word recognition in the written and audio-visual over spoken modality. No evidence was observed for such three-way interaction. Strong evidence was observed for an interaction between word difficulty and reading experience, such that for experienced readers in particular, recognition accuracy was increased for words of medium-high difficulty.

Concerning reaction times, we expected that the interaction effect of word difficulty and presentation modality on RTs would increase with reading experience. There was substantial evidence for an interaction between word difficulty and reading experience, but no evidence that this interaction was modulated by modality. This suggests that in all modalities, the word difficulty effect (faster recognition of easy words) is larger for highly experienced compared to less experienced readers.

The accuracy results of the present study suggested that successful word recognition across a range of easy and difficult words is not substantially influenced by presentation modality. Although easier words are more often recognized correctly than difficult words, presentation modality does not affect this. These null-effects regarding modality underscore patterns observed in the previous study (Chapter 2). This suggests that presentation modality does not influence the success of word recognition in a sample that is homogeneous with regard to educational background (Chapter 2) or a sample that has a more heterogeneous educational background (present study). Moreover, the null-effect of modality is present despite task differences between the two studies: in the present study the word recognition test was timed, whereas it was untimed in the previous study, and in the present study the average difficulty of the words was lower than in the previous study (present study: M prevalence = .85, range = .67 – .98, previous study: M prevalence = .75, range = .60 – .90). Thus, combining the evidence from these two studies, it seems that presentation modality has little to no effect on the success of word recognition.

The findings of the present study, however, do suggest that the ease (or difficulty) with which words are recognized, as indicated by RTs, is influenced by presentation modality. In particular, recognition of easy words seems to be most strongly facilitated in the written modality compared to the spoken modality. This finding is in line with previous studies that found the word difficulty effect to be larger in the written than spoken modality (Ferrand et al., 2018; Hasenäcker et al., 2019). An explanation for a written speed benefit for easy words could be that easy words had fewer letters than difficult words, as word length effects are known to be larger in the written than spoken modality (Hasenäcker et al., 2019). For the words in the present study, however, word difficulty (prevalence) and word length were not correlated ($r = -0.07$, $p = .29$), indicating that easy words were not shorter than difficult words. Moreover, when modelling RTs with the same model as described in the Results section, but with Word Length added as an additional predictor, strong evidence against an interaction between word difficulty and word length was found (BF10 = 0.01, BF01 = 111.11). This suggests that the effect of word difficulty on RTs did not differ for short or long words. Thus, the speed benefit in the written modality for more easy words was unlikely to be caused by our easy words consisting of fewer letters than our difficult words.

A different explanation for the speed benefit in the written compared to spoken modality is that for easy, more common words, orthographic representations of words are of higher quality than phonological representations. Based on this assumption, we expected that the difference in quality between orthographic and phonological representations would be most apparent for difficult words, as written language contains more difficult words than spoken language (Cunningham & Stanovich, 1998; Hayes & Ahrens, 1988). The written modality, however, allows for more encounters with *common* words as well. Brysbaert, Stevens, et al. (2016) calculated that in the hypothetical situation where one would only encounter words through social interaction during their waking hours, they would be exposed to 32.000 different word tokens per day, whereas if one would only read during their waking hours, they would be exposed to 288.000 different word tokens. Thus, not only does the written modality allow for more exposure with difficult words, also more easy words are encountered in their written compared to spoken form. Therefore, the quality difference between orthographic and phonological representation is probably present for difficult, less commonly known words, as well as for easy, more common words.

The speed benefit in the written modality may only occur for easy words though, and not difficult words as well, because mechanisms that facilitate recognition speed may be most effective for high quality representations. In the end, the overall quality of the representations of difficult words will always be lower compared to easy words, simply because difficult words are encountered less often. Computational models of word recognition assume that as words are encountered more often (and their representations improve in quality), a representation's activation threshold lowers, resulting in quicker activation (Coltheart et al., 2001; McClelland & Rumelhart, 1981; Morton, 1969). Moreover, increased precision of lexical representations causes lexical competitors to be quickly inhibited during word recognition, thereby improving the speed and increasing the accuracy of word recognition (Andrews, 1997; Andrews & Hersch, 2010; Perfetti, 1992). These facilitatory and inhibitory mechanisms during the word recognition process may have a stronger influence on word recognition for more commonly known words, which have high quality lexical representations, than more difficult words, which representations are of lower quality, and might explain why our observed patterns are stronger for more common words than difficult words.

The present study suggests that reading experience plays an important role in the success of word recognition. Experienced readers showed an overall accuracy benefit compared to less experienced readers. The interaction between reading experience and word difficulty suggested that the difference in accuracy between experienced and less experienced readers was particularly high for words of medium to high difficulty. This finding underscores patterns observed in our previous study (Chapter 2). A mechanism that may explain why experienced readers compared to less experienced readers are more successfully accessing lexical representations of difficult words is that due to their heightened experience with written language, experienced readers fine-tune their lexical representations (cf. lexical tuning hypothesis, Castles, 1999; Castles et al., 2007) of difficult words in particular, as these are more common in written language than in everyday speech. The fact that we did not observe presentation modality to influence the interaction between reading experience and word difficulty may indicate that this fine-tuning process is not constrained by presentation modality. It seems that even though experienced readers are most likely to encounter these difficult words in their written form, they are able to not only sharpen their orthographic representations of these words, but their phonological representations as well, allowing them to successfully access the lexical representations of difficult words regardless of the modality in which they are presented. One can speculate that crossmodal recoding/activation, an integral aspect in many theories of word reading (Coltheart et al., 2001) and reading acquisition (Ehri, 1995; Shankweiler, 1999; Share, 1995), may explain why experienced readers are able to sharpen both their orthographic as well as their phonological representations of words most often encountered in their written form only. It may be possible that experienced readers in particular, who have acquired strong grapheme-phoneme correspondences, crossmodally recode/activate phonological representations when encountering the written form, whereas this may be less automatic in less experienced readers who tend to have less strong grapheme-phoneme correspondences. However, such explanation requires further investigation.

Our study also indicated that reading experience influences the ease (or difficulty) with which words are recognized: the word difficulty effect (faster recognition of easy words) was particularly large for highly experienced compared to less experienced readers. This corroborates earlier findings that reported experienced readers to be faster responders when recognizing words (Chateau & Jared, 2000; Lee, Seong, Choi, & Lowder, 2019; Mainz et al., 2017; Sears, Siakaluk, Chow, & Buchanan, 2008; Yap et al., 2012). The influence of reading experience was only observed in RTs to words, but not nonwords. This seems to suggest that reading experience influences the speed of word recognition, but not the speed of the decision process, since nonword recognition does not require access to lexical representations. We did not observe evidence that the influence of reading experience on word recognition was modulated by presentation modality. Thus, it seems that regardless of the modality in which words are encountered, recognition of easy words is faster for experienced compared to less experienced readers. This may indicate that even though experienced readers expose themselves to language in the written modality in particular, both their orthographic as well as phonological representations increase in quality, ultimately allowing them to access both types of representations quicker than less experienced readers.

In sum, our study suggests that presentation modality does not influence the success of word recognition, as words recognition accuracy did not differ for words presented in different modalities. Reading experience does seem to influence the success of word recognition as experienced readers showed a recognition benefit over less experienced readers for words of medium to high difficulty. The ease with which words were recognized was influenced by both presentation modality and reading experience. The speed benefit associated with easy words compared to difficult words was larger in the written than spoken modality and also larger for experienced compared to less experienced readers.

The present study has several theoretical implications. First, the results of the present and previous study (Chapter 2) suggest that modality of presentation does not influence success of word recognition, but does affect the *speed* with which these lexical representations are retrieved. It could be the case that this modality effect on RTs is related to the fact that processing of written stimuli is faster in the written than spoken modality, due to a word's constituents being available at once in the written modality, but unfolding over time in the spoken modality. However, this cannot be the only explanation for the influence of presentation modality on RTs, as we found evidence for this modality effect to interact with word difficulty. In the written but not spoken modality, easy words were recognized faster than difficult words. If the speed benefit was only related to the inherent processing speed benefit of the written modality, we would have observed the speed benefit in the written modality to be similar for all words, not only for easy words. In this case, we would only have observed a main effect of modality, but no interaction with word difficulty. Theories of word recognition may need to reconsider how presentation modality influences word recognition. In particular, additional specifications are needed with respect to which stages presentation modality may play a role and how modality effects may differ for words varying in their difficulty.

A second theoretical implication of the present study is that our results indicate that reading experience influences the success of word recognition and, for easy words, the speed of word recognition. This corroborates emergentist approached to language (Dąbrowska, 2018; Lieven, 2016) which assume that individual variation in language experience can explain variation in language-related skills (see Kidd et al., 2020, 2018, for overviews). Thus, experience with written language improves word recognition abilities. Importantly, our study reported evidence for measures of an individual's experience to language in the one specific modality, in this case the written modality, explaining variance in processing both written and spoken linguistic input. This suggests that language experience, encountered in a specific modality, improves language processing, regardless of modality.

The findings of the present study also have practical implications. First, the finding that presentation modality does not affect the success of word recognition, but does influence the speed of word recognition may be of practical use within the field of psychometrics. Presentation modality may not affect performance on an untimed measure of vocabulary, but presentation modality should be carefully considered when developing timed vocabulary tests. Second, our findings underline the importance of considering the influence of individual differences in (written) language experience on language processing, both in the research as well as in the societal context.

3.5 Appendix I - Words and their prevalence (prev)

Words that occurred in this study but not in the previous study are displayed in bold.

Word	Prev	Word	Prev	Word	Prev	Word	Prev
rederijker	0.67	ooglijk	0.75	viltig	0.83	**secondair**	0.91
beknorren	0.67	heterofilie	0.75	exotisme	0.83	**grafkapel**	0.91
loosheid	0.67	quotatie	0.75	polytheen	0.83	**ruigharig**	0.91
cholerisch	0.67	welvoeglijk	0.75	doffig	0.83	**oplading**	0.91
polemologie	0.67	deugddoend	0.76	harpenist	0.84	**klerenborstel**	0.91
totalitarisme	0.67	routeren	0.76	inlassing	0.84	**nawinter**	0.91
biduur	0.67	postuleren	0.76	wijduit	0.84	**trotsheid**	0.92
vlieden	0.68	primeren	0.76	loffelijk	0.84	**psychoot**	0.92
verzoeten	0.68	xenofilie	0.76	aftroggelarij	0.84	**benutting**	0.92
nestvlieder	0.68	omdoping	0.77	ritualist	0.84	**schimmelig**	0.92
convertor	0.68	zinsbedrog	0.77	fenomenologie	0.84	**trouwhartig**	0.92
stribbeling	0.69	rotskunst	0.77	geruim	0.84	**manoeuvreren**	0.92
peigeren	0.69	abdominaal	0.77	condoleantie	0.84	**sproetig**	0.92
reconversie	0.69	spiritist	0.77	verwijven	0.85	**rasdier**	0.92
endemie	0.69	schalm	0.77	onbemerkt	0.85	**reglement**	0.92
decagram	0.69	ordinaat	0.77	onberecht	0.85	**wegaanleg**	0.93
demissie	0.69	eedbreuk	0.78	afratelen	0.85	**soppig**	0.93
lobberig	0.69	homeostase	0.78	autogram	0.85	**zenuwlijder**	0.93
erving	0.69	endotherm	0.78	replicator	0.85	**klarinettist**	0.93
cytologie	0.69	carbolineum	0.78	afstuiven	0.86	**dommerd**	0.93
collectioneren	0.69	legalisme	0.78	interferentie	0.86	**hoogteligging**	0.93
serafine	0.70	paleografie	0.78	lofprijzend	0.86	**overmand**	0.93
wiegelen	0.70	exclusie	0.78	hallucinant	0.86	**uitdenken**	0.94
scharren	0.70	afkukelen	0.78	lymf	0.86	**royalist**	0.94
sloffig	0.70	versmachten	0.79	zinnelijk	0.87	**dienstbetoon**	0.94
agronomie	0.70	weeklacht	0.79	cyclisme	0.87	**echtbreuk**	0.94
relevatie	0.70	rasperig	0.79	admissie	0.87	**bekoeling**	0.94
ressorteren	0.71	spitsig	0.79	reductionisme	0.87	**bemodderd**	0.94
bezemklas	0.71	patroneren	0.79	onduldbaar	0.87	**verhevene**	0.94
zwemblaas	0.71	schrijnwerker	0.79	brosheid	0.87	**bespeling**	0.94
geschulpt	0.71	oplaaiing	0.79	plichtig	0.87	**kropsla**	0.95
utilitarisme	0.71	historiek	0.80	curatief	0.87	**herbivoor**	0.95
secretarie	0.71	verificateur	0.80	cilindrisch	0.87	**bouwhal**	0.95
nomadisme	0.71	biogeen	0.80	solutie	0.87	**doorweken**	0.95
bijtreden	0.71	kwetsing	0.80	welgeaard	0.87	**acuut**	0.95
sculpturaal	0.72	signalisatie	0.80	smakker	0.88	**intypen**	0.95
gewemel	0.72	radiatie	0.80	veeweide	0.88	**grijptang**	0.95
heliocentrisch	0.72	vermaking	0.80	linkerrij	0.88	**jutezak**	0.96
insolide	0.72	pylon	0.80	vlotweg	0.88	**immens**	0.96
beroemen	0.72	schuieren	0.80	thuisloos	0.88	**klamheid**	0.96
havist	0.73	promoting	0.80	opraapsel	0.88	**smullerij**	0.96
assertie	0.73	fundatie	0.81	eruitzien	0.88	**reflectief**	0.96
fantasme	0.73	meerderwaardig	0.81	mediaan	0.89	**vleien**	0.96
verevenen	0.73	poolshoogte	0.81	tinkelen	0.89	**gebaat**	0.96
reticulair	0.73	reformisme	0.81	propeller	0.89	**verbluft**	0.96
knoeper	0.73	willigen	0.81	verzeilen	0.89	**sluieren**	0.97
verderfenis	0.74	scenarist	0.81	uitschateren	0.89	**fietsrit**	0.97
kroezelig	0.74	lamelle	0.81	onromantisch	0.89	**lokaliteit**	0.97
plagiator	0.74	obsederen	0.81	kletserig	0.90	**alligator**	0.97
uitloven	0.74	sjokkerig	0.82	spijten	0.90	**aanheffen**	0.97
picturaal	0.74	modereren	0.82	baatzucht	0.90	**flirterig**	0.97
wanbesef	0.74	fluoresceren	0.82	bezaaiing	0.90	**overtroeven**	0.97
traverseren	0.74	homologie	0.82	**secretie**	0.90	**visgerei**	0.97
roezemoezig	0.74	alreeds	0.82	wilsgebrek	0.90	**appartement**	0.98
deputatie	0.74	empirist	0.82	erkentenis	0.90	**schel**	0.98
temporeel	0.74	stranding	0.82	**idiotisme**	0.90	**waterpoel**	0.98
taxonoom	0.75	keutelig	0.82	**onbekeerd**	0.90	**lawaaierig**	0.98
kosterij	0.75	knapperen	0.83	**vin**	0.90	**cabine**	0.98
exciteren	0.75	druksel	0.83	**onheilig**	0.90	**attribuut**	0.98
postaal	0.75	nijverig	0.83	**logee**	0.90	**begroeiing**	0.98

3.6 Appendix II - Reading questionnaire

How often do you read the following reading materials for your own enjoyment?

5-point scale: Never of almost never, A few times a year, A few times a month, A few times a week, Daily

1. Newspapers
2. Fiction (novels, stories, poetry)
3. Non-fiction (biographies, informative books, travel guides, self-help books)
4. Comic books
5. Magazines

How often do you engage in the following online reading activities (using internet or an app):

5-point scale: Never of almost never, A few times a year, A few times a month, A few times a week, Daily

1. Using an online dictionary or encyclopaedia (such as Wikipedia)
2. Taking part in online group discussions on online web fora
3. Searching practical information online (for example timetables, events, tips, recipes)
4. Searching online information to learn something about a certain topic
5. Reading e-mails
6. Chatting/texting
7. Reading news online

How much time each day do you read for your own enjoyment?

5-point scale: I do not read for my own enjoyment, 30 minutes or less each day, Between 30-60 minutes each day, 1-2 hours each day, More than 2 hours each day

Indicate the extent to which the following statements pertain to you.

5-point scale: (Strongly disagree, Disagree, Agree nor disagree, Agree, Strongly agree)

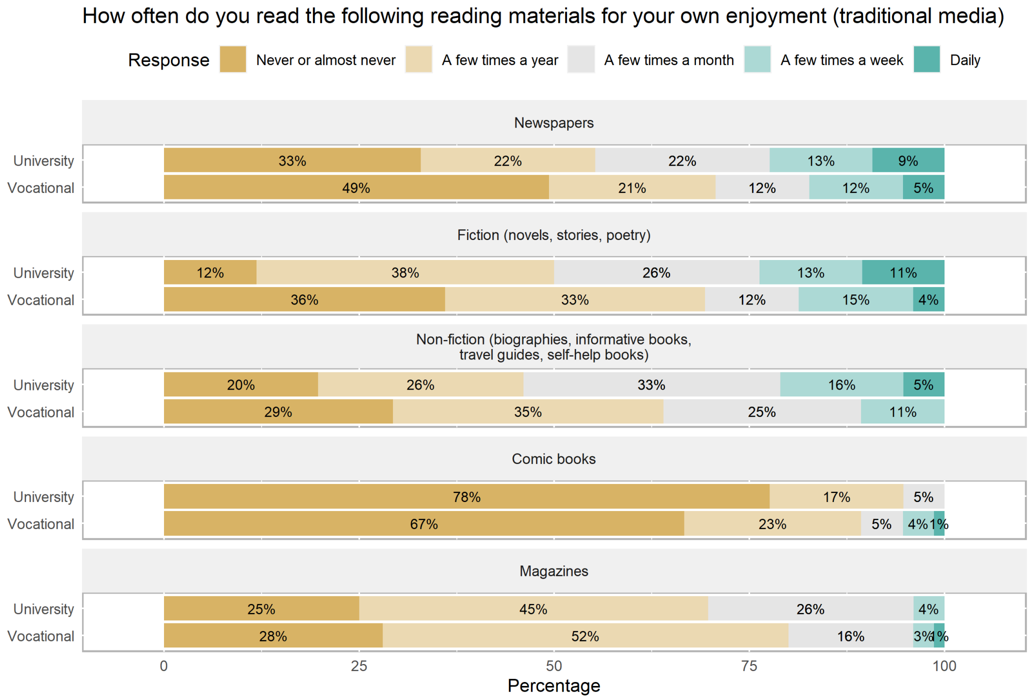
1. I only read when I have to
2. Reading is one of my favourite hobbies
3. I like talking about books with others
4. I find reading a waste of time
5. I only read to look up information I need

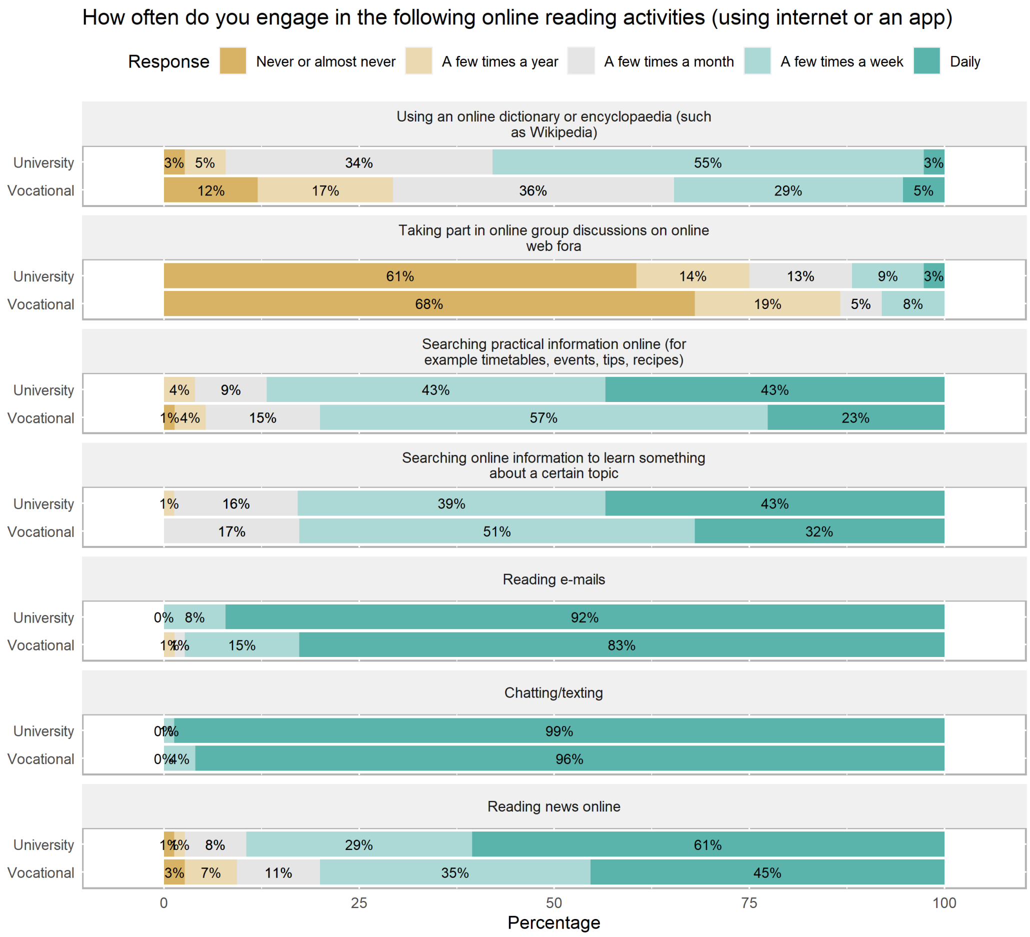
Indicate the extent to which the following statements pertain to you.

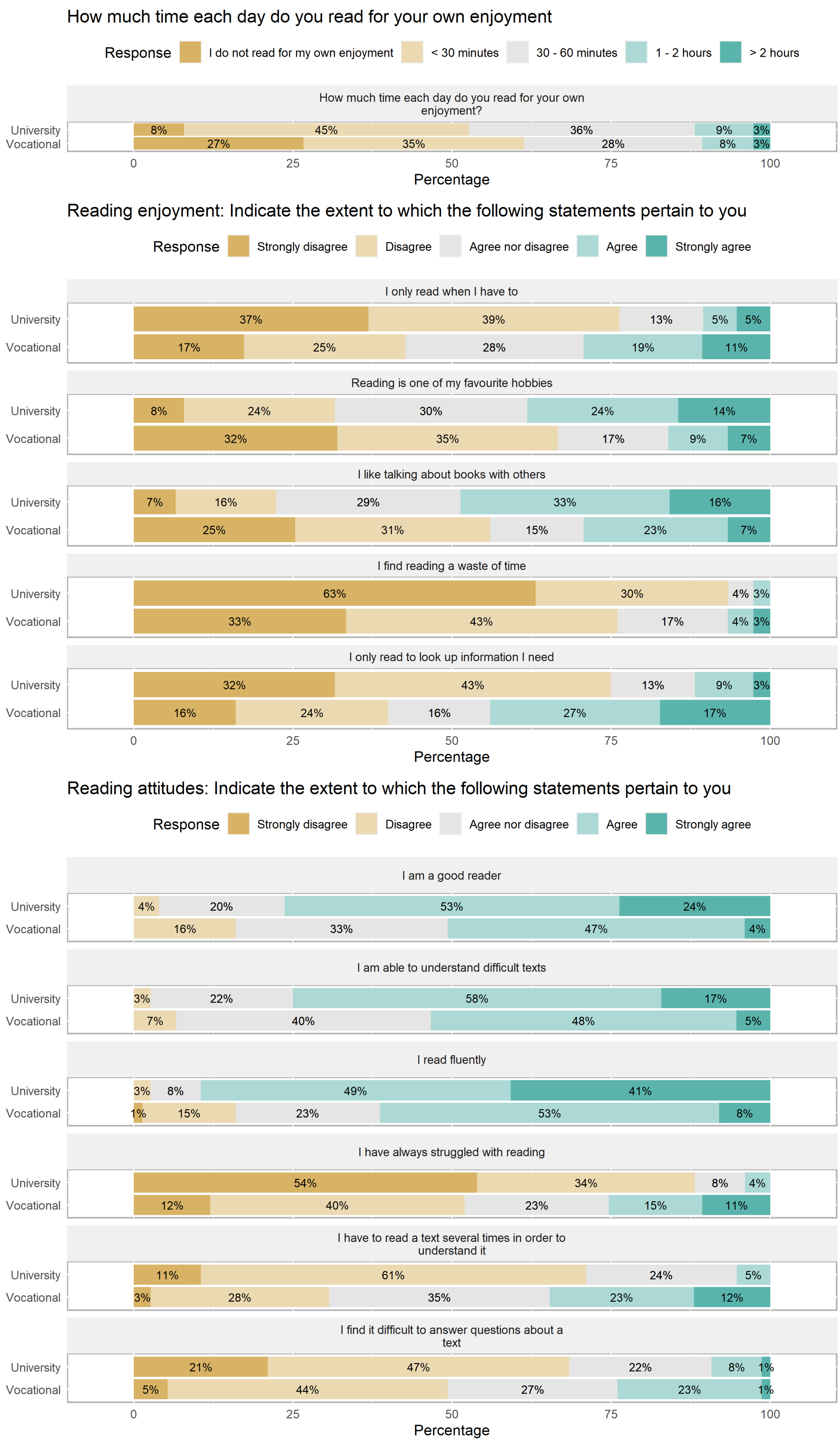
5-point scale: (Strongly disagree, Disagree, Agree nor disagree, Agree, Strongly agree)

1. I am a good reader
2. I am able to understand difficult texts
3. I read fluently
4. I have always struggled with reading
5. I have to read a text several times in order to understand it
6. I find it difficult to answer questions about a text

3.7 Appendix III - Results of reading questionnaire split by education level






How much time each day do you read for your own enjoyment
Response
I do not read for my own enjoyment
< 30 minutes
30 - 60 minutes
1 - 2 hours
> 2 hours
How much time each day do you read for your own enjoyment?
University 8% 45% 36% 9% 3%
Vocational 27% 35% 28% 8% 3%
0 25 50 75 100
Percentage
Reading enjoyment: Indicate the extent to which the following statements pertain to you
Response
Strongly disagree
Disagree
Agree nor disagree
Agree
Strongly agree
I only read when I have to
University 37% 39% 13% 5% 5%
Vocational 17% 25% 28% 19% 11%
Reading is one of my favourite hobbies
University 8% 24% 30% 24% 14%
Vocational 32% 35% 17% 9% 7%
I like talking about books with others
University 7% 16% 29% 33% 16%
Vocational 25% 31% 15% 23% 7%
I find reading a waste of time
University 63% 30% 4% 3%
Vocational 33% 43% 17% 4% 3%
I only read to look up information I need
University 32% 43% 13% 9% 3%
Vocational 16% 24% 16% 27% 17%
0 25 50 75 100
Percentage
Reading attitudes: Indicate the extent to which the following statements pertain to you
Response
Strongly disagree
Disagree
Agree nor disagree
Agree
Strongly agree
I am a good reader
University 4% 20% 53% 24%
Vocational 16% 33% 47% 4%
I am able to understand difficult texts
University 3% 22% 58% 17%
Vocational 7% 40% 48% 5%
I read fluently
University 3% 8% 49% 41%
Vocational 1% 15% 23% 53% 8%
I have always struggled with reading
University 54% 34% 8% 4%
Vocational 12% 40% 23% 15% 11%
I have to read a text several times in order to understand it
University 11% 61% 24% 5%
Vocational 3% 28% 35% 23% 12%
I find it difficult to answer questions about a text
University 21% 47% 22% 8% 1%
Vocational 5% 44% 27% 23% 1%
0 25 50 75 100
Percentage

4 The influence of literacy on spoken language processing: an individual difference and latent-variable approach

Abstract

Individuals vary substantially in their language abilities and these differences are thought to be due to differences in individuals' intrinsic cognitive capacity and variation in individual's exposure to linguistic input. In particular, one's literacy level, which arguably reflects one's exposure to written input, may influence language abilities, as written language is more varied, structured and complex compared to spoken language. In the present study, we re-analysed a large, publicly available dataset of cognitive and linguistic tasks run in participants with diverse educational backgrounds. Our first aim was to map out the relationships between tasks measuring various cognitive and linguistic constructs. Analyses using a latent-variable approach showed that tasks measuring individual cognitive and linguistic constructs correlated as expected. Our second aim was to investigate whether literacy explains variance in spoken language processing skills (word production and word comprehension), while accounting for individual differences in those general cognitive abilities that are likely involved in spoken language processing (nonverbal intelligence, processing speed). We found that literacy contributed to word production and word comprehension after accounting for the influence of nonverbal intelligence and processing speed. In sum, we showed that individuals' literacy level has a substantial influence on their spoken language processing ability.

4.1 Introduction

Individuals differ substantially in how they produce and comprehend spoken language. However, despite the prevalence of individual differences in daily language use, psycholinguistics has focused for a large part on characterizing participants' average or modal behaviour when investigating linguistic processing. Recently, this tradition has been complemented with the general interest in individual differences in various cognitive and linguistic processes. Given the indubitable presence of variation in people's language abilities, one of the main topics is how these individual differences in various language-related abilities come about. According to emergentist approaches to language (Dąbrowska, 2018; Lieven, 2016), differences among individuals' language ability are the result of variation in individuals' intrinsic capacity for learning and differences in individuals' exposure to linguistic input (see Kidd et al., 2020, 2018, for overviews). Linguistic input can vary in its modality: for example, it can be auditory when listening to spoken language and visual when reading written language. Written input requires a specific set of skills, specialized in processing written language, and knowledge pertaining to written language. In the present study, we examined if skills and knowledge pertaining to written input, which we summarize as 'literacy', contribute to individual differences in people's spoken language abilities, specifically their word production and comprehension skills, when accounting for the variation in general cognitive skills. Results will improve our understanding of how individual differences in language abilities come about.

Until recently, differences in individuals' linguistic abilities were not a main topic of interest in linguistic research, as traditional formal linguistic approaches did not consider it particularly important to investigate in order to ultimately understand language acquisition and language attainment. They argued that individual differences do not fundamentally influence the acquisition of the core representational properties of the linguistic system, because variability in these core representational properties is constrained by innate knowledge structures (Universal Grammar) (see Kidd et al., 2020, 2018, for overviews). Recently, emergentist approaches to language, such as the usage-based approach (Dąbrowska, 2018; Lieven, 2016), place a larger emphasis on the investigation of individual differences in linguistic abilities. These accounts claim that the properties of the language system are acquired through analyses of and generalization from input, with little influence of prior language-specific knowledge. In this framework, variation in individuals' intrinsic capacity for learning and differences in the quantity and quality of linguistic input are thought to have a considerable influence on language acquisition and language attainment (see Kidd et al., 2020, 2018, for overviews). From this perspective, investigating these individual differences and exploring how they come aboutis an endeavour worthwhile.

One specific hypothesis derived from emergentist approaches to language is that language exposure enhances the efficiency of language processing in language-related abilities. Indeed, experimental evidence indicates that people who have had more exposure to language show increased performance on various measures of language production, including speeded pronunciation (Yap et al., 2012), picture naming (Rodriguez-Aranda & Jakobsen, 2011) and verbal fluency (Rodriguez-Aranda & Jakobsen, 2011; Shao et al., 2014; N. Unsworth et al., 2011). This pattern is also visible for spoken word comprehension and spoken word recognition measured with a lexical decision task (Brysbaert, Lagrou, & Stevens, 2016; Diependaele et al., 2013; Janse & Jesse, 2014; Mainz et al., 2017), listening comprehension (Andringa et al., 2012), speech recognition (Bent et al., 2016) and use of predictive information in spoken context (Federmeier et al., 2002; Mani & Huettig, 2014).

The lexical entrenchment hypothesis (Brysbaert, Lagrou, & Stevens, 2016; Diependaele et al., 2013) provides an explanation as to why language exposure enhances language processing. According to this hypothesis, language exposure sharpens lexical representation (cf. lexical quality hypothesis, Perfetti, 2007). Having stable and more precise and complete lexical representations results in faster activation and less interference from similar representations and thereby more efficient language processing. Language exposure especially seems to improve processing of low-frequency words. People with more language exposure have negligible differences in processing high- and low-frequency words as a result of their precise and stable lexical representations. On the other hand, people with less language exposure have more difficulty processing low-frequency words compared to frequent words (Diependaele et al., 2013; Mainz et al., 2017; Yap & Balota, 2009).

Interestingly, the modality of language input influences the frequency of the words and (syntactic) structures that are encountered during an episode of language exposure. That is, written text contains more low-frequency linguistic elements and is generally more structured than speech. Concerning syntax, written text has more complex syntactic structures, such as passives, object relative clauses and participial phrases (Roland et al., 2007). With regard to the vocabulary used in written and spoken language, written text contains more infrequent words than spoken language (Cunningham & Stanovich, 1998; Hayes & Ahrens, 1988). Sentence structure tends to be more complex in written language, as for example subordination is 60% more common in written compared to spoken language (Kroll, 1977). Thus, written language exposure may sharpen lexical representations and thereby increase language processing efficiency more than spoken language exposure, because written language is more varied and complex in terms of syntax, structure and vocabulary.

Not a single study that previously investigated the relationship between language exposure and language processing has explicitly considered the modality of language exposure when operationalizing this construct. Arguably, extensive exposure to written language will have a larger influence on an individual's literacy level than exposure to spoken language. Literacy can be defined as a specific set of written language processing skills, such as reading ability, and knowledge pertaining to written language, such as spelling and vocabulary. In line with the entrenchment hypothesis (Brysbaert, Lagrou, & Stevens, 2016; Diependaele et al., 2013), literacy level, being a reflection of an individual's exposure to written language, may be a strong contributor to individual differences in people's spoken language processing abilities. Highly literate people, who have had extensive exposure to written language, have the processing skills necessary for quick and efficient intake of written input and possess considerable knowledge of different aspects of written language, such as the spelling and meaning of many different words. This quick and reliable access to and ever-expanding knowledge of written language results in precise and complete lexical representations. This may particularly be the case for infrequent words, which are more common in written than spoken input. Thus, literacy level may be a strong contributor to efficient language processing.

In the present study, we aimed to test this hypothesis and investigated to what extent literacy, referring to skills and knowledge pertaining to written language input, contributes to language processing. When studying broad constructs such as language processing or literacy, one must be aware that no single task can measure all aspects of a multi-facetted construct. Due to financial and time constraints, previous studies investigating these constructs often adopted a single task approach, where only one task is used to index a skill, ability or construct. This approach comes with validity problems, which makes interpretation of the results difficult. Problems arising from single task approaches are ameliorated by using a latent-variable approach. By administering multiple tasks per construct and statistically extracting what is common among different tasks, one can obtain a purer measurement of the construct (Miyake et al., 2000). Using this latent-variable approach, we approximated literacy by participants' performance on two reading tasks to capture written language processing skills, a spelling task, a measure of receptive vocabulary and the author recognition test to assess participant's knowledge pertaining to written language input. The latent-variable approach was also applied to the construct of spoken language processing. Spoken language processing was divided into two distinct constructs to facilitate the comparison of our results with previous literature, which investigated word production and word comprehension separately. To measure the construct word production, a picture naming task was used to measure participant's speed of lexical access during production, an antonym production task measure production following semantic activation, a verbal fluency task to measure production following both semantic and phonological activation and a speech rate task to measure participants' maximal speech rate. To capture the construct of word comprehension, an auditory lexical decision task was used to measure lexical access speed, a rhyme judgement task to measure phonological mapping abilities and a semantic categorization task to measure semantic access during spoken word recognition.

Emergentist approaches to language argue that individual differences in language processing ability are also caused by variability in individuals' intrinsic capacity for learning. Both genetic and environmental factors influence the brain's capacity to perceive, process and learn from language input. In individual-difference research, it is commonly observed that individuals' performance on one task of cognitive ability correlates strongly with performance on other cognitive ability tasks. This observation has led to the idea that a certain general cognitive factor underlies cognitive abilities. Ample experimental evidence points to the existence of such a general or nonverbal intelligence factor (see Deary, 2001, for a review). Even though this nonverbal intelligence factor is not linguistic in nature, it may influence performance on linguistic tasks, as language usage draws on general, non-linguistic cognitive processes as well (McQueen & Meyer, 2019). Thus, the examination of individual differences in linguistic abilities necessitates the inclusion of a measure of nonverbal intelligence. Moreover, since language processing depends on access to linguistic information in the mental lexicon, language processing efficiency relies on quick access. Faster processing allows individuals to allocate more time and cognitive resources to the next piece of information (Fernald, Perfors, & Marchman, 2006), thereby optimizing efficiency. Thus, processing speed is another skill to account for when exploring individual differences in language processing abilities.

The present study had two goals. First, we aimed to map out the relationships between various tasks that measure components of word production, word comprehension, literacy and processing speed. Our second aim was to investigate the influence of literacy, which refers to skills and knowledge pertaining to written language specifically, on spoken language processing, while accounting for the influence of general cognitive skills (nonverbal intelligence and processing speed). We focused on two forms of word-level spoken language processing, namely word production and word comprehension, to facilitate the comparison with previous studies. We re-analysed parts of a large, open access dataset that assessed linguistic and cognitive abilities in a sample of 112 adults with different educational backgrounds using 33 different tasks (Hintz, Dijkhuis, van 't Hoff, McQueen, & Meyer, 2020). Correlational analyses were used to assess the relationships between tasks that were designed to measure the constructs of word comprehension, word production, literacy and processing speed. Factor analyses were used to create single factor scores from the various measures used to operationalize the constructs. Regarding the first aim, we hypothesized that measures within a construct would correlate moderately. As for the second aim, we hypothesized that, in line with the entrenchment hypothesis, literacy would contribute to both aspects of word-level spoken language processing while accounting for the effect of nonverbal intelligence and processing speed. Two regression analyses were performed to investigate this hypothesis.

4.2 Methods

Participants

The dataset, made available by Hintz, Dijkhuis, et al. (2020), contained data from 112 participants (age: 21.8 years old, 73 female). The participants had no hearing or vision problems, and no history of speech, reading or other language-related pathology. The participants varied in their level of education, as one participant was a high-school graduate, 24 participants attended vocational education and 87 attended (applied) university. Participants were paid €92,- for their participation.

Tasks

Task descriptions are adapted from Hintz, Dijkhuis, et al. (2020). For speeded tasks, a pre-processing pipeline involving data trimming and outlier replacement was applied. Also, the signs of each individual's final scores on the speeded tasks were reversed such that higher scores reflected better performance. Participants performed all tasks of the test battery twice, which allowed measures of test-retest reliability to be calculated for each task. Pearson's correlations between the scores from the first and the second session were calculated; the coefficients are reported for each of the tests selected for the present analysis. For the analyses performed in this study, the data of the first session was used.

Word production

Picture naming. In the picture-naming task, participants were shown photographs and were asked to name these as quickly as possible. The target pictures were 40 photographs (300 x 300 pixels) of common objects taken from de Groot, Koelewijn, Huettig, and Olivers (2015) or retrieved via an online search engine. The object names varied in lexical frequency (average ZipfF[1] = 3.83, SD = 0.88, range = 2.04 – 5.39 in the SUBTLEX-NL corpus (Keuleers, Brysbaert, & New, 2010)), but were, according to the prevalence norms (the degree to which a given word is known by a representative sample of the Dutch speaking population (Keuleers et al., 2015), likely to be known by all participants (average prevalence 99.60%, SD = 0.40, range 97.70 – 100). The average number of phonological neighbours (sum of additions, substitutions, deletions of segments) of the object names was 4.05 (SD = 3.99, range = 0 – 18 (Marian, Bartolotti, Chabal, & Shook, 2012). Four practice trials preceded the task. During a trial, participants first saw a fixation cross for 800 ms, and then one of the pictures for three seconds during which participants were supposed to name the picture. After an inter-trial interval of one second, the next trial started. Participant's utterances were recorded. Naming accuracy and onset latencies were coded offline using Praat software (Boersma & Weenink, 2001). The final log-transformed score was participants' average onset latency for correctly named trials. Test-retest reliability of the task was high (r = .69).

Antonym production. During the antonym production task, participants were provided with words and were asked to produce the opposite meaning (e.g., cue: "hot", antonym: "cold"). This test was developed by Mainz et al. (2017). The 25 target words varied in frequency (M = 3.84, SD = 1.41, range = 1.70

[1] As recommended by van Heuven, Mandera, Keuleers, and Brysbaert (2014), Zipf-transformed word frequency values (ZipfF) were used, which can be operationalized as log10(frequency per million words) + 3

– 5.26) (Keuleers, Diependaele, & Brysbaert, 2010) and prevalence (M = 1.00, SD = 0.04, range = 0.85 – 1.00) (Keuleers et al., 2015). Three practice trials preceded the task. During a trial a fixation cross was first shown for 500 ms, after which a target word was presented both orthographically and auditory. Participants had to produce the antonym of the target word aloud and press a button on the keyboard to proceed to the next trial. Trial order was the same for each participant. Participants' utterances were and their naming accuracy was coded offline using Praat software (Boersma & Weenink, 2001). The final score was the proportion of correctly answered trials. Test-retest reliability of the task was high (r = .70).

Verbal fluency. In the verbal fluency task, previously run by Shao et al. (2014), participants had to name as many unique words as possible within one minute in a certain semantic category for the first part of the task ("animals" or "food/drinks") and starting with a certain letter ("M" or "S") for the second part of the task. During the task, participants first saw a countdown from three to zero, after which they were presented with the category. During participants' production, a timer that counting down from 60 to zero was shown at the bottom of the screen. After one minute, the second part of the trial started in which participants had to produce unique words starting with a certain letter. During the one-minute recordings, utterances were recorded and utterances were coded offline using Praat software (Boersma & Weenink, 2001). The final scores were the number of unique items produced in the correct category and the number of items starting with the correct letter. Test-retest reliability of the task was high (r = .72 for the semantic category part, r = .70 for the phonological part).

Maximal speech rate. In the maximal speech rate test, participants were asked to recite the months of the year as quickly as possible with clear pronunciation. Participants' speech was recorded and speech was coded offline using Praat software (Boersma & Weenink, 2001). The task was performed twice. The final score was the log-transformed average speech duration of the two attempts. In case only one attempt was correct (e.g., one or more months were skipped), the speech duration of the correct attempt was used as the final score. Test-retest reliability of the task was high (r = .88).

Word comprehension

Rhyme judgement. In this task, participants were presented with a set of two monosyllabic pseudowords and were asked to judge as fast as possible whether the words in the set rhymed. Rhyme overlap was defined as overlap in the vowel and following consonant(s). The 40 items consisted of 80 pseudowords that were constructed by changing one or more letters in existing words while still maintaining a phonological structure that adheres to Dutch phonotactics. Of the 40 items, 24 rhymed (e.g, “noost-woost”), eight sets were foils (pseudowords sharing the vowel but not the following consonants, e.g., “bruip-fluik”) and eight items were completely non-rhyming sets (e.g., “beus-fuug”). There were four practice trials. A trial consisted of a fixation cross presented for 500 ms and auditory presentation of the two pseudowords in an item with an interval of 500 ms in between nonwords. Participants had to indicate as fast as possible whether the two pseudowords rhymed or not, by pressing the associated button on the button box (right for a rhyming set, left for a non-rhyming set). The next trial started two seconds after a button was pressed. Rhyming trials, foil trials and non-rhyming trials were presented in a pseudo-random order. The final, log-transformed score was the average reaction time to the rhyming trials. Test-retest reliability of the task was high ($r = .79$).

Auditory lexical decision. In this test, participants heard spoken target words and had to indicate, using a button box, whether the word was an existing Dutch word or a non-existing word. Regarding the word trials, 60 Dutch words were selected from the SUBTLEX-NL corpus (Keuleers, Brysbaert, & New, 2010). The words varied substantially in word frequency (average ZipfF = 3.65, SD = 0.85, range = 2.04 – 5.66), but were highly prevalent and thus most likely known to all participants (average prevalence = 99.6, SD = 0.5, range = 97.3-100). The average number of phonological neighbours (Marian et al., 2012), defined as deletions, additions, and substitutions, was 2.8 (SD = 3, range = 0 – 12). The 60 nonwords were constructed using Wuggy (Keuleers & Brysbaert, 2010). There were three additional practice trials. During a trial, a fixation cross appeared for 300 ms, after which a target word was played. Participants decided on the lexicality of the word using a button box (right for an existing word, left for a pseudoword) on the lexicality of the target word as fast as possible. The next trial started 1000 ms after pressing a button to decide. The order of the existing and pseudoword trials was pseudo-randomized. The final, log-transformed, score was the average response latency for trials correctly identified as being an existing word. Test-retest reliability of the task was medium to high (r = .69).

Semantic categorization. In this task of semantic word comprehension, participants were presented with a spoken word and were asked to judge whether the word belongs to a certain semantic category ("professions" in the first part and "means of transportation" in the second part of the task). For each category, 32 words were selected from the SUBTLEX-NL corpus (Keuleers, Brysbaert, & New, 2010). Of the 32 words, 20 matched the category and 12 were unrelated to the category. Targets and distractors in each part were matched on word frequency (Keuleers, Diependaele, & Brysbaert, 2010) (professions: ZipfF M = 3.63, SD = 0.23, range = 3.25 – 4.07; means of transportation M = 3.27, SD = 0.51, range = 2.15 – 4.12). All words used in the test were highly prevalent (known to 99% – 100% of all people) (Keuleers et al., 2015). Each part was preceded by four practice trials. During a trial, a fixation cross was presented for 500 ms after which the spoken word was presented. Participants then indicated as fast as possible whether the word they had heard belonged to the assigned category or not, by pressing the associated key on the button box (right for "Yes, this word belongs to the category" and left for "No, this word does not belong to the category). The next trial started 2000 ms after pressing a button to decide. The final, log-transformed, score was participants' average reaction time for words correctly categorized as belonging to one of the two provided categories. Test-retest reliability of the task was medium (r = .62).

Literacy

Vocabulary. Participants' receptive vocabulary size was measured using a digitized version of the Dutch Peabody Picture Vocabulary Test (Dunn & Dunn, 1997; Schlichting, 2005, for the Dutch translation). In this task, participants heard a word and had to select the corresponding referent picture out of four pictures. Systematic evaluation of auditory and written vocabulary tests has shown high correlations between written and auditory vocabulary tests, indicating that the modality in which the test is administered has little influence on performance (Mainz et al., 2017). The task became progressively more difficult as the target words became progressively more infrequent. There were no practice trials. In a trial, four pictures appeared and participants pressed the space bar of the keyboard to play the word. They could replay the word as often as they felt necessary. The test consisted of 17 blocks and each block consisted of 12 items of roughly the same difficulty. The test started at block 13, which is the normed entry block for participants aged between 18 and 35. Based on their performance (four or fewer errors in block 13), participants' next block was either more difficult (block 14) or easier (block 12) than the entry block. The test was terminated when more than eight errors were made within a block that was not the starting block or when participants reached the last item of the test. The final score was the number of correct items minus the number of errors. Test-retest ability of the task was high (r = .91)

Spelling. Participants' spelling skills were measured using a test where participants were presented words and were asked to decide whether the words were spelled correctly. The 60 words were Dutch words that adults often find difficult to spell, for example the correct use of diaereses (i.e."bacteriën" [bacteria]), the use of double consonants in plural forms ("slimmeriken" [smart people]), and use of ei/ij (diphthong /εi/, i.e., "allerlei" [all kinds]). Half of the words were spelled correctly, and half was spelled incorrectly. The words were presented in three columns of 20 words. Participants were instructed to use their mouse to tick the boxes in front of the incorrectly spelled items. The final score was the proportion of correctly categorized misspelled words minus the proportion of incorrectly selected words that were spelled correctly. Test-retest reliability of the task was high (r = .85).

Literary experience. A Dutch version of the Author Recognition Test (Moore & Gordon, 2015) developed by Brysbaert et al. (2020), was used to measure literary experience. In this task, participants were provided with a list of names, including the names of several well-known (Dutch) authors and foil names, and had to indicate which of the listed names are known authors. Of the 132 names, 90 were existing authors (e.g., Roald Dahl, Nicci French) and 42 were foil names. The names were presented in pseudo-randomized order in three columns of 44 words. The final score was the proportion author names correctly recognized minus the proportion non-authors wrongly selected. Test-retest ability of the task was high (r = .95).

Word reading. A digitized version of the Eén-Minuut-Test [One-Minute-Test] (Brus & Voeten, 1973) was used to measure word reading. In this task, participants were instructed to read aloud a list of Dutch words as fast and correct as possible. The 116 Dutch words became progressively more difficult to read in terms of phoneme complexity and syllable length (range: 1 – 5 syllables). The words were presented in four columns of 29 words each. Participants had to read the words aloud as fast and precise as possible from the top left to bottom right within one minute. Participants' speech was recorded for one minute. Speech was scored offline. The order of the words was the same for all participants. The final score was the total number of words read within one minute minus the number of incorrectly pronounced words. Test-retest reliability of the task was high (r = .79).

Nonword reading. A digitized version of the Klepel (van den Bos, Spelberg, Scheepsma, & De Vries, 1994) was used to measure participant's nonword reading. In this task, participants were asked to read aloud a list of Dutch pseudowords as fast and correct as possible. The 116 pseudowords became progressively more difficult to read in terms of phoneme complexity and syllable length (range: 1 – 5 syllables). Words were presented in four columns of 29 words. Participants had to read the words aloud as fast and precise as possible from the top left to bottom right within two minutes. Participants' speech was recorded for two minutes. Speech was scored offline. The order of the words was the same for all participants. The final score was the total number of words read within one minute minus the number of incorrectly pronounced words. Test-retest reliability of the task was high (r = .88).

Processing speed

Auditory processing speed. In the auditory processing speed task, developed by Hintz, Jongman, et al. (2020) participants were asked to respond as fast as possible to the onset of an auditory stimulus. The task has two subtasks: the 'simple' and 'choice' task.

In the simple auditory processing speed task, participants had to press a button as soon as they heard a tone. The tone was a sine tone (550 Hz). There were 20 trials and eight practice trials. In a trial, participants first saw a fixation cross. After an interval of between one a three seconds, the tone was played for 400 ms. Participants had to press the right-hand button of a button box as soon as they heard the tone. One second after their button-press, the next trial started. The final, log-transformed score was calculated as the mean reaction time across all 20 trials. Test-retest ability of the task was medium (r = .59).

In the choice auditory processing speed task, participant heard a high or low tone and had to press buttons as quickly as possible to indicate whether the tone was high or low. The tones were sine tones (300 and 800 Hz). There were 40 trials and 16 practice trials. During a trial, first a fixation cross was presented. Then, after an interval varying between one and three seconds, one of the tones was played. Participants had to press the right button when they heard a high tone and the left button when they heard a low tone as quickly as possible. The next trial started one second after a button was pressed. The tones appeared equally often throughout the task and were presented in a pseudo-randomized order. The final, log-transformed, score was the mean reaction time of correct responses. Test-retest ability of the task was medium to high (r = .76).

Visual processing speed. During the visual processing speed, developed by Hintz, Jongman, et al. (2020), adapted from Deary, Liewald, and Nissan (2011), participants had to respond as quickly as possible to the onset of a visual stimulus. Similar to the auditory processing speed task, the task had two subtasks: the simple and choice task.

In the simple visual processing speed task, participants had to press a button as soon as they saw a line drawing of a triangle. The black-contoured triangle was 200 x 200 pixels. There were 20 trials and eight practice trials. During a trial, participants were first shown a fixation cross. After a varying interval between one and three seconds, the triangle was presented. Participants needed to press the right-hand button of a button box as soon as they saw the triangle. The next trial started one second after the button is pressed. The final, log-transformed score was the mean reaction time across all 20 trials. Test-retest reliability of the task was medium (r = .58).

For the choice auditory processing task, participants saw a line drawing of a star or circle and had to indicate which figure they saw using a button box. The black-contoured line drawings were 200 x 200 pixels. There were 40 trials and 16 practice trials. During a trial, first a fixation cross appeared. After an interval varying between one and three seconds, either a circle or star was presented. Participants had to press the right button when a circle was shown and the left button when a star appeared. The figures appeared equally often throughout the task and were shown in a pseudo-randomized order. The final, log-transformed score was the mean reaction time of correct responses. Test-retest ability of the task was high (r = .78).

Letter comparison. A digitized version of the letter comparison task (Huettig & Janse, 2016) was used as an additional processing speed test. Participants were asked to decide whether two presented letter strings were the identical or not. Half of the 48 pairs of letter strings consisted of three consonants (e.g., TZF) and the other half of six consonants (e.g., RNHKTG). The pairs were presented in a large monospaced font (font size 70) with a distance of 300 pixels between the letter strings in a pair. The first block consisted of the 24 three-letter pairs and the second block of the 24 six-letter pairs. Half of the trials contained of two identical letter strings and the other half contained non-identical strings, where one letter of the three or six was different. There were six three-letter practice trials. Each trial started with a fixation cross for 600 ms, after which the two strings of a pair were presented. Participants had to decide as quickly as possible whether the pairs were identical or not, by pressing the left button on a button for a non-identical pair, and the right button for an identical pair. 1000 ms after pressing a button, the next trial started. The final, log-transformed score was calculated as the mean reaction time of the correct responses. Test-retest reliability of the task was high (r = .83).

Nonverbal intelligence

Nonverbal intelligence was assessed using the digitized version of Ravens' advanced progressive matrices (Raven et al., 1998). Participants were shown a matrix of geometric patterns and were asked to complete the matrix by choosing the correct option from eight different options. The task, consisting of 36 matrices, became progressively more difficult. There were six practice trials. During the task, a timer counting down from twenty minutes to zero was shown in the corner of the screen. When participants clicked on one of the options that they thought completed the matrix correctly, the next trial started. Trials could be skipped by clicking a "skip"-button. These trials were then presented again at the end of the test. If the participant did not know the correct option again, they could click an "I don't know"-button, which resulted in the trial not being presented again. The final score was the proportion of correctly answered trials. Test-retest reliability of the task was high (r = .87).

Procedure

The full test session took 6.5 hours to complete and was divided into four sessions of 60 – 90 minutes. Between sessions, there were two breaks that lasted 15 – 20 minutes and one lunch break of 45 minutes. Participants were tested in groups of maximally eight individuals in a quiet room of 30 m^2. Participants were seated in a semicircle across the room facing the wall, with approximately 1m -– 1.5m space between them. Noise cancelling divider walls (height 1.80 m, width 1 m) were placed between participant desks. The walls in front of the participants were covered with curtains to absorb as much noise as possible. Each participant performed the task on an experimental laptop (Hewlett Packard ProBooks 640G1, 14-inch screens, running Windows 7), using a computer mouse, two-buttoned button box and Beyerdynamic DT790 headsets to play auditory stimuli and record speech. For the speech production task, participants additionally wore earplugs underneath their headsets. The experiments were either implemented in Presentation version 20.0 (Neurobehavioral Systems, 2017) or as web application in 'Frinex' (framework for interactive experiments) in a web browser (Chrome, version 75.0.3770.142). For each task involving the presentation of spoken stimuli, recordings were made by a native speaker of Dutch who was not the experiment leader to avoid voice familiarity effects. Recordings were made using a Sennheiser microphone sampling at a frequency of 44 kHz (16-bit resolution). The order of the tasks in the task battery was the same for each participant to minimize potential influences of the test procedure on participants' performance.

4.3 Results

Descriptive analyses

Data were checked for outliers and missing data. Table 4.1 displays the number of data points for each measure, the mean, standard deviation and the range of the scores.

Factor analyses

Factor analyses were conducted to determine the factor structure of the four constructs of word production, word comprehension, literacy and processing speed and their underlying measures. First, correlational analyses were performed. Then, principal component analyses were performed to determine the factor structure and calculate a factor score for each construct.

Correlations. Frequentist and Bayesian correlations were calculated using JASP version 0.14 (JASP Team, 2020). For Bayesian correlations, effects of predictors

Table 4.1: Summary descriptive statistics on the measures used in this study.

Construct	Measure	N	Missing	M	SD	Range
Word production	Picture naming	111	1	-2.95	0.06	-3.08, -2.74
	Antonym production	111	1	0.72	0.10	0.48, 0.92
	Verbal fluency - category	106	6	24.37	4.94	14, 39
	Verbal fluency - phonology	112	0	15.55	4.43	3, 30
	Speed of articulation	106	6	-3.60	0.09	-3.82, -3.39
Word comprehension	Rhyme judgement	109	3	-2.89	0.08	-3.12, -2.72
	Auditor lexical decision	112	0	-2.93	0.05	-3.10, -2.84
	Semantic categorization	109	3	-2.92	0.06	-3.12, -2.80
Literacy	Vocabulary	112	0	56.11	24.56	0, 95
	Spelling	112	0	0.56	0.18	0.10, 0.93
	Literary experience	112	0	0.20	0.12	-0.03, 0.60
	Word reading	111	1	62.8	12.08	34, 107
	Nonword reading	111	1	89.8	14.34	56, 116
Processing speed	Auditory processing speed simple	112	0	-2.35	0.08	-2.65, -2.20
	Auditory processing speed choice	112	0	-2.60	0.09	-2.86, -2.41
	Visual processing speed simple	112	0	-2.37	0.05	-2.55, -2.24
	Visual processing speed choice	112	0	-2.62	0.07	-2.86, -2.5
	Letter comparison	107	5	-3.02	0.08	-3.28, -2.86
Nonverbal intelligence		112	0	0.55	0.17	0.14, 0.89

were considered meaningful when the 95% Credible Intervals (CI) did not contain zero, which indicates that the correlation was non-zero effect with high certainty.

Frequentist Pearson's correlations between all measures are displayed in Appendix I. Regarding correlations between the measures indexing word production (Table 4.2) — picture naming, antonym production and verbal fluency (both category and phonology version) — all showed small to medium sized correlations (range: $r = .15 – .42$). Concerning word comprehension, for the measures rhyming, lexical decision and semantic categorization, correlations were high (range: $r = .69 – .75$) (Table 4.3). In addition, the measures of literacy (i.e., vocabulary, spelling, literary experience, word and nonword reading) all correlated to a medium to large degree (range: $r = .28 – .58$) (Table 4.4). For processing speed, visual processing speed tasks correlated highly with the auditory processing speed tasks (range: $r = .43 – .62$) and moderately with the letter comparison task (range: $r = .22 – .41$). Correlations between the auditory processing speed tasks and the letter comparison test were, however, lower (range: $r = .09 – .28$) (Table 4.5).

Table 4.2: Correlations between the measures of word production, with 95% credible interval in brackets.

	Picture naming	Antonym production	Verbal fluency – category	Verbal fluency – phonology
Picture naming	1			
Antonym production	.35** [0.17-0.50]	1		
Verbal fluency -– category	.40** [0.27-0.55]	.27* [0.09-0.44]	1	
Verbal fluency – phonology	.21* [0.02-0.38]	.30* [0.19-0.46]	.42** [0.26-0.56]	1
Maximal speech rate	0.15 [-0.05-0.32]	0.12 [-0.08-0.30]	.25* [0.05-0.41]	.31* [0.12-0.47]

Note. * p < .05, ** p < .001

Table 4.3: Correlations between the measures of word comprehension, with 95% credible interval in brackets.

	Rhyme judgement	Lexical decision
Rhyme judgement	1	
Lexical decision	.69** [0.58-0.78]	1
Semantic categorization	.69** [0.58-0.78]	.75** [0.64-0.82]

Note. * p < .05, ** p < .001

Table 4.4: Correlations between the measures of literacy, with 95% credible interval in brackets.

	Vocabulary	Spelling	Literary experience	Word reading
Vocabulary	1			
Spelling	.50** [0.34-0.62]	1		
Literary experience	.58** [0.44-0.69]	.55** [0.40-0.66]	1	
Word reading	.32* [0.14-0.47]	.51** [0.36-0.63]	.40** [0.22-0.54]	1
nonword reading	.28* [0.09-.44]	.39** [0.22-0.53]	.38** [0.21-0.52]	.56** [0.42-0.67]

Note. * p < .05, ** p < .001

Table 4.5: Correlations between the measures of processing speed, with 95% credible interval in brackets.

	Auditory processing speed	Auditory processing speed choice	Visual processing speed	Visual processing speed choice
Auditory processing speed simple	1			
Auditory processing speed choice	.62** [.48-.72]	1		
Visual processing speed simple	.43** [.26-.57]	.47** [.30-.60]	1	
Visual processing speed choice	.44** [.27-.57]	.61** [.47-.71]	.48** [.32-.60]	1
Letter comparison	0.09 [-.10-.27]	.28* [.09-.44]	.22* [.03-.39]	.41* [.23-.55]

Note. * $p < .05$, ** $p < .001$

Principal Component Analyses. For each of the four constructs, principal component analyses (PCA) were performed using the oblimin rotation, which allows for correlation between the factors in a construct's factor solution. It is likely that factors within a construct's factor solution correlate, given that all measures included in the analysis aim to approximate the same construct. PCA was performed in SPSS version 20.0 (IBM Corp, 2011). Assumptions for PCA were met: all variables were measured at the continuous level and were linearly related. Sampling adequacy was assessed by calculating the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy. KMO measures for the each construct PCAs ranged from .67 to .77 and KMO measures for each variable ranged from .63 to .83. These values are adequate (Kaiser, 1974), indicating that the PCA results will be reliable. Bartlett's Test of Sphericity was significant for each construct, indicating that the variables within a construct correlated adequately, which allows for data reduction through PCA.

For word production, the principal component analysis resulted in a single-factor solution that explained 41.50% of the variance. Factor loadings of the five measures on this single factor were .66 for picture naming, .58 for antonym production, .76 and .68 for verbal fluency category and phonology respectively, and .53 for maximal speech rate.

With regard to word comprehension again a single-factor solution was found, which explained 80.74% of the variance. Factor loadings on this single factor were .88 for rhyme judgement, .91 for auditory lexical decision and .91 for semantic categorization.

For literacy, a single-factor solution explained 55.80% of the variance. Factor loadings were .72 for vocabulary, .80 for spelling, .78 for literary experience, .75 for word reading and .69 for nonword reading.

Regarding the construct processing speed, a single-factor solution was found which explained 52.57% of the variance. Factor loadings on this single factor were .71 for simple auditory processing speed, .83 for choice auditory processing speed, .74 for simple visual processing speed, .81 for choice visual processing speed and .48 for letter comparison.

Based on the principal components analyses, a factor-score was calculated for each construct. Table 4.6 displays the correlations among the factor scores and non-verbal intelligence.

Table 4.6: Correlations among factor scores and nonverbal intelligence, with 95% credible interval in brackets.

	Production factor	Comprehension factor	Literacy factor	Processing speed factor
Production factor	1			
Comprehension factor	.24* [0.04-0.41]	1		
Literacy factor	.60** [0.44-0.70]	.39** [0.22-0.54]	1	
Processing speed factor	.21* [0.01-0.39]	.67** [0.54-0.76]	.39** [0.21-0.55]	1
Nonverbal intelligence	.35** [0.17-0.51]	0.15 [-0.04-0.32]	.49** [0.33-0.61]	.21* [0.02-0.38]

Note. * $p < .05$, ** $p < .001$

Regression analyses

In order to investigate the influence of literacy on word production and word comprehension after accounting for the influence of nonverbal intelligence and processing speed, two regression analyses were performed with respectively word production and word comprehension as dependent variables. Frequentist and Bayesian linear regression analysis was performed in JASP (JASP Team, 2020). In the frequentist regression analyses, literacy, non-verbal intelligence and processing speed were entered simultaneously (enter method).

Bayesian analyses are concerned with the likely magnitude of effects rather than statistical significance. For model comparison, Bayes Factors (BF10) were used as indicators of a model's likelihood (H1) compared to the likelihood of a null-model without predictors (H0). Following the classification scheme by Jeffreys (1961), a BF10 of 1 – 3 would indicate anecdotal evidence for H1 over H0, a BF10 value of 3 – 10 substantial evidence for H1, and a BF10 of > 10 strong evidence. This classification scheme also applied to a predictor's Bayes Factor of inclusion ($BF_{inclusion}$) when deciding whether the effect of the predictor was meaningful. Furthermore, effects of predictors were considered meaningful when the 95% Credible Intervals (CI) did not contain zero, which indicates that the predictor has a non-zero effect with high certainty. Regarding the priors, a Jeffreys–Zellner–Siow (JZS) prior was used, which is appropriate for analyses with a small number of predictors (Liang, Paulo, Molina, Clyde, & Berger, 2008). For model priors, beta-binomial model priors with $\alpha = \beta = 1$ were used, which are appropriate when there is little prior knowledge (Lynch, 2007).

Assumptions for linear regression were checked. Durbin-Watson values ranged between one and three, indicating that residual terms were uncorrelated (word production regression analysis: 2.43; word comprehension regression analysis: 2.07). Multicollinearity was not a concern as VIF scores did not exceed five (range word production regression analysis: 1.14 – 1.36, range word comprehension regression analysis: 0.74 – 0.88) and Tolerance values were below one (range word production regression analysis: 1.14 – 1.18, range word comprehension regression analysis: 0.71 – 0.85). Histograms and Q-Q plots of standardised residuals indicated that residuals were normally distributed. Scatterplots of standardised residuals indicated that the data was linear and that homoscedasticity was not an issue. Output from the frequentist linear regression analyses are displayed in Table 4.7. The output of the Bayesian linear regression analyses are displayed in Tables 4.8 and 4.9.

In the frequentist linear regression of word production, the predictors literacy, nonverbal intelligence and processing speed explained 38% of the variance ($F(3, 91) = 18.88$, $p < .001$, Cohen's $f^2 = .62$). Literacy contributed to a large degree to word production over and above the influence of nonverbal intelligence and processing speed ($\beta = 0.53$, $p < .001$, Cohen's $f^2 = .33$). The contributions of nonverbal intelligence ($\beta = 0.17$, $p = .07$, Cohen's $f^2 = .02$) and processing speed ($\beta = -0.01$, $p = .96$, Cohen's $f^2 = .00$) were small and did not reach significance. Similar findings were obtained in the Bayesian linear regression of word production: together, the three predictors explained 38% of the variance in word production. $BF_{inclusion}$ and 95% CI indicated that literacy contributed to a large degree to word production ($BF_{inclusion} = 270114.10$, 95% CI = [.37 – .76]). There was anecdotal evidence for the contribution of nonverbal intelligence to word production ($BF_{inclusion} = 1.04$, 95% CI = [0.02 – 1.84]). There was no evidence for a contribution of processing speed to word production ($BF_{inclusion} = 0.34$, 95% CI = [-.12 – .12]).

For word comprehension, the frequentist linear regression showed that 47% of the variance was explained by literacy, nonverbal intelligence and processing speed ($F(3, 102) = 30.06$, $p < .001$, Cohen's $f^2 = .88$). Literacy provided a small but significant contribution to word production after controlling for nonverbal intelligence and processing speed ($\beta = 0.18$, $p = .04$, Cohen's $f^2 = .04$). Moreover, processing speed predicted a large significant amount of variance in word comprehension, over and above the influence of literacy and nonverbal intelligence ($\beta = 0.61$, $p < .001$, Cohen's $f^2 = .31$). The contribution of nonverbal intelligence to word comprehension was small and did not reach significance ($\beta = -0.06$, $p = .50$, Cohen's $f^2 = .002$). In the Bayesian analyses, the three predictors explained 47% of the variance in word comprehension. Evidence for the contribution of processing speed was very strong ($BF_{inclusion} = 2.80e+09$, 95% CI = [0.46 – 0.79]). There was anecdotal evidence for the contribution of literacy to word comprehension ($BF_{inclusion} = 1.30$, 95% CI = [-0.00005 – 0.29]). There was no evidence for a contribution of nonverbal intelligence to word comprehension ($BF_{inclusion} = 0.36$, 95% CI = [-0.87 – 0.45]).

Table 4.7: Results of frequentist linear regression analyses for word production and word comprehension.

	Word production	Word comprehension
Predictor	β	β
Literacy factor	.53**	.18*
Nonverbal intelligence	.17	-.06
Processing speed factor	-.01	.61**
Total R^2	.38**	.47**

Note. * $p < .05$, ** $p < .001$

Table 4.8: Model comparison of Bayesian linear regression analyses for word production and word comprehension.

Word production			Word comprehension		
Model	BF10	R^2	Model	BF10	R^2
Null model	1	0	Null model	1	0
Literacy factor	6.24e+07	.36	Processing speed factor	8.55e+14	.45
Literacy factor + Nonverbal intelligence	4.72e+07	.38	Processing speed factor + Literacy factor	7.87e+11	.47
Literacy factor + Nonverbal intelligence + Processing speed factor	9.24e+06	.38	Processing speed factor + Literacy factor + Nonverbal intelligence	1.56e+11	.47

Table 4.9: Results of Bayesian linear regression analyses for word production and word comprehension.

	Word production	Word comprehension
Predictor	$BF_{inclusion}$	$BF_{inclusion}$
Literacy factor	270114.10 [0.37 – 0.76]	1.30 [-0.00005 – 0.29]
Nonverbal intelligence	1.04 [0.02 – 1.84]	0.36 [-0.87 – 0.45]
Processing speed factor	0.34 [-0.12 – 0.12]	2.80e+09 [0.46 – 0.79]

4.4 Discussion

The first aim of the present study was to map out the relationships between tasks that measure word comprehension, word production, literacy and processing speed. The second aim was to investigate the influence of literacy, operationalized as skills and knowledge acquired through exposure to written language, on two aspects of spoken language processing, namely word comprehension and word production, while accounting for individual differences in general cognitive abilities (i.e., nonverbal intelligence and processing speed). To this end, we re-analysed parts of an open-access dataset containing data from participants with diverse educational backgrounds on linguistic and cognitive skills tests. To reduce task impurity problems and in order to increase construct validity, we used a latent-variable approach rather than single task approach (Miyake et al., 2000).

With regard to the first aim, the relationships between tasks that measured the constructs word production, word comprehension, written language exposure and processing speed were investigated with correlation analyses. It was hypothesized that measures within a construct would correlate moderately. Regarding word production, picture naming, antonym production and verbal fluency correlated moderately as expected. The correlation between these tasks and the speed of articulation was not significant, probably due to the fact that the speed of articulation task specifically measures motor speed during pronunciation, whereas in the other tasks articulation speed was not measured (antonym production, verbal fluency) or only naming latency was measured (picture naming). With respect to word comprehension, the rhyming judgement, lexical decision and semantic categorization tasks correlations were high, indicating that the similar paradigms used in these tasks resulted in large amounts of shared variance. For literacy, the measures of vocabulary, spelling, word reading, non-word reading and literary experience correlations were moderate to large, as expected. The measures of visual and auditory processing speed correlated moderately to highly. Auditory processing speed measures did not correlate as highly as visual processing speed measures with the letter comparison measure, which is plausible given the shared modality across these tasks.

With regard to our second aim, we investigated the contribution of literacy on word-level processing while accounting for the influence of general cognitive abilities, with two regression analyses. It was hypothesized that literacy would explain variance in both components of spoken language processing, after accounting for nonverbal intelligence and processing speed. Results confirmed this hypothesis.

Interestingly, our results suggest that the influence of literacy and general cognitive skills differs for the word production and word comprehension tasks used in our study. Whereas literacy was the only contributor to word production, processing speed also predicted variance in word comprehension. There are studies reporting similar patterns of processing speed contributing to comprehension but not production tasks. For example, Mahr and Edwards (2018) reported processing speed to contribute to performance on a measure of receptive but not expressive vocabulary in infants. Moreover, performance on the nonword repetition task, which is assumed to capture processing capacity in a range of linguistic skills, has been found to correlate with measures of receptive but not expressive vocabulary in children (see Coady & Evans, 2008, for a review). Our findings are in line with these results, but there are currently no theoretical accounts that provide a detailed explanation for this pattern. As such, the relationship between processing speed and comprehension and production requires further investigation.

Our study is one of the first that explicitly considered the modality of language exposure when operationalizing this construct. An individual's literacy level is largely influenced by the exposure to written language. With extensive exposure to written language, individuals gain skills specialized in processing written input, such as reading ability, and also accumulate knowledge pertaining to written language, such as the spelling and meaning of many different words. Our results are the first to suggest that literacy contributes to two aspects of spoken language processing, namely word production and word comprehension. This corroborates earlier findings (Andringa et al., 2012; Bent et al., 2016; Brysbaert, Lagrou, & Stevens, 2016; Diependaele et al., 2013; Federmeier et al., 2002; Janse & Jesse, 2014; Mainz et al., 2017; Mani & Huettig, 2014; Rodriguez-Aranda & Jakobsen, 2011; Shao et al., 2014; N. Unsworth et al., 2011; Yap et al., 2012). However, these studies used a single task approach and assessed only one aspect of literacy, often vocabulary size. No single measure, however, can capture a broad construct such as literacy. The latent-variable approach used in our study minimised these task-impurity issues and achieved a purer measurement of the construct ((Miyake et al., 2000). Thus, our study provides a reliable confirmation of these earlier findings suggesting that variety in literacy level explains individual differences in spoken language processing.

One explanation for the positive effects of literacy on language processing can be derived from the entrenchment hypothesis (Brysbaert, Lagrou, & Stevens, 2016; Diependaele et al., 2013). This hypothesis argues that language exposure sharpens and stabilizes lexical representations, which will result in faster activation and less interference from similar representations and thereby more efficient language processing. Highly literate people, having received much expose to written language, have the skills and knowledge required to process written input quickly and efficiently. This would allow them to create and refine high quality lexical representations, resulting in faster activation and less interference from similar representations and thereby more efficient language processing. This explanation may particularly apply to low-frequency lexical items, since low frequency words are more common in written than in spoken language (Cunningham & Stanovich, 1998; Hayes & Ahrens, 1988). Indeed, it has been shown that highly literate people, as measured with a vocabulary task, are particularly efficient in processing low-frequency words compared to individuals with a lower literacy level (Diependaele et al., 2013; Mainz et al., 2017; Yap & Balota, 2009). However, vocabulary is only a single aspect of literacy, and thus more research assessing multiple aspects of literacy is required to fully understand the relationship between literacy and processing of low-frequency words.

Another hypothesis explaining the relationship between literacy and language processing efficiency has been proposed by Huettig and Pickering (2019). They argue that literacy enhances the prediction system that operates on both written and spoken input. In contrast to speech, information intake during reading is not tied to the (speech) pace of the interlocutor, but is self-paced instead. Thus, reading is eminently a skill for which one can improve one's information intake speed: by reading faster. One way to obtain this goal is to improve the speed and accuracy of the predictive system. There is a growing body of literature suggesting that prediction ability is an important contributor to efficient language processing (e.g., Pickering & Garrod, 2013): prediction allows for faster processing and reduced memory load, thereby increasing processing efficiency. There is currently no literature that has explored prediction ability as a mediator in the relationship between literacy and language processing efficiency. Investigating this hypothesis would be an avenue for future research.

Our study has implications for theory and practice. Theoretically, our findings support the emergentist approach (Dąbrowska, 2018; Lieven, 2016) and thereby underlines the importance of investigating individual differences in individuals' intrinsic capacity for learning and differences in the quantity and quality of linguistic input in order to ultimately understand language acquisition and final language attainment. Perhaps more importantly, our study addresses the importance of explicitly operationalizing the modality of language exposure, as it shows that individual differences in literacy, a reflection of an individual's exposure to written language, contributes to variation in language processing ability.

Our finding that literacy contributes to spoken language processing when accounting for one's general cognitive skills may be of importance in both educational and clinical practice. Both children and adults who show language deficiencies or delays may profit from interventions aimed to increase literacy. More specifically, our study underlines the importance of literacy for the development of linguistic abilities. Training literacy-related processing skills (reading ability) and knowledge (spelling, vocabulary) could therefore be a key aspect of preventive and intervening measures against language delays or problematic language development not caused by known (genetic) factors. Particularly initiatives that promote literacy and healthy reading habits in individuals who otherwise have difficulty accessing written materials are worthwhile to support.

4.5 Appendix I - Correlations between all tasks

The next table displays frequentist Pearson's correlations between all tasks.

	PN	AP	VF-C	VF-Ph	SoA	Rhy	LD	SC	Voc	Spel	LE	WR	nWR	APS	APS-C	VPS	VPS-C	LC
Picture naming	1																	
AP	.35**	1																
VF-C	.40**	.27*	1															
VF-Ph	.21*	.30*	.42**	1														
SoA	.14	.12	.25*	.31*	1													
Rhy	.35**	.23*	.15	.17	.03	1												
Lexical decision	.34**	.06	.23*	.16	-.05	.69**	1											
SC	.31*	.15	.16	.16	-.07	.69**	.75**	1										
Vocabulary	.33**	.64**	.20*	.31*	.12	.36**	.18*	.27*	1									
Spelling	.26**	.49**	.31**	.37**	.05	.22*	.14	.28*	.50**	1								
LE	.32*	.54**	.28*	.35**	-.02	.28*	.26*	.32*	.58**	.55**	1							
WR	.19	.27*	38**	.41**	.22*	.30*	.24*	.28*	.32**	.51**	.40**	1						
nWR	.25*	.22*	.34**	.49**	.38**	.31*	.22*	.33*	.28*	.39**	.38**	.56**	1					
APS-S	.24*	.19*	.05	21*	.19	.34*	.29*	.34**	.26*	.27*	.22*	.18	.21*	1				
APS-C	.27*	.19*	.14	.20*	.06	.54**	.51**	.56**	.34**	.35**	.23*	.24*	.22*	.62**	1			
VPS-S	.07	.04	.06	.15	.14	.38**	.33**	.26*	.13	.19*	.07	.15	.19*	.43**	.47**	1		
VPS-C	.21*	.21*	.01	.21*	-.04	.52**	.50**	.50**	.26**	.32*	.29*	.25*	.22*	.44**	.61**	.48**	1	
LC	.10	.12	.14	13	-.02	.43**	.47**	.48**	.28*	.36**	.16	.38**	.20*	.09	.28*	.22*	.41**	1
IQ	.13	.50**	.15	.34**	.17	.26*	.05	.10	.51**	.42**	.33**	.22*	.32*	.24*	.17	.06	.25*	.25*

Note. * $p < .05$, ** $p < .001$. PN = picture naming, AP = antonym production, VF-C = verbal fluency category, VF-Ph = verbal fluency phonology, SoA = speed of articulation, Rhy = rhyme judgement, LD = lexical decision, SC = semantic categorization, Voc = vocabulary, Spel = spelling, LE = literary experience, WR = word reading, nWR = nonword reading, APS-S = auditory processing speed simple, APS-C = auditory processing speed choice, VPS-S = visual processing speed simple, VPS-C = visual processing speed choice, LC = letter comparison, IQ = nonverbal intelligence.

5 A systematic review on the effect of literacy experience on spoken and written word recognition

Abstract

Research has shown that the acquisition of literacy induces fundamental changes in our language system. Importantly, individuals who are more experienced with literacy show facilitated word processing compared to individuals who are less experienced with written materials. This study examined literature on the topic of literacy experience and word recognition to determine how literacy experience influences the efficiency with which written and spoken words are processed. We explored how the relationship between literacy experience and word recognition differs for different types of lexical representations (semantic, phonology, orthography, syntax) and for different levels of representation (lexical, morpheme/syllable, sublexical). We examined studies of written and spoken word recognition separately to improve our understanding of how modality influences the effect of literacy experience on word recognition. Moreover, our review included both studies conducted with adults and those with children to shed light on how word recognition undergoes developmental changes that relate to the acquisition of literacy. From a literature search that initially provided 2377 papers, we reviewed 49 articles that explored the relationship between literacy experience and word level processing. We found conclusive evidence that literacy experience facilitated word recognition accuracy and speed at all levels of representation. Studies suggest that lexical representations regardless of type improved in quality, resulting in more efficient processing, and that results are similar across modalities. Several developmental trajectories were identified. Gaps in the literature are identified and recommendations for future research are provided.

5.1 Introduction

The invention of writing not only facilitated communication across vast distances and time, but also triggered systematic changes in humans' brain architecture, particularly pertaining to language. The acquisition of literacy induces structural changes in the human brain (Dehaene et al., 2015; Horowitz-Kraus & Hutton, 2015) that influence individuals' linguistic abilities. That is, as shown in previous chapters, participants who have more experience with written materials show higher proficiency in their word recognition abilities, even when accounting for the influence of general cognitive skills. The aim of the current study is to review the current literature on how our experience with written materials, also known as literacy experience, influences word recognition abilities. This review focused on several aspects related to word recognition.

First, word recognition depends on the efficiency of the activation of a word's lexical representation (representational quality), and literacy experience may influence this. A lexical representation consists of different types of information of the same lexical entry, namely semantics, phonology, orthography and syntax. According the Lexical Quality Hypothesis (Perfetti, 2007), the quality (i.e., level of precision and completeness) of representational types may differ within a single lexical entry. Thus, it may be the case that the semantic and phonological representation of a lexical entry are of high quality and therefore quickly activated, whereas the orthographic representation of the same lexical entry may be of lower quality, resulting in less efficient activation. Individuals with little literacy experience may show large quality difference between representation types. For example, they may have a particular difficulty activating an orthographical representation, because the orthographic representation is of low quality due to limited experience with written materials. At the same time, accessing the phonological or semantic representations of the same word may be efficient, since the quality of these representations does not necessarily depend on experience with written materials. We will review how literacy experience influences different types of representation – semantic, phonological, orthographical and syntactic.

Not only may literacy experience influence the quality of different representational types, but also the quality of the binding between different types. For example, the binding between a word's phonological and orthographical representation may be very strong in an individual who has a lot of reading experience. Grapheme-phoneme correspondences are thought to be strongly established in experienced readers (Ehri, 1995; Share, 1995), which would cause phonological representations to activate the orthographical representation and vice versa (cf. interactive models of word recognition, Grainger & Ferrand, 1994; McClelland & Rumelhart, 1981). This process of interactive activation between different representational types of the same lexical entry may be less efficient in less experienced readers. In these individuals, the binding between the orthographical and other representational types may be of low quality due to their limited experience with written materials. We will review how literacy experience influences the quality of binding between different types of representation.

Second, processes of word recognition operate on different levels of processing and therefore it may be possible that literacy experience influences word recognition differently at different levels. Some word recognition processes takes place on the word or lexical level, where a word's full orthographic form, phonological pronunciation, semantic meaning and syntactic information is accessed. Some recognition processes take place on the sublexical level, as word-level lexical representations consist of smaller, sublexical units (Grainger & Ferrand, 1994). For example, orthographical representations are a collection of grapheme units, and phonological representations are built from multiple phoneme units. Finally, although models of word recognition do not describe a representational level between the lexical and sublexical level, research on word production (Bock & Levelt, 1994; Levelt, 2001; Levelt, Roelofs, & Meyer, 1999) indicates that language processes can occur on a level that is in between the lexical and sublexical level. This "in-between" level is said to contain information regarding morphology, syllabification and stress patterns. Because experienced readers are thought to adopt a whole-word recognition strategy rather than a decoding strategy (Ehri, 1995), it may be the case that word recognition of highly literate individuals depends predominantly on lexical rather than sub-lexical level processes, whereas low literacy experienced individuals may tend to rely more on syllable/morphological or sublexical processes. The present review thus seeks to describe and understand at which level of representation (lexical, syllable/morphological, sublexical) literacy experience influences word recognition.

Third, we also examine how literacy experience influences word recognition separately for the spoken and written modality. From a theoretical point of view, it is important to make a distinction between modalities for two reasons. First, written and spoken word recognition are considered distinct processes, each described by their own set of theoretical models and frameworks. Due to the unique characteristics of the two modalities, written and spoken word recognition depend on different processes and operate on different types of representations and possible at different levels of representation. For example, written word recognition entails visual processing, relies on the activation of orthographic representations and may depend more on whole-word representations than sublexical representations, due to a written word's constituents being processed simultaneously rather than in a serial fashion. Because of the difference in word recognition processes posed by the inherent differences between the two modalities, we must review the literature for each modality separately. A second reason to distinguish written and spoken word recognition when examining the influence of literacy experience is that it informs us of how written and spoken language processing are related. If literacy experiences is found to improve recognition efficiency in not only the written but also spoken modality, it indicates that the influence of whichever aspects of literacy experience that causes this facilitation is not limited to the written modality, but transfers to the spoken modality. This would mean that exposure to written language affects the organisation of the language system and processes that operate on this system regardless of modality. By reviewing the influence of literacy experience on written and spoken word recognition separately, we may provide new insights that may unify modality-specific models of word recognition.

Finally, we decided to not only look at the effect of literacy on word recognition in adults that vary in their literacy experience, but also at how word recognition develops as children grow older and become more experienced with literacy. Children learn to read around the age of six (Seymour et al., 2003; Vaessen et al., 2010) and from that age onwards their development of language-related skills and processes, such as word recognition, coincides with their literacy acquisition. Arguably, direct comparisons of adults with children, or children of different ages are confounded by many differences between the samples. It is therefore difficult to distinguish the effect of literacy experience on word recognition from the effects of cognitive maturation, experience with language in general or increasing world knowledge. Instead of not reviewing this child literature at all due to these interpretation difficulties, we decided to review the literature on this topic, as we believe it to be worthwhile to review for two reasons. First, reviewing the child literature allows us to explore developmental trajectories related to literacy acquisition in word recognition for different types of representations, different levels of representations and different modalities. Second, a systematic review that describes the children's literature on this topic will provide insights into gaps in the current body of literature and will result in useful recommendations that may change how this topic is currently studied. The reader must, however, consider the results with regard to the children's literature as obtained from this systematic review in the light that we currently cannot distinguish the effect of literacy experience on word recognition from effects of maturation, experience with (spoken) language or increasing world knowledge.

In sum, the aim of the present study was to review the literature on the influence of literacy experience on word recognition ability. We explored the effects of literacy experience on 1) types of representation (semantic, phonological, orthographical, syntactic) and 2) levels of representation (lexical, syllable/morphological, sublexical). This was examined for written and spoken word recognition separately, and for both adults and children, to determine how word recognition is subjected to developmental changes related to literacy acquisition. The scope of the review is restricted to word recognition in alphabetic languages and individuals from neurotypical samples that are neither diagnosed with impairments in the domains of language, speech or reading, nor diagnosed with a learning or intellectual disability.

5.2 Methods

In line with the PRISMA statement (Moher, Liberati, Tetzlaff, & Altman, 2009), a systematic search of ERIC, Scopus, PsychINFO and Web of Science was performed to identify relevant articles published up January 2021. For the literacy experience part of the search term, we included the search terms "literacy", "print exposure" and "reading experience", which are all terms to describe an individual's accumulated experience with written materials. For the word recognition part, we included the terms "word comprehension", "word recognition", "speech comprehension" and "speech recognition" to capture word-level processing of both spoken and written language. Table 5.1 provides the search terms for each search engine and the number of hits. These searches yielded 3187 results including duplicates (see Figure 5.1).

Search results were exported and loaded into Zotero (version 5.0.23). After duplicate removal, 2286 unique publications remained (see Figure 5.1). An additional search was performed, because the authors noticed that some articles referred to word recognition as "lexical processing" or only referred to word recognition using the term "lexical decision task". A second search was performed where the literacy experience part of the search term was the same as in the previous search, but the word recognition part of the search term included "lexical decision" and "lexical processing". Table 5.2 provides the exact search terms and number of hits for this second search. This search resulted in 304 additional articles, or 180 after removing duplicates from within the second search and 91 after removing duplicates from the original search. These two searches resulted in 2377 unique publications (see Figure 5.1).

Table 5.1: Search terms of original search.

Search engine	Settings	Within	Search term	Hits
Web of Science	1900 to present; all languages; all document types	Titles	TS=(literacy OR print NEAR/0 exposure OR reading NEAR/0 experience) AND TS=(word NEAR/0 comprehension OR word NEAR/0 recognition OR speech NEAR/0 comprehension OR speech NEAR/0 recognition)	968
Eric	-	-	(literacy OR "print exposure" OR "reading experience") AND ("word comprehension" OR "word recognition" OR "speech comprehension" OR "speech recognition")	957
PsychInfo	1806 to present; via Ovid database	Keywords	(literacy OR "print exposure" OR "reading experience") AND ("word comprehension" or "word recognition" or "speech comprehension" or "speech recognition")	589
Scopus	All documents	Titles, abstract or keyword	(TITLE-ABS-KEY (literacy OR “print exposure” OR “reading experience”)) AND (TITLE-ABS-KEY ("word comprehension" OR "word recognition" OR "speech comprehension" OR "speech recognition"))	673

Table 5.2: Search terms of second, additional search.

Search engine	Settings	Within	Search term	Hits
Web of Science	1900 to present; all languages; all document types	Titles	TS=(literacy OR print NEAR/0 exposure OR reading NEAR/0 experience) AND TS=(lexical NEAR/0 decision OR lexical NEAR/0 processing	100
Eric	-	-	(literacy OR "print exposure" OR "reading experience") AND ("lexical decision" OR "lexical processing")	31
PsychInfo	1806 to present; via Ovid database	Keywords	(literacy OR "print exposure" OR "reading experience") AND ("lexical decision" OR "lexical processing")	93
Scopus	All documents	Titles, abstract or keyword	(TITLE-ABS-KEY (literacy OR "print exposure" OR "reading experience")) AND (TITLE-ABS-KEY ("lexical decision" OR "lexical processing"))	80

In the next step, titles and abstracts were screened based on several inclusion criteria. Articles were included if:

- The sample included participants over the age of 18, or, if the study included participants below 18 years old, there were at least two different age groups that were directly compared to each other. If participants were below 18 years old and only children from a single grade level were tested, the study was included if a measure of literacy experience was administered.

- Participants were neurotypical and did not have any developmental, intellectual, learning, reading, speech, language, cognitive or neurological disabilities or disorders, or any problems with hearing or sight unless fully corrected.
- The design of the study included a measure of spoken and/or written word recognition on the word level, where participants performed a silent word recognition task where they were required to fully activate a word-level lexical entry. For example, participants judged the lexicality of the word in a lexical decision task or had to categorize words based on semantic knowledge. Studies that used tasks such as phoneme or letter monitoring, were excluded, because these tasks do not require activation of the full word, but can be performed by only sublexical activation (Foss & Blank, 1980; Segui, Frauenfelder, & Mehler, 1981). In lexical decision or semantic categorization tasks, in comparison, the word-level lexical representation must be accessed to make a decision. Since we were interested in recognition on the word level, studies using lexical decision or semantic categorization tasks were included in the review.
- The design of the study included a measure of literacy experience, such as print exposure or reading frequency questionnaire. Print exposure measures, such as the Author Recognition Test, are a validated, recognized proxy of literacy experience. (Acheson et al., 2008; Dąbrowska, 2018; James et al., 2018; Mar & Rain, 2015; Payne et al., 2012; Stanovich & West, 1989). Studies that administered reading frequency questionnaires, although more subjective than measure such as the Author Recognition Test, were also included, as such measures may have been more often used in less literate populations, such as children or low-literate adults. Finally, studies were also included if they directly compared at least two groups of participants with distinct levels of literacy experience (literate participants vs. illiterate/low-literate participates, children of different ages or children vs. adults).
- The study was administered in a language with an alphabetic script.
- The study was administered in the native language of participants. Studies on second language learning (L2) were excluded.

- The article was written in English
- A full-text was available.

After this first screening, 72 articles were identified to be eligible for full-text review. Full-texts of these papers were retrieved and reviewed for inclusion. Twenty-four additional papers were identified by going through the reference list of these 72 articles. In the end, 49 papers were included in the final review, as displayed in Figure 5.1 below.

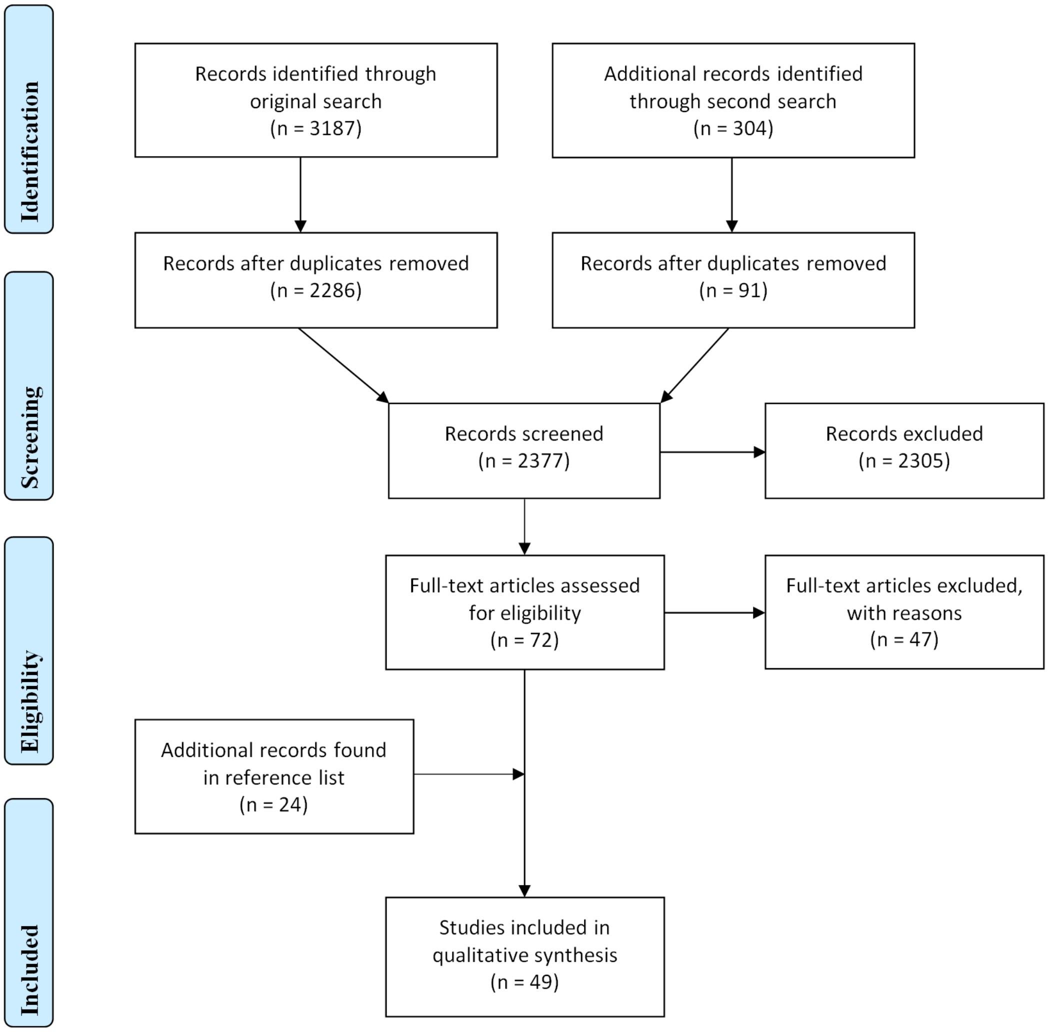


Figure 5.1: Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart of the literature search.

5.3 Results

This review is divided alongside two major themes. The first theme is the sample age that was under examination in the studies: adults aged 18 years or older varying in their literacy experience are discussed separately from studies comparing adults and children, or children of different age groups. The second theme is modality of word recognition: studies of spoken vs. written word recognition are discussed separately. Within each theme, studies exploring the same the same type or level of representation are grouped together. Appendix I displays descriptive information of each study, including language, sample information and task descriptions.

Adults aged > 18

Spoken word recognition.

Only one study examined the effect of literacy experience on spoken word recognition in adults. In their study, Kosmidis, Tsapkini, and Folia (2006) investigated word recognition in a group of Greek elderly females (M age = 67.91) that were illiterate and low educated, literate and low educated, or literate and highly educated. Using a spoken lexical decision task, they found that literacy experience influenced word recognition, as literate and highly educated participants recognized more nonwords (but not words) than participants who were illiterate. Education level also influenced word recognition, as highly educated participants showed better nonword recognition than low educated participants. These findings indicate that literacy itself as well as extensive literacy experience through education facilitates nonword recognition.

Written word recognition.

Eleven studies examined written word recognition in adults. All but one study were performed in English. The exception was a study in Korean (n = 1), a language with a transparent orthography that uses a (featural) alphabetic script called Hangul. All studies sampled from the university population, except for one, which looked at age effects in word recognition and sampled adults within the age range of 18 – 86 years (Cohen-Shikora & Balota, 2016). This was also the only study that differentiated variation in participants' literacy experience using age as a proxy, whereas all others utilized the Author Recognition Test (Stanovich & West, 1989). One study also used measures of vocabulary and reading comprehension to distinguish high and low literacy experienced participants (S. J. Unsworth & Pexman, 2003), and one study used a composite measure of the Author and Magazine Recognition test and a reading experience questionnaire (Lewellen, Goldinger, Pisoni, & Greene, 1993). All studies used a lexical decision task, and one study used an additional semantic categorization task (D. Hansen, Siakaluk, & Pexman, 2012), and one study administered an additional animacy judgement task (Cohen-Shikora & Balota, 2016). All studies found that literacy experience improved written word recognition (Chateau & Jared, 2000; Cohen-Shikora & Balota, 2016; D. Hansen et al., 2012; Lee et al., 2019; Lewellen et al., 1993; Michael, 2008; Sears et al., 2008; Thompson, 2011; S. J. Unsworth & Pexman, 2003; Welcome & Trammel, 2019), except for one (Kennedy, 1996). However, the lack of association found in this study may be due to the fact that the influence of other variables, including vocabulary and reading fluency, on word recognition were not partialled out. Given the high associations between these variables and literacy experience ($r = .40$ and $r = .48$ respectively) reported by Kennedy (1996), it is likely that the inclusion of these variables reduced the relative influence of literacy experience on written word recognition.

Semantic representations. D. Hansen et al. (2012) explored how literacy experience influences the activation of semantic representations, particularly sensorimotor semantic knowledge, during written word recognition. In this study, the body-object-interaction (BOI) value of words was manipulated. BOI represents perceptions of the ease with which a human body can physically interact with an object. Examples of high BOI words are “map”, “boot” or “mask” and examples of low BOI words are “mist”, “moon” or “ship”. In general, high BOI words are more easily recognized. D. Hansen et al. (2012) used a semantic categorization task where participants had to decide how easily imaginable a word was, and a phonological lexical decision task (“Does this written word *sound* like a real word?”). In additional analyses, the authors controlled for the relative speed benefit of high compared to low literacy experienced individuals.

Evidence was found for an effect of literacy experience on activation of sensorimotor semantic representations. The evidence suggested that high literacy experienced individuals have high quality semantic representations, regardless of whether a word is high or low BOI. The facilitative effect of high BOI words was found to be smaller in high literacy experienced compared to low literacy experienced individuals. This suggests that high literacy experienced participants do not show a recognition benefit for words that are easily interacted with (high BOI), because representations of less interactive words (low BOI) are already of high quality and thus easily accessed. Low literacy experienced individuals do benefit from their sensorimotor semantic knowledge, because it enables them to activate their lower quality semantic lexical representations more efficiently.

Moreover, evidence indicated that the quality of the binding between orthographic, semantic and phonological representations is higher for individuals with high literacy experience. In the phonological lexical decision task, the authors found a larger BOI effect for high literacy experienced than low literacy experienced participants. Thus, in literacy experienced individuals the activation of semantic representations, elicited by high BOI words, efficiently feeds forward towards the phonological representations, resulting in a recognition benefit. In low literacy experienced individuals, the binding between the different types of representations is not as strong and causes less activation of the phonological representation from the semantic representation, resulting in a smaller recognition benefit.

Phonological representations. Three studies investigated the influence of literacy experience on the activation of phonological representations during written word recognition (Chateau & Jared, 2000; Lewellen et al., 1993; S. J. Unsworth & Pexman, 2003). Phonology was manipulated in various ways. Some studies utilized homophones, which are words where a single phonological representation maps onto multiple distinct orthographical representations, e.g., /blu/ maps onto "blew" and "blue" (Lewellen et al., 1993; S. J. Unsworth & Pexman, 2003). Homophones are recognized slower than 'normal' words, because the orthographic representation ("blew") activates the phonological representation (/blu/), which then activates the 'other' orthographic representation ("blue"). Moreover, phonology was manipulated using homographs and irregular words (S. J. Unsworth & Pexman, 2003). Homographs refer to orthographic representations that map onto multiple distinct phonological representations, e.g., "tear" maps onto /tɛər/ and /tɪər/). Irregular words are words for which the correspondence between a word's graphemes and phonemes is uncommon. For example, "pint" is pronounced /paɪnt/, whereas the grapheme "i" more commonly corresponds to the phoneme /ɪ/ (as in "mint"). Homographs and irregular words are recognized slower than 'normal' words, because the orthographic form activates multiple phonological representations (e.g., /tɛər/ and /tɪər/ for the homograph "tear" and /aɪnt/ and /ɪnt/ for the irregular spelling "[p]int"). Another study (Chateau & Jared, 2000) used a primed lexical decision paradigm with primes that were either orthographically similar but phonologically dissimilar to the target (touch /tʌʧ/ - couch /kaʊʧ/), or unrelated (shall-couch). The orthographically similar but phonologically dissimilar prime “touch” activates the phonological representation /tʌʧ/, which impairs recognition of the target “couch” (/kaʊʧ/).

Evidence was reported for a reduced influence of phonological processing on written word recognition with increased literacy experience. Semantic categorization of homophones (Lewellen et al., 1993) and lexical decisions to homophones (“blew”), homographs (“tear”) and irregular words (“pint”) (S. J. Unsworth & Pexman, 2003) were faster in high than low literacy experienced participants. This indicates that the disadvantageous activation of the ‘other’ orthographic representation (“blue”) through the phonological representation (“blew”- /blu/), and the disadvantageous activation of multiple phonological competitors (/tɛər/ and /tɪər/ through “tear”, /aɪnt/ and /ɪnt/ through “pint”) was smaller in individuals with high literacy experience. In these individuals, written word recognition seems to be less strongly affected by phonological processing than individuals with low literacy experience.

Although the evidence suggested that increased literacy experience resulted in written word recognition to be less affected by phonological processing, the studies also suggested that the quality of the binding between orthographic and phonological representations increased with literacy experience. S. J. Unsworth and Pexman (2003) reported that when using a phonological lexical decision task (“Does this written word *sound* like a real word?”), literacy experienced participants were slower at recognizing homographs compared to normal words, whereas there was no difference for participants with less literacy experience. Thus, when participants were required to consider the lexicality of phonological representations (/tɛər/ and /tɪər/) that have been activated through an orthographic representation (“tear”), high literacy experienced participants were slower, because the strong binding between orthographic and phonological representations caused strong competition between the activated phonological candidates (/tɛər/ and /tɪər/). Moreover, Chateau and Jared (2000) reported that in low literacy experienced individuals a 30 ms or 60 ms prime that was orthographically similar but phonologically dissimilar (“touch”) did not influence word recognition of the target (“couch”), but that in high literacy experienced adults, the 60 ms prime inhibited word recognition. This indicated that in both groups the phonological representation of the prime was not activated initially (30 ms), but that it became strongly activated later (60 ms) in high literacy experienced individuals only, thereby impairing word recognition of the orthographically similar but phonologically dissimilar target (“couch”- /kaʊʧ/). This means that for high but not low literacy experienced individuals, strong bindings between phonological and orthographic representations result in fast and automatized feedback between the two types of representations.

Orthographic representations. Three studies explored how literacy experience influences the activation of orthographic representations during written word recognition (Chateau & Jared, 2000; Sears et al., 2008; Welcome & Trammel, 2019). The studies varied in how orthography was manipulated. One study used an orthographic priming paradigm, where targets paired with an orthographically related prime (jonis-joins) are assumed to be recognized more efficiently than target words paired with orthographically primes (mmkaes-north) (Welcome & Trammel, 2019). The other two studies used a lexical decision task with pseudohomophones as nonwords, manipulating orthographic typicality of these pseudohomophones (Chateau & Jared, 2000; Sears et al., 2008). Pseudohomophones are nonwords that sound like real words but are written differently, such as “fale” (derived from “fail”). Recognition of pseudohomophones is less efficient compared to ‘normal’ nonwords, because the orthography of the pseudohomophones (“fale”) activates a phonological representation (/feɪl/), which then activates the orthographic representation of a real word (“fail”). In these studies, some of these pseudohomophones were spelled using typical orthographic patterns (“fale” derived from “fail”), whereas others were orthographically atypical (“cawph”, derived from “cough”).

The studies reported evidence for increased activation of orthographic representation with extensive literacy experience. Participants with high compared to low literacy experience showed more efficient recognition of pseudohomophones (Chateau & Jared, 2000; Sears et al., 2008). More importantly, the effect of orthographic atypicality was smaller in individuals that were more literacy experienced. Low literacy experienced individuals recognized pseudohomophones, that were orthographical atypical and therefore very unword-like (“cawph”), faster than orthographical typically, more word-like pseudohomophones (“fale”), whereas the speed difference was smaller in literacy experienced individuals. This indicated that orthographic processing is less efficient in low literacy individuals, as they show difficulty recognizing nonwords that are spelled according a language’s typical orthographic patterns.

Evidence was provided suggesting that participants with lower literacy skills compensate for their suboptimal orthographic processing with their phonological processing skills. Sears et al. (2008) only reported an effect of literacy experience when using a lexical decision task where pseudohomophones ("brane", derived from "brain") were used as nonwords, but not when using a task that used 'normal' words as nonwords (e.g., "grun"). In lexical decision tasks with pseudohomophones as nonwords, participants can only rely on their orthographic knowledge to decide whether the word is a real word ("brain") or a pseudohomophone ("brane") and therefore cannot use their phonological knowledge as a 'back-up' source of information. Thus, in the phonological lexical decision task, low literacy experienced participants could not use their phonological skills to compensate for their suboptimal orthographic processing ability, and found it difficult to recognize pseudohomophones ("brane") as a nonwords.

Findings also indicated that orthographic representations become more precisely specified with extensive literacy experience. Welcome and Trammel (2019) found that individuals who were less experienced with literacy showed a larger recognition benefit than high literacy experienced individuals for an orthographically related prime (jonis-joins) compared to an unrelated prime (mmkaes-north). This indicates that orthographic representations in high literacy experienced individuals are complete and precisely specified, as a prime that deviates slightly from the target will not activate the target word. As a result, their word recognition of the target is not facilitated. Participants with less literacy experience will have representations of lower quality, and thus the presentation of a prime that slightly deviates from the target will activate the target word, improving their word recognition.

Finally, in their ERP-study, Welcome and Trammel (2019) shed light on the stage of the written word recognition where literacy experience may have an effect. During the early, pre-lexical orthographic activation stage of written word recognition (125 – 300 ms) literacy experience did not affect the size of an orthographic priming effect, but during the stage of lexical activation (400 – 600 ms), literacy experience affected the size of the priming effect: individuals who were more experienced with literacy showed smaller priming effects. This indicates that initial visual/orthographic processing is not affected by literacy experience, but that stage of lexical access literacy experience seems to have a facilitatory effect.

Word level: word frequency, word familiarity and neighbourhood size. Five studies examined whether literacy experience influences recognition of written words that vary with respect to word-level characteristics, namely word frequency, word familiarity and neighbourhood size (Chateau & Jared, 2000;

Cohen-Shikora & Balota, 2016; Lee et al., 2019; Lewellen et al., 1993; Sears et al., 2008). To approximate the these characteristics, authors used measures of word frequency or neighbourhood size from corpora (Chateau & Jared, 2000; Cohen-Shikora & Balota, 2016; Lewellen et al., 1993; Sears et al., 2008), or obtained subjective familiarity ratings in a separate task (Lee et al., 2019). In general, frequent and familiar words are more efficiently recognized than infrequent and unfamiliar words. Moreover, words with many orthographic neighbourhood are recognized more efficiently than words with few orthographic neighbours. Studies were administered in English, except for one which was performed in Korean (Lee et al., 2019). Most studies sampled college-age adults, except for one study that looked at written word recognition across the adult life span (age range: 18-86 years) (Cohen-Shikora & Balota, 2016).

Results suggested that literacy experience increases the quality of lexical representations of infrequent and unfamiliar words, as most studies reported the difference in recognizing frequent-infrequent or familiar-unfamiliar words to decrease as participants' literacy experience increased (Chateau & Jared, 2000; Cohen-Shikora & Balota, 2016; Lee et al., 2019). One study did not report this trend (Lewellen et al., 1993).

Furthermore, the studies indicated that literacy experience improves the quality of orthographic representations of words with small neighbourhood sizes. Initial findings were inconsistent, with Chateau and Jared (2000) reporting that for infrequent words, differences in recognizing words with many or few orthographic neighbours decreased as participants had more literacy experience, but Lewellen et al. (1993) reporting no such effect of literacy experience. Crucially, Sears et al. (2008) showed that these contrasting findings were a result of task differences between the two studies. They replicated the null-effect of literacy experience only when using a regular lexical decision task, like Lewellen et al. (1993). However, when administering a lexical decision task with pseudohomophones as nonwords ("brane"), such as the one Chateau and Jared (2000) used, Sears et al. (2008) replicated their finding that with literacy experience, the neighbourhood size effect for infrequent words decreased. Thus, when low literacy experienced participants cannot rely on their phonological representations do decide whether a target is a word or nonword ("brane"), their word recognition is facilitated by the activation of many orthographic neighbours. High literacy experienced participants' orthographic representations are of higher quality, causing these individuals to rely less on phonological representations during word recognition, and allowing for successful lexical access even if there are only few orthographic neighbours to contribute to the activation of the correct lexical candidate.

Comparison adults and children, or children of different age groups

All of the following results must be considered while keeping in mind that, for these studies, it is impossible to distinguish effects of literacy experience on word recognition from effects of cognitive maturation, general (spoken) language experience and world knowledge, as children and adults, and children of different ages inherently differ on these aspects.

Spoken word recognition.

In total six studies examined spoken word recognition in children of different age groups, or children and adults. Studied languages were German (n = 2), English, French, Dutch and Portuguese (all n = 1). Children's ages varied between five (pre-reading age) and eleven years old. Three studies also tested college-aged adults. Five studies used a lexical decision task, and one study used a semantic categorization task. In general, studies found that word recognition was more accurate and faster in older children and adults (Gijsel, Ormel, Hermans, Verhoeven, & Bosman, 2011; Lin, Wang, Newman, & Li, 2018; Pattamadilok, Morais, De Vylder, Ventura, & Kolinsky, 2009; Schild, Becker, & Friedrich, 2014; Schild, Roder, & Friedrich, 2011; Ventura, Morais, & Kolinsky, 2007).

Semantic representations. Gijsel et al. (2011) explored the activation of semantic representations during word recognition in Dutch children aged seven to eleven. Both horizontal semantic knowledge, where one has to match lexical entries from the same taxonomic category (e.g., "Choose the word that best matches orange", when presented with "cherry", "egg", "snow" and "ball"), and vertical semantic knowledge, where one has to choose a lexical entry when presented with a taxonomic category ("choose a fruit", when presented with "cherry", "egg", "snow" and "ball") were studied. The task was performed with spoken targets, written targets or pictorial targets.

They reported that as children grew older, their semantic lexical representations became more precise, as older children were more accurate and faster at categorizing words in both the vertical and horizontal task. Moreover, this study showed that the bindings between semantic, phonological and orthographic representations are quite strong at already an early age, because the effect of age on word recognition was not influenced by the modality in which the words were presented.

Phonological representations. Two studies investigated the activation of phonological representations during spoken word recognition in college-aged adults and children aged six to ten (Lin et al., 2018; Schild et al., 2014). Schild et al. (2014) also included a group of pre-reading kindergartners, who were not able to read. Amongst other variables, phonology was manipulated. In the paradigm of Lin et al. (2018), nonwords were created by changing one or two phonemes of real words ("calin" derived from "cabin"). In Schild et al. (2014)'s primed lexical decision paradigm, primes were manipulated to match or mismatch the target word (e.g., "monster") with regard to phoneme onset (e.g., matching prime: "mon", mismatching prime: "tep").

The studies showed that younger children compared to older children and adults rely more on phonological processing during word recognition. Schild et al. (2014) reported the facilitatory effect of phoneme overlap to be larger in six-year old pre-reading kindergartners and seven-year olds who were able to read compared to adults. Lin et al. (2018) reported that six- and eight year olds were worse at recognizing nonwords created by changing one or two phonemes of real words than ten-year old children and adults, indicating that the older children and adults relied less on phonological processing during the recognition of the nonwords.

Moreover, it seemed that in younger children, activation of phonological representations lingers on longer during the word recognition processes. Schild et al. (2014) reported that age affected an ERP component thought to capture late phonological priming effects (400 – 1000 ms after stimulus presentation). For pre-kindergartners and six-year olds compared to seven-year olds and adults, phoneme-matching priming conditions (mon-monster) elicited more negative amplitudes than phoneme-mismatching priming conditions (tep-monster) over right posterior regions. This unique ERP component capturing late effects of phonological priming disappeared with age, indicating that only in younger children phonological information remains activated for a longer period of time.

Orthographic representations. Two studies investigated the activation of orthographic representations during spoken word recognition in college-aged adults and children aged five to ten (Pattamadilok et al., 2009; Ventura et al., 2007).

Both studies manipulated the rime consistency of the words. Some target words ended with a rime that could only be spelled in one way (consistent rime, e.g., the Portuguese /um/ can only be spelled as “ume”), whereas other target words ended with a rime that could be spelled in multiple ways (inconsistent rime, e.g., /ɛl/ can be spelled with both “ele” or “el”). Word recognition is generally facilitated for words that have consistent rimes (i.e., can only be spelled in one way), indicating that spoken word recognition is influenced by orthographic representations.

The studies reported seemingly inconsistent results regarding age effects on the activation of orthographic representations during spoken word recognition. Ventura et al. (2007) found the influence of orthographic representations to decline with age, as children (pre-reading kindergartners, children aged eight to ten) showed faster word recognition facilitation for rime-consistent words, but adults’ recognition speed did not differ for consistent or inconsistent words. Pattamadilok et al. (2009) reported overall little evidence for activation of orthographic representations, except for a sudden surge around the age of nine: only children aged nine, but not children aged eight or ten or adults, showed an accuracy benefit for consistent rimes.

Interestingly, Ventura et al. (2007) provided evidence that suggest that orthographic activation differs when recognizing words and nonwords, and that this pattern may be subjected to age effects. In this analysis, they also controlled for the relative speed benefit of adults over children. They reported that children’s word and nonword recognition was facilitated by rime consistency, whereas in adults rime consistency only improved word but not nonword recognition. This suggests that regardless of age, orthographic representations are activated when recognizing words. However, when recognizing nonwords, orthographic representations are only activated in children, but not adults. This may indicate, according to the authors, that the binding between phonological and orthographical units are more automatically and easily triggered (i.e., have a lower threshold of activation) in children than adults. A possible cause of this may be that children rely more on a decoding strategy, which requires swift activation and strong connections between phonological and orthographic units. In children, even when recognizing spoken nonwords, the word recognition system automatically activates of orthographic units. Adults, on the other hand, do not rely as much on such decoding recognition strategies, and the disuse of the connections between phonological and orthographic units may lead to a less automatic flow of activation during spoken word recognition.

Lexical stress level. Processing on the level of lexical stress during word recognition was studied in two studies in English and German speaking children ranging from the age of six (pre-reading age) to ten years old and college-aged adults (Lin et al., 2018; Schild et al., 2014). Both studies manipulated, amongst others, stress, either by creating nonwords from real words by changing the stress pattern (e.g., nonword "cabin" derived from cabin) (Lin et al., 2018), or manipulating the overlap between a prime and target's stress (overlap: "mon-monster", no overlap: "mon-monster") (Schild et al., 2014).

Results regarding the effect of age on stress-level processing were inconclusive, with Lin et al. (2018) reporting stress-level processing during word recognition to increase with age, as adults and ten-year olds compared to six and eight year olds showed improved nonword recognition for nonwords that were real words with a changed stress pattern ("cabin"). On the other hand, Schild et al. (2014) reported no evidence for stress-level processing in any age group, as word recognition was not facilitated by stress-matching primes ("mon-monster") compared to stress-mismatching primes ("mon-monster"). These difference findings are unlikely due to the difference in language, since English and German languages use very similar stress systems (Domahs, Plag, & Carroll, 2014). It may be that the stress manipulation in the priming paradigm used by Schild et al. (2014) was too subtle, even for adults, and therefore did not influence word recognition.

Written word recognition.

Thirty-two studies examined written word recognition in children of different age groups, or children and adults. Studies were performed in English (n = 12), French (n = 7), German (n = 4), Spanish (n = 4), Italian (n = 2), Dutch (n = 2) in Hebrew (n = 1). Samples included children ranging from the age of seven to eighteen, and 15 studies also included college-aged adults. Three studies followed a longitudinal design, testing the same group of children over the course of two to four years. Fifteen studies used a primed lexical decision paradigm, eleven studies used a regular lexical decision task, two studies administered a go/no go lexical decision task and four studies used a semantic matching or categorization task. All studies found word recognition speed and accuracy to be higher in older children and adults (Beyersmann, Castles, & Coltheart, 2012; Beyersmann, Grainger, Casalis, & Ziegler, 2015; Beyersmann, Grainger, & Castles, 2019; Burani, Marcolini, & Stella, 2002; Castles, 1999; Castles et al., 2007; Colombo, Sulpizio, & Peressotti, 2017; D'Alessio, Wilson, & Jaichenco, 2019;

C. Davis, Castles, & Iakovidis, 1998; Dawson, Rastle, & Ricketts, 2018; Ducrot, Pynte, Ghio, & Lété, 2013; Fleischhauer, Bruns, & Grosche, 2021; Gijsel et al., 2011; Grainger, Lété, Bertand, Dufau, & Ziegler, 2012; Hasenäcker, Beyersmann, & Schroeder, 2020; Hasenäcker & Schroeder, 2017b; Hasenäcker, Schröter, & Schroeder, 2017a; Kezilas, McKague, Kohnen, Badcock, & Castles, 2017; Luque et al., 2020; Lázaro, Acha, de la Rosa, García, & Sainz, 2017; McCutchen, Logan, & Biangardi-Orpe, 2009; Perdijk, Schreuder, Baayen, & Verhoeven, 2012; Polse & Reilly, 2015; Quémart, Casalis, & Colé, 2011; Quémart, Casalis, & Duncan, 2012; Quémart, Gonnerman, Downing, & Deacon, 2018; Samuels, LaBerge, & Bremer, 1978; Schiff, Raveh, & Fighel, 2012; Sprenger-Charolles, Siegel, Béchennec, & Serniclaes, 2003; Waldie & Mosley, 2000; Ziegler, Bertrand, Lété, & Grainger, 2014). Although one study (Acha & Perea, 2008) did not report a main effect of age, their descriptive statistics (p.251, Table 1) show clear difference between different age groups, with word recognition improving with age.

Semantic representations. Seven studies examined the activation of semantic representations during written word recognition in college-aged adults and children aged seven to fourteen (Beyersmann et al., 2019; Fleischhauer et al., 2021; Gijsel et al., 2011; McCutchen et al., 2009; Polse & Reilly, 2015; Quémart et al., 2011, 2018). Most studies used a primed lexical decision paradigm with, amongst others, semantically related primes (e.g., “strategy-plan”) (Beyersmann et al., 2019; Fleischhauer et al., 2021; McCutchen et al., 2009; Quémart et al., 2011, 2018). Gijsel et al. (2011) and Polse and Reilly (2015) administered a semantic categorization task where participants had to semantically match a word (e.g., “woman” or “water”) to a target (“lady”). The task of Gijsel et al. (2011) has already been discussed in this review’s section on spoken word recognition, as target words were either presented auditorily, orthographically or pictorially.

Studies were somewhat inconsistent regarding the effect of age on semantic processes during word recognition. One study reports no evidence for semantic processing (Fleischhauer et al., 2021). Some studies found that the influence of semantic knowledge on word recognition efficiency increases with age, with more efficient semantic categorization (Gijsel et al., 2011; Polse & Reilly, 2015), and larger semantic priming effects in adults and older children (age ten-eighteen) than younger children (age nine-eleven) (Beyersmann et al., 2012). Other studies reported semantic knowledge to affect word recognition regardless of age (McCutchen et al., 2009; Quémart et al., 2011, 2018).

This disparity in findings may be influenced by task type. The studies that did not report an age effect or no semantic effects at all manipulated not only semantic overlap between primes and targets, but also morphological and orthographic overlap. Morphological primes also tended to overlap orthographically. Thus, in these studies primes were orthographically related in the majority of the conditions (e.g., orthographically related: "abricot-abri" [apricot-shelter]; morphologically related: "tablette-table" [tablet-table), pseudomorphologically related: "baguette-bague" [bread-ring]), except for in the semantic condition ("tulipe-fleur" [tulip-flower]). Orthographic overlap was therefore much more salient than semantic overlap, which may explain the reduction of semantic effects and the absence of an age effect.

A possible mechanism that may explain the increased semantic effects with age reported by some studies is that adults may have quicker access to semantic information. Quémart et al. (2011) showed that children nine to fifteen showed semantic priming effects when a semantically related prime was shown for 800 ms, but not when the prime was shown for 250 ms. Crucially, adults showed a trending ($p = .09$) semantic priming effect when the prime was shown for 250 ms. This may suggest that adults access semantic representations faster than children.

Moreover, evidence suggested that in young children (aged seven and eight) word recognition relies heavily on the orthographic processing system, and that with age they increasingly use their semantic representations during the word recognition process. Polse and Reilly (2015) found that younger children aged seven and eight showed efficient word recognition when categorising words orthographically (matching "lady" with "lazy"), but less efficient word recognition when semantic categorization was required ("lady" with "woman"), particularly when an orthographic foil was present ("lazy"). Children aged nine and ten became increasingly more efficient regarding semantic processing even in the presence of an orthographic foil. This indicates that with age, they rely less on orthographic representations and more on semantic representations during written word recognition.

Phonological representations. Five studies investigated the activation of phonological representations in written word recognition in college-aged adults and children aged seven to fourteen (C. Davis et al., 1998; Grainger et al., 2012; McCutchen et al., 2009; Sprenger-Charolles et al., 2003; Ziegler et al., 2014). Three studies administered a primed lexical decision task where primes were, amongst others, homophones ("maid-made") (C. Davis et al., 1998), homophones ("vaze-vase") (Ziegler et al., 2014) or varied in the degree of phonological overlap with the target ("planner-plan" vs. "dept-deep") (McCutchen et al., 2009). Other studies used a lexical decision paradigm (Grainger et al., 2012) or semantic categorization task (Sprenger-Charolles et al., 2003) with, amongst others, pseudohomophones as nonwords ("trane", derived from "train").

With regard to the effect of age on phonological processing during word recognition, contrasting findings were reported. Some studies reported the effect of phonological processing during word recognition to remain stable across age (C. Davis et al., 1998; McCutchen et al., 2009; Ziegler et al., 2014), whereas others reported the effect of phonological manipulations to decrease (Grainger et al., 2012; Sprenger-Charolles et al., 2003), indicating an increase in the efficiency of phonological processing. Note that the authors that reported a decrease included younger participants (minimally aged seven), and that two studies that found no developmental trajectories only tested older children (minimally aged eleven) (C. Davis et al., 1998; McCutchen et al., 2009). Ziegler et al. (2014) did not report a developmental trajectory either and tested children as young as seven, but this null-effect might be related to their paradigm. A primed lexical decision task, with phonologically manipulated primes presented for only 70 ms, might be less suitable for young, beginning readers than a regular lexical decision task (Grainger et al., 2012) or a semantic categorization task (Sprenger-Charolles et al., 2003). Thus, when taking into account the differences in age range and paradigm between studies, the literature suggests that as children become older, they depend less on phonological processes (i.e., decoding) during word recognition, as shown by findings that with age, their word recognition is affected to a lesser degree by phonological manipulations.

Orthographic representations. Seventeen studies examined activation of orthographic representations during written word recognition in college-aged adults and children aged seven to twelve. Eight of these studies specifically looked at morpho-orthographic processes and will be discussed later in this review's section on written word recognition on the morphological level (Beyersmann et al., 2012, 2015; Fleischhauer et al., 2021; Hasenäcker et al., 2020; McCutchen et al., 2009; Quémart et al., 2011, 2018; Schiff et al., 2012). In this part of the review, we will discuss the remaining nine studies (Acha & Perea, 2008; Castles, 1999; Castles et al., 2007; Colombo et al., 2017; Grainger et al., 2012; Kezilas et al., 2017; Polse & Reilly, 2015; Sprenger-Charolles et al., 2003; Ziegler et al., 2014). Five studies used a primed lexical decision task, manipulating, amongst others, the orthographic relationship between the prime and target (Acha & Perea, 2008; Castles, 1999; Castles et al., 2007; Kezilas et al., 2017; Ziegler et al., 2014). This manipulation occurred either by transposing two letters of the target (e.g., “aminal-animal”) or by substituting two letters of the target (e.g., “arisal-animal”). Two studies used a regular lexical decision paradigm, with, amongst others, nonwords that were made from real words by transposing two letters (e.g., “talbe”) (Colombo et al., 2017; Grainger et al., 2012). Finally, two studies administered a semantic categorization task (Sprenger-Charolles et al., 2003) with, amongst others, orthographic nonwords (e.g., “rouqe”, derived from “rouge” [red]), or a semantic matching task (Polse & Reilly, 2015) with the relationship between targets and foils being manipulated with regard to, amongst others, orthography (e.g., match: “lady-lady”, mismatch: “lady-lazy”).

Findings indicated that orthographic representations are already of sufficient quality to enable efficient word recognition at an early age, and that the quality of these representations seems to increase with age. Already in children as young as seven years old, orthographic overlap between a prime and target improved word recognition, and nonwords that orthographically overlapped with real words were more difficult to recognize as nonwords (Polse & Reilly, 2015). Moreover, effects of orthographic priming increased with age (Sprenger-Charolles et al., 2003).

Importantly, studies suggested that one particular aspect of orthographic representations increased in quality with age: namely letter identity. In younger children aged nine, but not older children and adults, specific letter identities within lexical representations were not yet strongly established. Only younger children showed word recognition facilitation for words (“animal”) that were preceded by a prime where one or more letters are substituted (“arisal”) (Castles et al., 2007). In older children and adults, the letter identities within an orthographic representation are precise and complete and as a result their word recognition is not facilitated by a prime that deviates from the target.

Furthermore, evidence suggests that, in the words of Colombo et al. (2017), (p.56): "(...) there is an increasing reliance [during written word recognition] on an orthographic representation in which letter position is not specified". Only older children and adults, but not younger children showed facilitation effects for primes where the position of two letters are switched (“aminal”, derived from “animal”). This facilitative transposed letter effect is found to increase from age seven to age eleven and in adults (Kezilas et al., 2017; Ziegler et al., 2014), or, in one case, shows a decrease until age nine, only to be increased again at age eleven and in adults (Grainger et al., 2012). This suggest that with age, word recognition relies less on specific letter positions within a lexical representation and that word recognition can be efficient even if the letters within a word are not placed in the correct order. Only Acha and Perea (2008) reported no increase of the transposed letter effect. However, this may be the result of their older age group, as they tested children aged nine and twelve, whereas the other studies also tested children as young as seven years old.

Finally, evidence was reported which suggested that with age nonword recognition increasingly relies on serial processing strategies. Colombo et al. (2017) found transposition and substitution effects to be larger when they occurred at the beginning of nonwords (“ablergo/acmergo”, derived from "albergo" [inn]) rather than at the end of nonwords (“leopadro/leopatso”, derived from "leopardo" [leopard]) and that this effect increased with age. This indicates that as children grow older, they use a serial scanning or decoding mechanism when encountering nonwords, thereby detecting transpositions or replacements faster when they occur at the beginning compared to the end of a word.

Morphological level. In total 17 studies, of which 13 obtained through references, investigated processing on the morphological level during word recognition in college-aged adults and children aged seven to seventeen (Beyersmann et al., 2012, 2015, 2019; Burani et al., 2002; D’Alessio et al., 2019; Dawson et al., 2018; Fleischhauer et al., 2021; Hasenäcker et al., 2020, 2017a, 2019; Lázaro et al., 2017; McCutchen et al., 2009; Perdijk et al., 2012; Quémart et al., 2011, 2012, 2018; Schiff et al., 2012). Studies were administered in English, German, French, Dutch, Italian and Hebrew. Nine studies used a priming paradigm and manipulated the morphological relationship between the target and the prime (Beyersmann et al., 2012, 2015, 2019; Fleischhauer et al., 2021; Hasenäcker et al., 2020; McCutchen et al., 2009; Quémart et al., 2011, 2018; Schiff et al., 2012). One study used a go/no go lexical decision task (Lázaro et al., 2017) and the other studies a ’regular lexical decision task’ (Burani et al., 2002; D’Alessio et al., 2019; Dawson et al., 2018; Hasenäcker & Schroeder, 2017b; Hasenäcker et al., 2017a; Perdijk et al., 2012; Quémart et al., 2012). Morphology was manipulated in various ways, including morphological complexity (monomorphemic vs. multimorphemic words), morpheme frequency or morphological family size. Studies examined lexical decisions on words, nonwords, or compound words.

Regarding the influence of age on morphology-level processes during word recognition, some studies reported changes with age (Beyersmann et al., 2012; Dawson et al., 2018; Fleischhauer et al., 2021; Hasenäcker et al., 2020; Hasenäcker & Schroeder, 2017b; Hasenäcker et al., 2017a; Perdijk et al., 2012; Quémart et al., 2011, 2012; Schiff et al., 2012), whereas others reported effects to be stable across ages (Beyersmann et al., 2015, 2019; Burani et al., 2002; D'Alessio et al., 2019; Lázaro et al., 2017; McCutchen et al., 2009; Quémart et al., 2011). Of the studies that did not find an effect of age on morphological processing during word recognition, two were administered in Spanish (D'Alessio et al., 2019; Lázaro et al., 2017), and one in Italian (Burani et al., 2002). Both languages are transparent languages with simple syllabic structures (Seymour et al., 2003) and rich morphology (Bane, 2008; Moscoso del Prado, 2011). The lack of an effect of age on morphological processing during word recognition, may indicate that in these transparent languages with simple syllable structures but rich morphological system, sophisticated morphology-level processing is already acquired at an early age.

In the literature on morphological processes during word recognition, a distinction is made between different levels of morphological processing: morpho-semantic and morpho-orthographic processing (Rastle & Davis, 2008). Morpho-semantic processing entails morphological decomposition of words based on semantic overlap between full words and stems. For example, "singer" is decomposed in the stem *sing* and the suffix *-er*, because "singer" and "sing" share a semantic relationship. The word "corner", however, is not decomposed in the stem *corn* and the suffix *-er*, because "corner" and "corn" do not share a semantic relationship. On the other hand, morpho-orthographic processing refers to morphological decomposition in the absence of semantic relationships. At this level of processing, both "singer" and "corner" are decomposed in the stem and the suffix *-er*, regardless of the semantic relationship between the full word and the stem. It is thought that morpho-orthographic decomposition is performed through affix-stripping, where the affix (-er) is stripped from the stem (sing/corn). In addition, evidence has been found for an additional morphological process called embedded stem activation, where morphological decomposition takes place on the basis of stems rather than affixes. For example, "turnip" is decomposed in the stem *turn* (not semantically related to the full word) and *-ip*, which is not a true suffix. Discussing the debate to what extent adults engage in morpho-semantic processing, morpho-orthographic processing through affix stripping, or embedded stem activation during word recognition is beyond the scope of this review.

Rather, it is worthwhile to examine this from a developmental perspective and review the literature regarding the effect of age on these types of morphological processing.

There is major evidence that morpho-orthographic processing is only acquired at a later age, with younger children relying more on morpho-semantic processing. Studies reported that morpho-orthographic processing was only present in children aged ten or older and adults, but not in younger children, whereas at all ages morpho-semantic processing took place (Beyersmann et al., 2012; Fleischhauer et al., 2021; Quémart et al., 2018; Schiff et al., 2012). These results were obtained in English-, Hebrew- and German-speaking participants. However, morpho-orthographic processing may already be acquired at an earlier age in languages, such as French, with simple syllabic structures but complex morphological systems (Bane, 2008; Moscoso del Prado, 2011), suggested by studies in French that reported young children aged nine to already show evidence of morpho-orthographic processing (Quémart et al., 2011).

Studies also reported evidence for morpho-orthographic decomposition based on affix-stripping being acquired at a later stage of development than morphological processing based on embedded stem activation. Embedded stem activation was found to be already present in children aged eight (Beyersmann et al., 2015; Quémart et al., 2012), whereas evidence for affix-stripping was only present around the age of nine (Hasenäcker et al., 2020). This is also corroborated by studies who reported efficient processing of compound word, which are essentially words build from stems only, already in children from the age of eight (Beyersmann et al., 2019; Hasenäcker et al., 2017a), whereas efficient processing of affixes only occurred at around the age of ten (Hasenäcker et al., 2017a). Moreover, it seems that simultaneous processing of a stem and an affix has detrimental effects on young children's (aged nine), but not older children's (aged eleven) word recognition (Quémart et al., 2012), indicating that morphological processes related to affixes are indeed acquired at a later age.

Syllable level. Two studies examined written word recognition on the syllable level in German- and Spanish-speaking adults and children aged eight to twelve (Hasenäcker & Schroeder, 2017b; Luque et al., 2020). One study used a lexical decision task manipulating syllable boundary (congruent "Fah:rer" [fa:ther] or incongruent "Fa:hrer" [fath:er]) (Hasenäcker & Schroeder, 2017b), and the other manipulated the frequency of the initial syllable (Luque et al., 2020).

The results suggest that a language's syllabic complexity has a large influence on age effects with regard to syllable-level processing during word recognition. In language with a more complex syllabic structure, such as German(Seymour et al., 2003), a clear developmental trajectory of syllable processing is reported with more efficient syllable processing with age. Hasenäcker and Schroeder (2017b) reported that adults were able to recognize words regardless whether a word was visually parsed (using ":") along the true syllable boundary of the word ("Fah:rer"), whereas word recognition was hampered in children aged eight and ten if the word contained an incongruent syllable boundary ("Fa:hrer"). Also, in these complex syllabic languages, syllable and morphological processing seem to become separated with age. Hasenäcker and Schroeder (2017b) showed that eight-year-old children showed difficulties recognizing nonwords with incongruent syllable boundaries, regardless of whether the word contained multiple morphemes. Ten-year-old children only showed recognition difficulty for nonwords with incongruent syllable boundaries if these nonwords contained multiple morphemes ("Hel:ber" [dri:ber]) but not for monomorphemic words ("Dos:tor" [dor:tar]). Adults' nonword recognition was not influenced by neither syllable congruency nor morphological complexity.

In languages with more simple syllabic structures such as Spanish (Seymour et al., 2003), no developmental trajectory of syllable-level processing during word recognition was reported. This suggests that in these languages efficient syllabic processing during word recognition is acquired early. Luque et al. (2020) reported that eight- and ten-year old children were equally sensitive to syllable frequency. At both ages, children found it difficult to respond to low-frequency words with a high frequent initial syllable. Word recognition of such words is disrupted, because the highly frequent syllables activate many other lexical candidates, increasing competition among lexical candidates. The effect is particularly strong for low frequency words, because these words have a lower activation threshold than frequent words (Alvarez, Carreiras, & De Vega, 2000).

Word level: word length. Two studies looked at how recognition of written words in college-aged adults and children aged eight to twelve differs for words varying in one specific word-level characteristics, namely word length (Acha & Perea, 2008; Samuels et al., 1978). In general, shorter words are recognized faster and more accurately than longer words.

Studies indicated that with age a shift occurs in word recognition strategy from a segmentation and (phonological) decomposition strategy to a more holistic, whole-word recognition strategy. Both studies reported that with age, the word length effect decreased. Younger children showed large speed and accuracy differences at recognizing short vs. long words, but older children and adults recognized words efficiently, irrespective of length. Thus, with age, individuals adopt a word recognition strategy that relies more on whole-word representations.

Word level: word frequency and neighbourhood size. Five studies examined how recognition of written words in college-aged adults and children aged eight to twelve differs for words varying in word frequency and neighbourhood size (Burani et al., 2002; Castles, 1999; D'Alessio et al., 2019; Ducrot et al., 2013; Luque et al., 2020). All studies manipulated, amongst others, word frequency or subjective ratings of a word's neighbourhood size. The frequency or neighbourhood size effect refers to frequent words or words with more orthographic neighbours being recognized more accurately and faster than infrequent words or words with few orthographic neighbours.

Studies suggest an U-shaped developmental trajectory of word frequency effects, with frequency effects being present in young children aged seven, but decreasing with age (Ducrot et al., 2013). The word frequency effect, although small in size, still remains present (Burani et al., 2002; Luque et al., 2020), and increases again around the age of twelve (D'Alessio et al., 2019). This U-shaped trajectory may indicate that young children show recognition facilitation for high frequency words, because they might not yet know the low frequency words. With age, their knowledge of low frequency words catches up, but there is little difference in the quality of high and low frequency words, resulting in the word frequency effect to decrease in this age group compared to younger children. Around the age of twelve, however, high frequency words increase in quality compared to low frequency words, resulting in an increased word frequency effect around this age.

With regard to the neighbourhood size effect, it seems that only in adulthood lexical representations have become of sufficient quality to show a neighbourhood size effects. Castles (1999) only found neighbourhood size effects in an analysis including both children (aged eight to twelve) and adults, but not when analysing the children data only. This may indicate that at only at a later age lexical representation of words with many orthographic neighbours are have become of high quality, thereby facilitating word recognition. In younger children, lexical representations of words with many or few orthographic neighbours may not yet differ enough in quality to elicit facilitation effects.

Lateralization. One study investigated lateralization of word recognition processes in children aged seven and twelve (Waldie & Mosley, 2000). In literate adults, language processes are left-lateralized and the authors aimed to investigate how this lateralization occurs as children grow older. In the study, children of seven and twelve years old saw words in their left or right visual field. They had to press a button with their left of right hand if the words was a real word, and had to refrain from responding if the word was a nonword.

The authors report increased lateralization with age: older children showed evidence for left hemispheric lateralization, whereas younger children did not. In older children, words presented in the left visual field (processed by the right hemisphere) were recognized faster with the right hand (controlled by left hemisphere) than left hand (controlled by the right hemisphere). This indicates that in order for the word to be processed, transcallosal relay from the right to the left hemisphere was required. For younger children, however, words presented in the left visual field (processed by the right hemisphere) were recognized faster with the left hand (controlled by the right hemisphere) than the right hand (controlled by the left hemisphere). This indicates that transcallosal relay was not required and that both left and right hemisphere contributed to the word recognition. The results from this study show that left-lateralization for language processes such as written word recognition occurs as children age. The authors suggest that the lateralization reflects a change in reading strategy from quick pattern recognition during perceptual processing, which is thought to be right-hemispheric, to left-hemispheric more automatized decoding based on grapheme-phoneme correspondences (Zaidel & Schweiger, 1984).

5.4 Discussion

Summary of results

In this study, we reviewed the literature that investigates the relationship between literacy experience and word recognition to understand how literacy experience influences word recognition efficiency. Specifically, this relationship was explored for different types of representation (semantic, phonological, orthographical, syntactic) and different levels of representation (lexical, syllable/morphological, sublexical). Studies examining spoken word recognition were discussed separately from studies discussing written word recognition. This allowed for the interpretation of results within modality-specific theories and frameworks of word recognition, and and for the investigation of differences and overlap between written and spoken language processing. Furthermore, studies that investigated word recognition in adults were discussed separately from studies that looked at word recognition in children of different ages or sampled from both the children and adult population. Even though it is not possible to directly compare groups that vary in their cognitive maturation, reviewing the literature carried out in both samples illuminates our understanding of developmental trajectories induced by literacy acquisition. In total 49 relevant articles were identified.

Overall, the evidence points to a substantial facilitative effect of literacy experience on word recognition. For all types and levels of representations, and in both the spoken and written modality, individuals with higher literacy experience outperformed individuals with lower literacy experience. We will now summarize the evidence for the influence of literacy experience on word recognition for the different types and levels of representations. Within each type and level of representation notable findings regarding modality or developmental trajectories are summarized. Then, global findings regarding the modality in which word recognition takes place and evidence for developmental changes related to literacy acquisition in word recognition are discussed. Limitations with regard to the generalizability of our findings are presented, and finally we provide recommendations and suggestions for future research.

Types of representations and their binding

Semantic representations. Literacy experience improves the quality of semantic representations and thereby increases both written and spoken recognition speed and accuracy. We also found evidence for a developmental trajectory of written word recognition. Initially, children's written word recognition seems to rely more on orthographic representations, but, with age, word recognition started to depend more on semantic representations instead. Access to semantic representations during written word recognition was faster in older children and adults and slower in young children.

Phonological representations. Phonological processing was more important for word recognition in adults with less literacy experience compared to adults with high literacy experience, and this pattern was reported for both the written and spoken modality. There was also evidence for developmental changes in that older children and adults relied less on phonological processes than younger children. ERP evidence also indicated that in young children (aged six, both pre-reading and beginning readers), phonological representations during spoken word recognition tended to remain active for a longer period of time (400 – 1000 ms after stimulus presentation). In older children (age seven onwards) and adults, however, this late ERP component related to phonological processing was not observed.

Orthographic representations. In contrast to the findings obtained for phonological processing, the reviewed literature suggested an *increased* reliance on orthographic representations with increased literacy experience, at least during written word recognition. For spoken word recognition, the results are less clear: there is no literature that investigated this particular relationship in adults, and the evidence for a developmental trajectory related to the acquisition of literacy is inconsistent. For written word recognition, clear evidence for developmental trajectories were observed. In particular, research in the children population provided more insight into which aspects of an orthographic representation undergoes changes as children become more literate. Letter identities within an orthographic representation were found to become more precisely and completely specified as children grew older. At the same time, the evidence suggested that the specific position of a letter within an orthographic representation does not undergo a process of specification, as written word recognition of older children and adults was still efficient even if the letters were in the wrong position.

An interesting finding from the adult literature was that people with lower literacy skills may be able to use their phonological skills to compensate for their suboptimal orthographic processing skills. This enables them to recognize written words efficiently, despite lower quality orthographic representations. However, when these individuals could not compensate for their orthographic processing difficulties due to phonological information being unreliable, their written word recognition was found to be impaired. For example, identifying a written nonword ("brane") that sounds like a real word ("brain") proved to be very difficult for low literacy experienced individuals.

Syntactic representations. Within the boundaries of this systematic review, no literature was identified that investigated the effect of literacy experience on syntactic representations during written or spoken word recognition.

Binding between types of representations. Bindings between semantic and orthographic representations, and semantic and phonological representations, seemed to increase in quality as individuals become more experienced with literacy. Regarding the binding between orthographic and phonological representations, an intriguing developmental trajectory was found. In adults, evidence suggested that the binding between these representational types increased in quality as a result of literacy experience, but the connection between orthographic and phonological representations seemed to be even stronger in younger children. This finding may be a reflection of younger children relying more on a decoding strategy than a whole-word recognition strategy (Ehri, 1995), whereas this is the other way around in older individuals and adults. Such dependence on a decoding strategy would allow the bindings between orthographical and phonological representations to be easily excited due to activation thresholds being low. More holistic word recognition strategies, in comparison to decomposition strategies, do not require these easily excitable connections between orthographical and phonological representations. What is striking is that this explanation is derived from theoretical frameworks of reading acquisition (Ehri, 1995), but that the evidence for connections between orthographic and phonological representations to be more easily excited in young children than in adults was obtained in the spoken modality. Thus, it seems that literacy acquisition has such a strong influence on language processing strategies that evidence for a shift in processing strategy can even be observed in the other modality.

Levels of representation

Lexical level. At the word level, evidence suggests that literacy experience improved recognition of whole words in adults. Moreover, whole-word recognition was improved in adults compared to children, suggesting a developmental trajectory induced by age and/or the acquisition of literacy. Our review also provided insights in how the effect of several word-level characteristics on word recognition was influenced by literacy experience or undergoes developmental changes.

A developmental effect was found on the word length effect in the written modality. This effect refers to shorter words generally being recognized faster than longer words. The word length effect was reported to be smaller in older children and adults compared to young children. This may reflect the shift from a decoding word recognition strategy to a holistic whole-word recognition strategy that has been suggested in frameworks of reading acquisition (Ehri, 1995).

By processing the word as a whole, rather than decomposing its constituents, processing duration of longer words speeds up, and speed differences between processing long and short words are reduced in size.

The word frequency effect and word familiarity effect (i.e., high frequent/ more familiar words are recognized faster) in the written modality was reported to be smaller in individuals who have more literacy experience. This seems to indicate that, as a result of increased exposure to written materials, lexical representations of low frequent or unfamiliar words improve in quality, thereby facilitating their recognition. Interestingly, the studies with children suggested a U-shaped developmental trajectory of the word frequency effect in the written modality. Studies reported a word frequency effect in young children (aged seven) that decreased in size in the upcoming years, but with a subsequent further increase in children aged twelve. This may indicate that the quality of lexical representations of frequent vs. infrequent words is subject to change during the process of literacy acquisition. Young children may show benefits in recognizing high frequency words, because they do not know the low frequency words yet. As they age, they learn the low frequency words and are able to recognize them, but given the decreased size of the word frequency effect, it seems that representations of high and low frequency words do not differ much in quality. At a later age, the word frequency effect increases again, possibly because representations of high and low frequency words start to differ more in their quality.

With regard to the neighbourhood size effect (i.e., words with many orthographic neighbours are recognized faster than words with few neighbours), we observed that, in the written modality, the effect decreased as adults became more literacy experienced. This suggested that literacy experience specifies lexical representations to such an extent that words with few neighbours are accessed as efficiently as words with many neighbours, even though a word with few neighbours does not have the facilitation benefit that comes with being in a dense neighbourhood. The neighbourhood size effect was not observed in children (aged eight to twelve), suggesting that the lexical representations of words with few neighbours only become of high quality in adulthood, or that the organization of the mental lexicon is subjected to change during literacy development.

Morphological/syllable level. Several studies manipulated word characteristics on a level between the lexical and sublexical level. This gives insights into how literacy experience influences processing of word parts (morphemes, syllables). We could not identify any studies that were administered in the adult population only. The following results were thus obtained in studies comparing children of different ages, or children and adults.

With regard to lexical stress, evidence was somewhat inconclusive, with some studies finding no evidence of lexical stress processing in spoken word recognition in adults or children. The null-results obtained in some studies may be due to their task design, which was too difficult or included a manipulation that was too subtle. Studies with less subtle manipulations of lexical stress did report that adults and older children process lexical stress to a larger extent than younger children. This may indicate that, in younger children, the lexical representations that include information about stress may not yet be fully developed or at least does not influence spoken word recognition.

Also with regard to morphemes, it was reported that older children and adults show more sophisticated processing during written word recognition. More precisely, a shift was described from morpho-semantic processing to morpho-orthographic processing, which was reported to occur around the age of ten. These effects, however, depended on the language in which the study was administered. In languages with a simple syllabic structure and rich morphology, such as Spanish or Italian, young children already made sophisticated used of their morphological knowledge during word recognition and already seemed to be able to process morphemic information morpho-orthographically at a young age.

Finally, for syllable processing, evidence indicated that processing was more sophisticated in older children and adults. Again, this effect was influenced by language, as such that developmental effects were only found in languages with complex syllable structures, whereas children speaking languages with simple syllable structures tended to show sophisticated syllable processing already at a younger age.

Sublexical level. Particularly studies that examined orthographic and phonological representations used manipulations on the sublexical level, by changing graphemes or phonemes. The evidence indicated that written and spoken word recognition on the sublexical level was influenced by literacy experience. Orthographical effects increased with literacy experience, whereas phonological effect decreased. This pattern was also found in studies with children, suggesting a developmental pattern. Younger children were more strongly influenced by manipulations of phonemes, whereas word recognition in older children and adults was affected more strongly by manipulations of graphemes. Moreover, evidence was reported in the written modality that the precise identity of sublexical features such as graphemes is subjected to developmental changes, as in older children and adults, but not younger children, the letter identities within an orthographical representation were fully defined.

Spoken and written word recognition

In this review, we examined the effect of literacy experience on word recognition separately for the written and spoken modality, because written and spoken word recognition are described by modality-specific theoretical models and frameworks, which makes it difficult to directly compare studies that examine word recognition in different modalities. Due to the demands of each modality, processing information in these modalities differs and therefore, literacy experience may affect processing in the spoken and written modality differently. Our review, however, showed that the effects of literacy on the types and levels of representations were very similar across modalities. No stark differences were observed in the pattern of findings between the modalities. This finding has several important theoretical implications. First, it provides support for theories that hypothesise interactivity within the mental lexicon between modality-specific representations. Models of written word recognition, in particular, assume that orthographical and phonological representations are closely connected and that activation spreads between the types of representations even if information is presented in only one modality (Grainger & Ferrand, 1994; McClelland & Rumelhart, 1981; Seidenberg & McClelland, 1989).

A second theoretical implication is that literacy experience does not only affect skills related to the written modality, such as reading fluency, reading comprehension, or spelling (Mol & Bus, 2011) or 'amodal' language skills, such as vocabulary size or declarative knowledge (Cunningham & Stanovich, 1991; Stanovich, West, & Harrison, 1995), but also skills related to processing information in the spoken modality. A mechanism that may account for this cross-modal influence may be the aforementioned connections between orthographic and phonological representations that are activated during the word recognition stage where lexical representations are accessed. Through these connections, which were reported to strengthen as a result of literacy experience, the quality of phonological representations may improve when encountering input in the written modality. ERP studies included in the present review, indeed provide some evidence that literacy experience facilitated word recognition at the stage of lexical access, but not at other stages that are more specific to processing information in a certain modality, such as the early stages where acoustic or visual information is initially processed. Thus, it seems that the connectivity between representations in the mental lexicon does not only allow literacy experience to influence written word recognition, but also transfers its facilitatory effects to word recognition in the spoken modality.

Developmental trajectories

One limitation of the present review, and the children's literature in general, is that we cannot distinguish the effects of literacy experience or literacy acquisition on word recognition from effects of cognitive maturation, (spoken) language experience, or increasing world knowledge, due to the inherent differences between adults and children, and children of different ages with regard to these factors. However, by including studies that examined word recognition in children of different ages, and studies comparing children and adults, we were able to shed light on some developmental changes in word recognition for different types of representations, different levels of representations and different modalities. Several developmental trajectories were summarized in the previous parts of this discussion. One key developmental change, reported by many studies in this review, is the occurrence of a shift in the processing route of written word recognition. Dual-route approaches to silent written word recognition assume that there are two routes towards lexical access during silent written word recognition: a direct route through the orthography representation of the printed word, and an indirect route through the phonological representations associated with the printed word (Grainger & Ziegler, 2011). Evidence from our review suggested that, as children get older, they shifted from the indirect, phonological route to a direct, orthographic route, as evidenced by the fact that phonological effects were found to reduce with age, whereas orthographic effects on written word recognition increased. Moreover, older children's word recognition seems to rely more on whole-word processing strategies rather than decomposition processing strategies. These findings are in line with theoretical frameworks on the development of sight word reading. Beginning readers tend to utilize a decomposition strategy in which they serially decode graphemes to phonemes. Only at a later stage in their reading acquisition, they are able to access the orthographic representation of the full word at once (Ehri, 1995).

Although the large differences in cognitive maturation between children of different ages, and between adults and children, hinder direct comparisons between these populations, we found patterns of results that were very similar across different populations. First, there was major evidence that regardless of the type of representations, levels of representations or presentation modality, word recognition was facilitated in older children compared to younger children. Moreover, studies comparing younger children, older children and adults reported that older children tend to outperform younger children, and that adults tend to outperform children. This indicates that word recognition is subjected to a long developmental trajectory which is present when children start to learn to read and continues into the teenage years into adulthood.

Limitations

For the present review, we limited our search with respect to several aspects, which must be kept in mind when interpreting the finding. For example, we focused on alphabetic languages and only included studies that sampled from neurotypical populations with no history of impairments in the domains of general development, language, speech or reading. Most importantly, we focused on word-level processing. Only studies in which participants performed a silent word recognition task where they were required to fully activate a word-level lexical entry were included. Studies that used tasks focused on sublexical language units, for example phoneme or letter identification tasks, were excluded. In order to engage in these tasks, activation of the full lexical representation is not required. Instead, the tasks can be performed by only activating sublexical units in the mental lexicon prior to word-level lexical access (Foss & Blank, 1980; Newman & Dell, 1978; Segui et al., 1981). This does not mean that there is no input from the lexical level during these tasks, as, according to theories of word recognition, activation between the different levels of representations is thought to flow freely (Grainger & Ferrand, 1994; McClelland & Elman, 1986; McClelland & Rumelhart, 1981). However, performance on these tasks is thought to rely more strongly on sublexical than lexical representations, especially if there is little predictive context, which is often the case in phoneme or letter monitoring tasks (cf. dual code hypothesis, Foss & Blank, 1980). A review on literacy experience and word processing including tasks measured at a sublexical level would help us understand whether literacy experience also influences sublexical language processing and how it influences the connections between the different levels of representation.

We also excluded studies that used tasks with larger language units, such as sentences and texts. Studies using these larger language units often examine more holistic comprehension processes, inference making processes or prediction processes. Our review does not cover the effect of literacy experience on these processes and mechanisms. It may be worthwhile to examine the effect of literacy experience on processing language units larger than words as there is evidence that suggests a relationship between literacy experience and prediction of upcoming speech (Huettig & Pickering, 2019) and literacy experience and text comprehension (Mol & Bus, 2011).

The results with regard to type and levels of representations must be interpreted with two aspects in mind. First, for most levels, with the exception being the syllable/morphological level, the authors of the studies often did not precisely specify which type or level of representation they attempted to examine. We therefore often inferred this from the information (research questions, hypotheses, study design, tasks) available in the text. Thus, some allocations are subjective in nature. Second, many theories of word recognition assume that word processing does not occur at a single level or representational type, but that information flows freely through connections between different types and levels of representations (Grainger & Ferrand, 1994; McClelland & Elman, 1986; McClelland & Rumelhart, 1981). With clever manipulations, studies aimed to look at the effect of literacy experience on word recognition for one particular type or at a particular level of representation, but that does not mean that we can interpret such an effect as only taking place for that particular type of representation or at that particular level, as it likely also asserts its influence to other types and levels of representations as well.

One must keep in mind that this systematic review covers studies with a publication year ranging from 1978 to 2021. Within this time period, statistical treatment of data underwent changes. Most studies used an Analysis of Variance approach (ANOVA), whereas only some of the newer studies analysed the data with linear mixed effect models (LME). The latter may be particularly suitable analysis strategy for studies examining the effect of literacy experience on word recognition, as variation between individuals due to other factors can be accounted for using random effects. Also, in linguistic tasks, where there are often multiple trials per condition, LME does not require observations to be averaged within condition, thereby reducing information loss compared to analysing the data with an ANOVA. Another methodological issue that often arises in older studies, is the use of very small samples. The studies reviewed here did not have very small sample sizes, with an average sample size of 122 participants ($SD = 89.61$, range = 38 – 627). Unfortunately, because studies did not report power analyses, we do not know whether the sample sizes were sufficient. Despite these issues with regard to methodology and statistics, the studies produced conclusive findings and stable patterns with regard to the influence of literacy experience on word recognition. Unfortunately, studies often did not report effect sizes, which leaves us unable to compare the size of the effect of literacy experience on word recognition across studies. As such, a meta-analysis, where effect sizes are calculated with data available from the full-texts of the article, might be worthwhile to pursue. It is, however, unlikely that such meta-analysis would produce patterns of observations and ultimate conclusions that are very different from those obtained in this present systematic review, but a meta-analysis may provide insights into the size of the effect of literacy experience on written and spoken word recognition.

Recommendations & future directions

This review identified points of attention that should be taken into account when venturing into the relationship between literacy experience and word recognition. Most importantly, the studies of the children's population that were included in this review directly compared response times between adults and children of different ages. Adults compared to children are generally faster responders (Hale, 1990) and this speed difference should be controlled for by, for example, standardizing reaction times (RTs) (Faust, Balota, Spieler, & Ferraro, 1999). Only two studies included in the review compared children and adults accounted for the relative speed benefit (D. Hansen et al., 2012; Ventura et al., 2007). Such practice allows for determining whether the reported effects are due to the variable of interest, for example word frequency, or relative processing speed differences between the two groups. We recommend that researchers that aim to compare response speed of adults and children consider whether correcting for differences in processing speed is required, depending on their question of interest and study design.

Our review also established new directions for future research. First, and perhaps one of the most striking findings of this review, the research on the effect of literacy experience on spoken word recognition is very limited. In the adult population, we identified only one study that examined spoken word recognition, versus eleven studies in the written modality. In the child population, more studies in the spoken modality were identified (six), but there was still a large contrast compared to the number of studies in the written modality (thirty-two studies). This suggests that there is an imbalance in the literature favouring written word recognition over spoken word recognition. In particular, our results indicate that there is a relationship between literacy experience and word recognition in the spoken modality that suggests that experience with written materials also impacts processing spoken information. However, due to the lack of studies, we do not know which aspects in the spoken word recognition process are affected by literacy experience, and how this relationship forms as individuals gain experience with the written modality. More data on this relationship, both in the adult and children population, will provide a better understanding of how our language system processes, organizes and operates on modality-specific input.

Second, there is relatively little literature on the influence of literacy experience on syntactic representations and syntactic processing during word recognition. Most research was focused on semantic, orthographic or phonological representations, or the binding between these three representational types. There is evidence that grammatical knowledge varies as a function of literacy experience (Favier & Huettig, 2021) but these studies are most often administered within a sentence context, rather than on the word level. Thus, investigating the effect of literacy experience on syntactic representations or syntactic processing during word recognition may be a new avenue of research that can inform us of how literacy experiences shapes syntactic representations and syntactic processing.

Third, the current literature can only provide a glimpse of understanding as to how literacy experience affects the word recognition at different stages of the written and spoken word recognition process. The literature suggests that literacy experiences influences word recognition at the stage where lexical representations are accessed, but that this influence may not be present during initial visual or acoustic processing. More research is required to investigate the influence of literacy experience on these initial stages of processing to get a better understanding of these initial observations. It is also worth establishing at which stages at the lexical level literacy experience facilitates processing. For example, it seems to be the case that individuals with high literacy experience have a more automated flow of information between different types of representations during the stage of lexical access but we do not know whether literacy experience also influences other lexical stages, such as the lexical competition stage. Adopting neuroscientific methodologies such as ERP could provide a clearer image of the mechanisms that allow literacy experience to facilitate word recognition during the process of recognition.

Fourth, and finally, this review showed that literacy experience was related to facilitated word recognition in adults and that word recognition undergoes developmental changes related to literacy acquisition. However, the current literature cannot provide insights with respect to whether variation in literacy experience in children of the same age is reflected in differences in their word recognition efficiency. The word recognition studies in the child population most often did not administer measures of literacy experience. However, since already at a young age children show differences with respect to how much they read (Juel, 1988), it is possible that already at a young age variation in a group of children's literacy experience is reflected in differences in their word recognition efficiency. Furthermore, research has shown that differences in literacy experience are stable over time: that is, the children who read a lot when they are young tend also read a lot when they grow older (Wigfield & Guthrie, 1997). Thus, it may be the case that the effects of literacy experience on word recognition that are observed in adults already start from the moment a child learns to read, and remain stable as the child grows into adulthood. This would mean that the internal and external factors that result in the groups of literacy experienced and less experienced individuals we observe in the adult population may start to assert their influence from the very first moments of reading acquisition. Uncovering these factors could help determine how and when interventions that promote literacy can be most impactful.

For such research to commence, however, it is necessary to develop tools with which literacy experience can be measured in children across their development. As also shown in this review, the Author Recognition Task is the most widely used measure to assess literacy experience. Due to its objective nature, it is thought to be a valid and reliable measure (Acheson et al., 2008; Dąbrowska, 2018; James et al., 2018; Mar & Rain, 2015; Payne et al., 2012; Stanovich & West, 1989). The task might be confounded though by factors such as whether readers pay attention to authors during reading, or the fact that selected authors are authors from popular genres only. However, unlike more subjective measures, the Author Recognition Task is not influenced by social desirability, as are questionnaires about reading behavior, and neither is it influenced by personal biases such as the over- or underestimation of reading time estimates. Even objective measures of literacy experience, such as reading times or number of pages read, are susceptible to distortions, as they rely on the type and complexity of the material read and the personal speed with which one reads. Thus, despite some shortcomings, the Author Recognition Test is a very useful tool to approximate literacy experience.

For children, several objective measures of literacy experience have been developed. Examples are the Title Recognition Test (Cunningham & Stanovich, 1990) tested in children aged nine to ten, the Comic Recognition Test (Allen, Cipielewski, & Stanovich, 1992) tested in children aged eleven, and the UK Children's Author Recognition Test (Stainthorp, 1997) tested in children aged nine to eleven. The Adult Author Recognition Test has also been administered to children aged nine (Cunningham & Stanovich, 1990). These measures are, however, not widely used and their relative validity or reliability across children's development has not been researched much. Questions as to what type of recognition test (author, title) or what type of genre (books, magazine, comics) are most suitable for children populations are not yet answered. Moreover, the age range for which these tests are suitable is not established. In addition, the tests are currently only available in English, and, to our knowledge, objective measures of literacy experience suitable for children are not yet developed in other languages. Thus, it would be worthwhile to research psychometric aspects of (newly developed) tests measuring literacy experience in children in tandem with further investigations of developmental trajectories of literacy experience on language processing.

5.5 Appendix I - Study descriptives

The table below provides a detailed overview of all studies included in this review. To structure the review, the authors categorized the studies by themes and subthemes. For each study, the table describes the language in which the study was conducted, the number and age range of the participants, the task with which literacy experience was assessed, and a short description of the word recognition task.

Theme	Subtheme	Author	Type	Language	N	Age	Literacy experience measure	Word recognition measure
Age > 18, Spoken WR	General literacy experience	Kosmidis et al. (2006)	journal article	Greek	58	adults M = 67.91)	literacy + education level (low-literate + low education, literate + low education, literate + high education)	lexical decision
Age > 18, Written WR	General literacy experience	Kennedy (1996)	dissertation chapter (Chapter 3)	English	45/40	adults	ART	lexical decision
Age > 18, Written WR	General literacy experience	Michael (2008)	dissertation chapter (Chapter 2 (experiment 1))	English	140	adult	ART	lexical decision manipulating word placement in left or right visual field
Age 18, Written WR	General literacy experience	Thompson (2011)	dissertation chapter (Chapter 5)	English	122	adults	ART	lexical decision with manipulation of imageability, homophony, regularity in words and pseudohomophony, bigram frequency, orthographic neighbourhood in nonwords
Age > 18, Written WR	orthography	Welcome and Trammel (2019)	journal article	English	38	adults	ART	lexical decision with orthographic priming. Primes were anagrams (pronounceable (jonis-joins) or unpronounceable (cdoes-codes)) or unrelated target string (pronounceable (bulid-doubt) or unpronounceable (mmkaes-north))

Theme	Subtheme	Author	Type	Language	N	Age	Literacy experience measure	Word recognition measure
Age > 18, Written WR	phonology (homographs/ regularity/ homophones)	S. J. Unsworth and Pexman (2003)	journal article	English	49/50	adults	ART (+ vocabulary and reading comprehension)	lexical decision tasks. Experiment 1: normal LDT. Experiment 2: phonological LDT (does it sound like a word?). In both manipulations of whether targets sounded like other real words (homophones "reel" vs. nonhomophones "seal"), looked like other real words (homographs "tear" vs. nonhomographs "clear") and whether the targets were exception ("deaf") or regular words ("beam")
Age > 18, Written WR	semantic	D. Hansen et al. (2012)	journal article	English	92/72	adults	ART	Experiment 1: semantic categorization task ("is this word easily imaginable?"). Experiment 2: lexical decision (phonological: does this sound like a word?). Both manipulated BOI of the words (= body-object interaction, which represents perceptions as to whether a human body can physically interact with the object)
Age > 18, Written WR	word familiarity (prevalence/frequency)	Lee et al. (2019)	journal article	Korean	104	adults	ART	lexical decision

Theme	Subtheme	Author	Type	Language	N	Age	Literacy experience measure	Word recognition measure
Age > 18, Written WR	word frequency	Cohen-Shikora and Balota (2016)	journal article	English	148	adults	age (18-86 years)	lexical decision (manipulating word frequency), animacy judgement task
Age > 18, Written WR	word frequency, orthography (lexical decision task) and phonology (form priming lexical decision)	Chateau and Jared (2000)	journal article	English	64	adults	ART	Task 1: lexical decision: manipulating word frequency and neighbourhood density, using pseudohomophones as nonwords, manipulating orthographic typicality. Task 2: lexical decision form priming, with primes that were orthographically similar but phonologically dissimilar (touch-couch) or unrelated primes (shall-couch), manipulating prime duration (30 or 60 ms) and prime frequency
Age > 18, Written WR	word frequency/ neighbourhood size (experiment 1), orthography (experiment 2)	Sears et al. (2008)	journal article	English	120/120	adult	ART	lexical decision manipulating word frequency/ neighbourhood size. Experiment 1: normal nonwords, manipulating neighbourhood size. Experiment 2: nonwords were pseudohomophones, manipulating orthographic typicality

Theme	Subtheme	Author	Type	Lan-guage	N	Age	Literacy experience measure	Word recognition measure
Age > 18, Written WR	word frequency/ neighbourhood size (experiment 2), phonology (experiment 3)	Lewellen et al. (1993)	journal article	English	70/70	adults	ART, MRT, language experience questionnaire	Experiment 2: lexical decision manipulating word frequency/ neighbourhood density. Experiment 3: semantic categorization (things to eat "peach", body parts, animals) with homophone foils ("pair") + controls ("pier") and semantic foils ("stove") + control ("stone")
Age < 18, Spoken WR	orthography	Pattamadilok et al. (2009)	journal article	French	90	adults + grade 2, 3, 4	age	lexical decision orthographic rime consistency manipulation: consistent rime (rime has no orthographic competitor , e.g., -age can only be spelled as age) or inconsistent rime (rime has orthographic competitor, e.g., -ac can be spelled with ac or aque)
Age < 18, Spoken WR	orthography	Schild et al. (2011)	journal article	German	51	kinder-garten children (pre-readers and readers) and children grade 2	age	lexical decision go/no go primed with word fragment: identity condition (mon-monster), variation condition (non -monster) and control condition (dack-monster)

Theme	Subtheme	Author	Type	Language	N	Age	Literacy experience measure	Word recognition measure
Age < 18, Spoken WR	orthography	Ventura et al. (2007)	journal article	Portuguese	90	adults, Kindergarten children (prereaders), children grade 2, 3, 4	age	lexical decision orthographic rime consistency manipulation: consistent rime (rime has no orthographic competitor , e.g., /um/ can only be spelled as ume) or inconsistent rime (rime has orthographic competitor, e.g., /mɛl/ can be spelled with ele or el)
Age < 18, Spoken WR	stress, phonology	Lin et al. (2018)	journal article	English	88	adults + children age 6,8,10	age	lexical decision with stress (real word made nonword by changing stress, e.g., cabin-cabin) and phoneme manipulation (real word made nonword changing 1 letter, e.g., cabin-calin)
Age < 18, Spoken WR	stress, phonology	Schild et al. (2014)	journal article	German	69	adults, Kindergarten children, children age 7	age	lexical decision primed manipulating primes' stress overlap (match/mismatch with target) and phoneme overlap (phonological onset of targets/primes matched/mismatched)
Age < 18, Spoken WR AND Age 18, Written WR	semantics	Gijsel et al. (2011)	journal article	Dutch	141	children grade 1-6	age	semantic categorization task with auditory or written presented words. Semantic task was either horizontal (Exemplar-level task) or vertical (Superordinate-level task)

Theme	Subtheme	Author	Type	Language	N	Age	Literacy experience measure	Word recognition measure
Age < 18, Written WR	lateralization	Waldie and Mosley (2000)	journal article	English	55	children aged 7, 12 (right-handed)	age	lexical decision go/no go manipulating word placement in left or right visual field and hand (left/right) to press button. Nonwords were either pronounceable (orthographically regular "deks") or unpronounceable (orthographical irregular "gaot")
Age < 18, Written WR	letter position, word frequency	Ducrot et al. (2013)	journal article	French	107	children grade 1-5	age	lexical decision, manipulating fixation place in word (experiment 1: left vs right parafoveal presentation, experiment 2: fixation on one of the five letter slots) and word frequency
Age < 18, Written WR	morphology	Dawson et al. (2018)	journal article	English	154	adults + children aged 7-9, 12-13 and 16-17	age	lexical decision manipulating and morphology of nonwords (morphological nonword "earist" or true nonword "earlit")
Age < 18, Written WR	morphology	Lázaro et al. (2017)	journal article	Spanish	90	children grade 2, 4, 6	age	lexical decision go/no go manipulating suffix frequency
Age < 18, Written WR	morphology	Perdijk et al. (2012)	journal article	Dutch	118	children grade 2, 4	age	lexical decision manipulating morphological family size

Theme	Subtheme	Author	Type	Language	N	Age	Literacy experience measure	Word recognition measure
Age < 18, Written WR	morphology	Quémart et al. (2012)	journal article	French	60	children grade 3, 5	age	lexical decision with words and nonwords manipulated regarding presence of base and suffix (B+S+ "pêcheur", B-S+ "janvier', B+S- "barque', B-S- "brousse")
Age < 18, Written WR	morphology, compound words	Hasenäcker et al. (2017a)	journal article	German	627	adults + children grade 2-6	age	lexical decision with monomorphemic (Laterne), compound (Segelboot) and derivation (Lehrer) words
Age < 18, Written WR	morphology, orthography	Beyersmann et al. (2012)	journal article	English	134	adults (experiment 1) + children grade 3, 5 (experiment 2)	age	lexical decision (primed, 60 ms) manipulating prime (morphological related "golden-gold", pseudoderivation "mother-moth", orthographically related "spinach-spin")
Age < 18, Written WR	morphology, orthography	Beyersmann et al. (2015)	journal article	French	191	children grade 2-5	age	lexical decision (primed, 50 ms) manipulating prime (morphological related "tristesse-triste", morphological nonwords "tristerie-triste", nonsuffixed orthographically related nonwords "tristald-triste" and unrelated nonwords "direction-triste")

Theme	Subtheme	Author	Type	Language	N	Age	Literacy experience measure	Word recognition measure
Age < 18, Written WR	morphology, orthography	Hasenäcker et al. (2020)	journal article	German	98	children grade 2-4 (longitudinally)	age	lexical decision (primed, 50 ms) manipulating prime (morphological related "kleidchen-kleid", morphological nonwords "kleitum-kleid", nonsuffixed orthographically related nonwords "kleidekt-kleid" and unrelated nonwords "träumerei-kleid")
Age < 18, Written WR	morphology, orthography	Schiff et al. (2012)	journal article	Hebrew	80	children grade 4, 7	age	lexical decision (primed, 57 ms) manipulating prime (identity, morphologically and semantically related, morphologically related but semantically unrelated (pseudoderivation), orthographically (and phonologically) related but morphologically unrelated)
Age < 18, Written WR	morphology, semantic, orthography	Fleischhauer et al. (2021)	journal article	German	254	adults + children grade 1-4	age	lexical decision (primed, 67 ms) manipulating morphological overlap of primes (morpho-semantic "Leser-lesen", morpho-orthographic "Messer-messen", orthographic "Nagel-nagen", semantic overlap "Zeitung-lesen", or no relation ("Ärger-lesen")

Theme	Subtheme	Author	Type	Language	N	Age	Literacy experience measure	Word recognition measure
Age < 18, Written WR	morphology, semantic, orthography	Quémart et al. (2011)	journal article	French	220	(experiment 1 + 2) adults + (Experiment 1, 2, 3) children grade 3, 5, 7	age	lexical decision (primed, 60 ms) manipulating prime (morphological related "tablette-table", pseudoderivation "baguette-bague", orthographically related "abricot-abri", semantically related "tulipe-fleur") and prime duration (experiment 1: 60, experiment 2: 250, experiment 3: 800 ms)
Age < 18, Written WR	morphology, semantics, compound words	Beyersmann et al. (2019)	journal article	English	162	adults + children grade 3, 5, 6-12	age	lexical decision (primed, 50 ms) manipulating compound transparency (transparent "farmhouse-farm", opaque "butterfly-butter" or non-compound "sandwich-sand") and semantic relatedness of prime/target
Age < 18, Written WR	morphology, semantics, orthography	Quémart et al. (2018)	journal article	English	108	children grade 3, 5	age	lexical decision (primed, auditory) manipulating prime (morphological and semantically related (low "belly-bell", moderate "lately-late", and high similarity relations ’boldly-bold", semantically related "garbage-trash", orthographically related "spinach-spin")

Theme	Subtheme	Author	Type	Language	N	Age	Literacy experience measure	Word recognition measure
Age < 18, Written WR	morphology, semantics, orthography, phonology	McCutchen et al. (2009)	journal article	English	163	children grade 5, 8	age	lexical decision (primed, continuous) manipulating prime type (morphological (phonological overlap "planner-plan" or not "dept-deep)", semantic "strategy-plan", orthographical (phonological overlap "farmer-far" or not "depend-deep"), unrelated "salad-plan")
Age < 18, Written WR	morphology, syllable	Hasenäcker and Schroeder (2017b)	journal article	German	81	adults + children grade 2, 4	age	lexical decision manipulating number of morphemes (monomorphemic "Fahrer" or multimorphemic "Spinat") and syllable boundary (congruent "Fah:rer"/"Spi:nat" or incongruent "Fahr:er"/"Spin:at")
Age < 18, Written WR	morphology, word frequency	Burani et al. (2002)	journal article	Italian	90	children grade 3, 4, 5 (experiment 2)	age	lexical decision manipulating frequency of words, word length, and morphology of nonwords (morphological nonword or true nonword)
Age < 18, Written WR	morphology, word frequency	D'Alessio et al. (2019)	journal article	Spanish	90	children grade 2, 4, 6	age	lexical decision manipulating morphological complexity (simple and suffixed words) and word frequency (high and low)

Theme	Subtheme	Author	Type	Language	N	Age	Literacy experience measure	Word recognition measure
Age < 18, Written WR	neighbourhood size, orthography	Castles (1999)	journal article	English	103	adults, children grade 2, 4, 6	age	lexical decision (masked priming, 57 ms) manipulating neighbourhood size of targets and orthographic relationship between prime and target (identity: "ball-ball", form overlap: "dall-ball", control: "lift-ball"
Age < 18, Written WR	orthography	Castles et al. (2007)	journal article	English	47	adults + children grade 3, 5 (longitudinally)	age	lexical decision (masked priming, 57 ms) manipulated prime–target orthographic similarity: one letter different (rlay-PLAY), transposed letters (lpay-PLAY) and control primes with all different letters (meit-play)
Age < 18, Written WR	orthography	Colombo et al. (2017)	journal article	Italian	140	adults + children grade 2, 3, 5	age	lexical decision manipulating nonwords in length (short vs long) and orthographic similarity between words and nonwords: two-letter-transposed (codra-corda) or two-letter-different (dolba-corda), and for the long words, place of transposition (beginning "ablergo/acmergo (from albergo)" or end of word "leopadro/leopatso (from leopardo)"

Theme	Subtheme	Author	Type	Language	N	Age	Literacy experience measure	Word recognition measure
Age < 18, Written WR	orthography	Kezilas et al. (2017)	journal article	English	122	adults + children grade 2-6	age	lexical decision (primed, 50 ms) manipulating orthographic relationship between prime/target: transposed letter (litsen-listen), 2 transposed letters (lidfen-listen), all different (rodfup-listen) and identity prime
Age < 18, Written WR	orthography, phonology	Grainger et al. (2012)	journal article	French	163	adults + children grade 1-5	age	lexical decision with nonwords that were either pseudohomophones (trane) or orthographic control (trand) or had transposed letters (talbe) and their orthographic control (tarpe)
Age < 18, Written WR	orthography, phonology	Sprenger-Charolles et al. (2003)	journal article	French	60	children grade 1-4 (longitudinally)	age	semantic categorization task with correct words and pseudohomophones (rouje, oto) and visual foils (rouqe, outo) as foils
Age < 18, Written WR	orthography, phonology	Ziegler et al., 2014	journal article	French	284	children grade 1-5	age	lexical decision (primed, 70 ms) manipulating orthographic relationship between prime/target using pseudohomophones (vaze-vase) vs control (vare-vaze) and words with transposed letters (talbe-table) vs. control (tarfe-table)

Theme	Subtheme	Author	Type	Language	N	Age	Literacy experience measure	Word recognition measure
Age < 18, Written WR	orthography, semantic	Polse and Reilly (2015)	journal article	English	80	children grade 1-4	age	(Experiment 1): semantic matching task, with target (sick) and foils being manipulated in orthography (baseline lady [target], lady [match], water [foil]; orthographic condition lady [target], lady [match], lazy [foil], semantic condition lady [target], woman [match], water [foil], orthographic-semantic condition lady [target], woman [match], lazy [foil]
Age < 18, Written WR	orthography, word length	Acha and Perea (2008)	journal article	Spanish	119	adults + children grade 3, 6	age	lexical decision (primed, 50 ms) manipulating word length (short vs. long) and orthographic relationship prime/target: transposition (aminal-animal), 2-letter substitution (arisal-animal)

Theme	Subtheme	Author	Type	Language	N	Age	Literacy experience measure	Word recognition measure
Age < 18, Written WR	phonology	C. Davis et al. (1998)	journal article	English	112	adults (experiment 1 + 3) + children grade 4 (experiment 2 + 4)	age	lexical decision (masked priming, 57 ms) with identity "made-made", homophone "maid-made", orthographic control that shared the same degree of orthographic overlap as homophone prime "maud-made" or all letter different prime "flea-made" (experiment 3 + 4, go/no go in exp1+2).In experiment 1 + 2 they differed 2 or more letters from the target, and in experiment 3 + 4 the homophone and orthographic control either only differed 1 letter from the target "wosh-wash" or "wesh-wash", or 2 or more letters (like in experiment 1 + 2)
Age < 18, Written WR	syllable frequency, word frequency	Luque et al. (2020)	journal article	Spanish	80	children grade 2, 4	age	lexical decision manipulating word frequency (high-low) and frequency of initial syllable (high-low).
Age < 18, Written WR	word length	Samuels et al. (1978)	journal article	English	80	adults + children grade 2, 4, 6	age	semantic categorization (animals) manipulating word length (3, 5, 6 letter words)

6 | Modality effects in novel word learning[1]

Abstract

New words can be encountered in at least two modalities: spoken and written. Several theoretical frameworks produce contradicting predictions regarding in which modality learning is most efficient, and there is experimental evidence for both a spoken and written learning benefit. The present study investigated the modality effect in word learning using a new paradigm. Dutch adults were presented with pseudowords, either visually or auditorily, each paired with a picture of an unknown object. Following the training task, participants performed an old-new matching test in the written or spoken modality. In Experiment 1, no main effect of training modality was found, but an interaction indicated that the written training – spoken test condition yielded the most efficient learning. Experiment 2 was aimed at investigating modality effects as consolidation progresses, with test sessions a day and a week after the initial training. No modality effects were observed at any time point. In Experiment 3, the number of participants per condition was quadrupled allowing for the detection of small modality effects. There was no evidence for an effect. To conclude, it seems that highly literate adults are able to learn novel words equally efficiently in the spoken and written modality.

[1]Experiment 1 adapted from Wolf, M. C., Smith, A. C., Meyer, A. S., & Rowland, C. F. (2019). Modality effects in vocabulary acquisition. In A. Goel, C. Seifert, & C. Freksa (Eds.), *Proceedings of the 41st Annual Meeting of the Cognitive Science Society (CogSci 2019)* (pp. 1212-1218). Austin, TX: Cognitive Science Society.

6.1 Introduction

From the moment reading instruction commences around the age of six, humans are able to encounter and learn novel words not only while listening to speech, but also during reading. Gaining the ability of reading has an important consequence for the representational system of language. The reading skill allows for storing two representations of the same lexical item in the mental lexicon: a phonological and an orthographic representation. It is currently unclear whether the modality in which words are encountered — spoken or written — influences the efficiency with which these words are learned. Are the representations learned in one modality more precise and complete than representations obtained in the other modality? Is the activation of lexical entries initially learned in one modality more accurate than in the other modality? The present study uses a new word learning paradigm to answer the question how modality of input influences word learning efficiency. Results will provide in new insights regarding the mechanisms of written and spoken word learning. Moreover, it will improve our understanding of how orthographic, phonological and semantic aspects of words are stored, connected and retrieved in the mental lexicon.

In the adult word learning literature, studies directly comparing written and spoken word learning efficiency are scarce. Table 6.1 provides a detailed overview of the current literature. A study by Bakker et al. (2014) reported evidence for a spoken learning benefit. Lexical competition, an indicator of successful lexical integration of novel words, was stronger for words learned in the spoken modality, and recall and recognition of these words was better. Indeed, the statistical learning framework proposed by Frost, Armstrong, Siegelman, and Christiansen (2015), predicts a benefit when learning spoken words. A word's identity is determined by the sequence of its constituents (phonemes or graphemes). Shuffling the order of the constituents may result in an entirely different word with a different meaning. Thus, learning the correct sequence of phonemes is of utmost importance during novel word learning. According to Frost et al. (2015)'s framework, the serial nature of auditory information has made the auditory cortex sensitive to the sequential aspect of input, which benefits statistical learning of sequential information in the auditory modality (Conway & Christiansen, 2005; Robinson & Sloutsky, 2007; Saffran, 2002). In contrast, the spatial nature of visual information sensitized the visual cortex to spatial cues, creating an advantage for statistical learning of spatial information in the visual modality Conway, Pisoni, and Kronenberger (2009); Saffran (2002). Their framework thus predicts a spoken learning benefit, because a crucial part of novel word learning, namely learning the correct sequence of its constituents, is particularly optimized in the spoken modality.

On the other hand, written learning benefits are also described in the literature. Balass et al. (2010) reported higher accuracy and faster response latencies for words learned in the written modality. Moreover, using a primed lexical decision task, van der Ven et al. (2015) found larger priming effects for words learned in written modality, indicating that these words were better integrated in the mental lexicon. They found no evidence for a modality effect for recall and recognition. Theories of word reading (Coltheart et al., 2001) and reading acquisition (Ehri, 1995; Shankweiler, 1999; Share, 1995) indeed predict a word learning benefit in the written modality. These theories argue that upon a written encounter with a novel word, readers mentally recode the graphemes into phonemes, resulting in precise and complete orthographic and phonological representations of the novel word. The storage of two high quality representations allows for two routes towards efficient and accurate activation of the novel word upon the next encounter. Regarding the spoken modality, evidence suggests that crossmodal recoding results in less precise crossmodal representations. In particular, Johnston, McKague, and Pratt (2004) found large differences in the strength of the priming effect in English 5-letter known words and pseudowords learned in the spoken modality. For known words, using a prime that was the same as the target (identity prime) facilitated recognition with 71 ms compared to a prime with random letters. A prime that differed two or more letters from the target word had a much smaller facilitative effect of 16 ms. For words learned in the spoken modality, the identity prime facilitated recognition as expected (with 86 ms compared to a prime with random letters). Crucially, for words learned in the spoken modality, the prime that differed two or more letters from the target also facilitated recognition of the target; with 58 ms compared to a prime with random letters. Thus, orthographic representations created during a spoken encounter are already activated when exposed to a prime where only half of the letters match with the target. This indicates that such representations are less precise and less complete. In a situation without priming, these low quality orthographical lexical representations would increase lexical competition between the newly learned word and orthographical neighbours, thereby impeding efficient and accurate activation in the spoken modality.

Alongside these contradicting findings and contrasting predictions, there are also studies that report no modality effect whatsoever (Dean et al., 1988; Nelson et al., 2005). All these seemingly contrasting findings may be the result of several substantial differences in studies' methodology. First, studies differed in the extent to which they controlled for the encoding specificity principle (Tulving & Thomson, 1973), which states that recall is enhanced if conditions during encoding match the conditions during retrieval. That is, some studies manipulated modality during the training session, but administered the test session in one modality only. Indeed, these studies report learning to be enhanced in the condition where the modality during the training session matches the modality of the test session (Balass et al., 2010; van der Ven et al., 2015). Second, studies differed substantially in training regimes. Some studies used self-paced learning (Balass et al., 2010; Nelson et al., 2005; van der Ven et al., 2015), whereas exposure time was strictly controlled in other studies (Bakker et al., 2014; Dean et al., 1988). This results in large differences regarding the exposure duration to the novel words, which complicates the comparison of findings. Moreover, the modality effects in the studies may be difficult to interpret. For self-paced learning studies, exposure was not recorded, so participants could have exposed themselves more to words in one modality than the other. The reported learning benefit may be the result of extended self-exposure rather than training modality per se. Even if exposure duration to the written and spoken stimuli was equalized in the strictly controlled studies, by presenting the written word for the duration of the spoken word, exposure duration between training conditions is not comparable. People are able to read faster than the time it takes for the speech to unfold, so when exposure duration is equalized, participants still have relatively more time with the written stimuli, which may boost learning efficiency.

The aforementioned methodological differences complicate interpretation of the results of previous studies. Therefore, the present study aimed to investigate modality effects on novel word learning in a controlled experimental setting. In a new paradigm, participants had to learn 24 Dutch-like, fully transparent pseudowords and an accompanying meaning in the form of a picture of a non-existing object in an implicit, fast-paced training task.

Controlling crossmodal transparency of the words ensured that the orthographic and phonological forms of the words were equally learnable. The use of pictorial rather than linguistic information (e.g., dictionary definition) as semantic context decreases cognitive load (Sweller, van Merrienboer, & Paas, 1998). Learning two pieces of information (word form and meaning) in the same format (linguistic), increases the likelihood of overloading working memory capacity and consequently inhibit learning. Moreover, using a linguistic format would complicate the design of the paradigm, since meanings must be presented in either the written or the spoken modality. The training task was implicit in nature such that participants were not informed that they were required to learn the words and would be tested later. This ensured that participants did not use modality-specific learning strategies. The fast-paced training task asserted control over the exposure duration to written and spoken learning materials. It was also checked whether the fact that written input is available simultaneously, whereas spoken input unfolds over time, affected learning. Therefore, the experimental design yielded three training conditions: spoken, equal written exposure, where written word forms were presented for the duration of the recording of their spoken counterparts, and reduced written exposure, where exposure time is equated rather than equalized by reducing written exposure duration significantly, but still to the point that the participants could comfortably read the written words.

After a 20-minute period of consolidation in the form of a purely visual task, with no orthographical or phonological input, participants were tested on their knowledge of the learned word forms and meanings. In this matching task, participants had to decide whether a picture matched with a previously learned word or not. Accuracy was recorded. This test was administered in either the written or spoken modality to control for the effect of the encoding specificity principle (Tulving & Thomson, 1973).

To control whether the participants in the different experimental groups differed in their language abilities and general intelligence, a nonverbal IQ task (Raven's Advanced Progressive Matrices), vocabulary task (Peabody Picture Vocabulary Test) and word- and nonword reading tasks (One-Minute-Test and Klepel test respectively) were administered.

Based on previous literature, we predicted that training modality would affect word learning. However, since the literature produces contradicting predictions and reports contrasting findings, we could not formulate a clear hypothesis concerning the direction of the modality effect.

Table 6.1: Literature overview modality effects in novel word learning.

Study	N	Design	Training procedure	Test procedure	Outcome variable	Findings
Bakker et al. (2014)	107	4 conditions, between-subjects	20 written words, 20 spoken words. Spoken training: phoneme monitoring. Each word 36x. Written training; letter monitoring. Each word 36x. Written word shown for duration of spoken word. Training on Day 1 and 2	Spoken words: pause detection task (20 learned words, 60 fillers) Written words: semantic decision task (20 learned words, 80 fillers) Free recall task (spoken). 2AFC word recognition task (half of words presented written, half spoken). Tests on Day 1, 2 and 8	Pause detection task: response latencies. Semantic decision task: response latencies. Free recall and 2AFC word recognition tasks: number of words correctly recognized	Lexical competition effect larger for words learned in spoken modality (pause detection task) on Day 2 and 8. Also recall and recognition better for words learned in spoken modality on all days
Balass et al. (2010)	37	3 conditions, within-subjects	105 uncommon words. 35 written + meaning, 35 spoken + meaning, 35 written + spoken. Meaning = definition. Meanings presented written. Self-paced learning for 2.5 hours or 100% correct recognition	Semantic relatedness judgment task. 105 trained words, 105 new words. All words presented written	Accuracy, response latencies, ERPs	Higher accuracy and faster response latencies for words learned in written modality
van der Ven et al. (2015)	64	2 conditions, between-subjects	65 new, 65 known words + meaning. Meaning = definition. Words presented written. Meanings presented written or spoken (equal exposure duration). Self-paced learning: participant goes through list, indicates if they want to see a trial again later. Learning for 2 hours or no more trials are left participants wanted to see again. Training on Day 1	Primed lexical decision task (written prime and written target words). Meaning recall task (written target words). Meaning recognition task (written target words, spoken/ written meaning). Tests on Day 1 and 2	Primed lexical decision task: response latencies. Meaning recall recognition tasks: proportion correct	Larger priming effects for words learned in written modality. No modality effect for meaning recall and recognition
Dean et al. (1988), Experiment 2	80	2 between-subjects test modality conditions, training modality within subjects	40 words, 20 written and 20 spoken. Words shown for 1 second. 1 exposure per word	Old/new recognition task. 20 written training words, 20 spoken training words, 40 new words. Half of words presented written, half spoken	Total number of trained words correctly recognized	No effect of training modality
Nelson et al. (2005)	35	3 conditions, within-subjects	105 uncommon words. 35 written + meaning, 35 spoken + meaning, 35 written + spoken. Meaning = definition. Meanings presented written. Self-paced learning for 2.5 hours or 100% correct recognition	Old/new recognition task. 35 written training words; 35 spoken training words, 35 new words. Half of words presented written, half spoken	Proportion of trained words correctly recognized, corrected for response bias (A')	No effect of training modality

6.2 Experiment 1

6.2.1 Methods

Participants

Ninety-one participants (M = 22.96 years, SD = 2.45, 73 female) were recruited from the participant database of the Max Planck Institute for Psycholinguistics. All participants were students at the Radboud University or HAN University of Applied Sciences. In addition, all were right-handed, with no language, sight or hearing disorders. They had normal or corrected-to-normal vision and hearing and gave written informed consent prior to testing. Ethical approval to conduct the study was provided by the ethics committee of the Faculty of Social Sciences at Radboud University. Participants received € 10,- as compensation for their participation.

Design

The two between-subjects factors were training modality and the test modality. The training modality factor had three levels: *spoken*, *equal written exposure*, where words were presented for the duration of the spoken word, and *reduced written exposure*, where the words were presented for 300 ms. The test modality factor had two levels: written and spoken. There were therefore six between-subjects conditions (Table 6.2). Participants were semi-randomly assigned to a condition.

Table 6.2: Design of Experiment 1.

		Modality	
		Spoken	Written
Training phase	Spoken Equal written Reduced written	Spoken-Spoken Equal written-Spoken Reduced written–Spoken	Spoken-Written Equal written-Written Reduced written-Written

Materials

Pseudowords. Bisyllabic, transparent Dutch pseudowords were created using Wuggy (Keuleers & Brysbaert, 2010). The Wuggy algorithm calculates bigram frequencies of an input list of words, and uses these bigram frequencies of the input list and the bigram frequencies of words in a lexical database of the preferred language to generate new words. For the Dutch language, Wuggy uses the CELEX database (Baayen, Piepenbrock, & Gulikers, 1995). The input list for Wuggy contained pseudowords and real Dutch words that only contained phonologically and orthographical transparent consonants and vowels. Based on this input, the Wuggy algorithm generated 10 bisyllabic pseudowords per input word that matched the morphological, phonological and orthographical rules of Dutch. From this output, 144 words were selected that only consisted of phonological and orthographical consistent consonants and vowels and if they appeared Dutch-like according to native Dutch. From these words, 24 words were selected that were orthographical and phonologically transparent, had a Levenshtein's distance (Levenshtein, 1966) of above three, did not contain phonologically and orthographically confusable phonemes or graphemes, and were not closely reminiscent of existing Dutch words. The number of phonemes and graphemes varied between four and eight and the number of graphemes between five and nine. The 24 pseudowords used in the experiment can be found in Appendix I.

Meaning. We selected 24 pictures of unknown objects from the Novel Object and Unusual Name (NOUN) Database (Horst & Hout, 2016). The pictures were chosen to be visually dissimilar to each other. Pictures were randomly paired with one of the words for each group of six participants (one participant in every condition).

Procedure

The experiment took place in a single test session (Figure 6.1). For the training phase, participants were explained that several word form-picture pairs would be presented and that these pairs were only shown shortly. It was not mentioned that the participants would be tested at a later stage; participants were only instructed to pay close attention to the pairs. The instructions of the filler task were provided before the training task to ensure that participants did not receive any written or spoken input immediately after learning. After the filler task, participants took part in the test phase. All instructions prior to the test phase were presented in both the spoken and the written modality. Following the test phase, participants completed a questionnaire, a word and a nonword reading task, a vocabulary task, and a word retyping task. Note that the participants in the reduced written training condition were tested two months after the participants in the spoken and equal written training conditions due to the summer break.

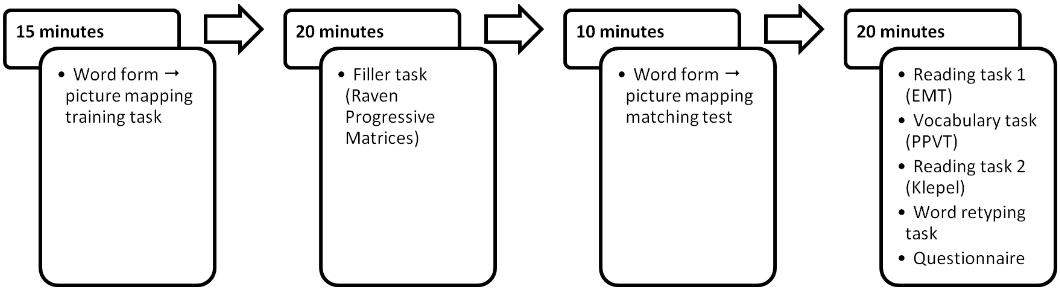


Figure 6.1: Experimental procedure of Experiment 1.

Tasks

Training task. Using a flashing task, participants were exposed to the picture-word form pairs (Figure 6.2). Participants first saw a fixation cross at the centre of the screen for 250 ms. Then a picture of a non-object appeared for 1000 ms. The exposure to the picture was short, so that it could only be visually processed. Following picture presentation, a fixation cross appeared for 250 ms. Then, participants either saw the written word form of the corresponding word or heard the target word. The trial ended with a visual and auditory mask.

For the equal written exposure condition, exposure duration to the word form was limited to the speech duration of the spoken word form. The speech durations and thus written exposure duration of the pseudowords in the equal written condition varied between 664 and 993 ms (M = 860). In the reduced written training condition, exposure time was reduced to 300 ms. Previous reports in the literature (Rayner, Pollatsek, Ashby, & Clifton Jr, 2012; Schilling, Rayner, & Chumbley, 1998; Sereno, Rayner, & Posner, 1998) and a pilot study indicated that short, frequent words are read within around 100 ms, whereas longer, infrequent words take about 200 – 250 ms to read. Because our learning materials consisted of pseudowords, which may take longer to read, we decided to set the exposure duration for the reduced written training condition at 300 ms. This is a reduction of one-third compared to the equal written training condition (M = 860 ms), but most likely enough time to fully read the word for participants varying in their reading ability.

Reducing exposure time to written words on in the reduced written exposure condition inevitably led to reduced trial durations compared to the other two trial types. To avoid confounds in memory consolidation, we decided to add 280 ms to both the fixation cross at the beginning of the trial (250 + 280 = 530 ms) and to the mask at the end of the trial (500 + 280 = 780 ms) for these trials (bottom Figure 6.2).

A careful pilot study where number of exposures to a picture-word form pair and blocking procedure was manipulated informed our decision to use seven exposures per word and blocked learning in the form of three blocks with three sets of eight words. In each block, one set of pairs was shown four times, alternated with the two other sets. Order constraints ensured that subsequent word forms were not phonologically or orthographically similar (Levenshtein's distance above three and different onset). There were no breaks between the blocks. The training task started with eight practice trials with known words and pictures (e.g., "train", "bread") so that participants became familiar with the fast pacing of the task.

To ensure that participants were paying attention to the training task, which required them only to look at the flashing word form-picture pairs, an additional attention task was implemented in the training task. Each of the eight pictures from the practice trials appeared in-between the learning trials. At the sight of these pictures, participants were required to press a button within 2000 ms. The attention control trials were randomly distributed within the learning trials, but the place and order of the attention control trial was the same for every participant. participants' responses to the attention control was recorded.

Filler task: nonverbal IQ. The computerized 20-minutes version of the Raven's Advanced Progressive Matrices (Raven et al., 1998) was used as the filler task. On each trial, participants saw a panel of eight geometrical figures, with the space for a ninth figure left blank. From a set of eight candidates shown at the bottom section of the screen, they had to select the figure that completed the sequence. They indicated their choice by clicking on the chosen item with a mouse. Participants could skip items by clicking on a button a labelled "Skip"; these items were shown again at the end of the test. When they did not know the answer to a skipped item, participants could click on an "I don't know"-button. There were six practice items and 36 test items, increasing in difficulty. Participants had 20 minutes to complete the test. Throughout the test, a clock in the right top corner of the screen showed the time remaining. A participant's score was their number of correct responses.

Test phase. In the matching test, participants saw a pair of a picture and the corresponding word form from the training task (matching trial), or a pair of one of the other 23 pictures and a word form the training task (mismatching trial). Participants were required to decide whether the pair was a match or

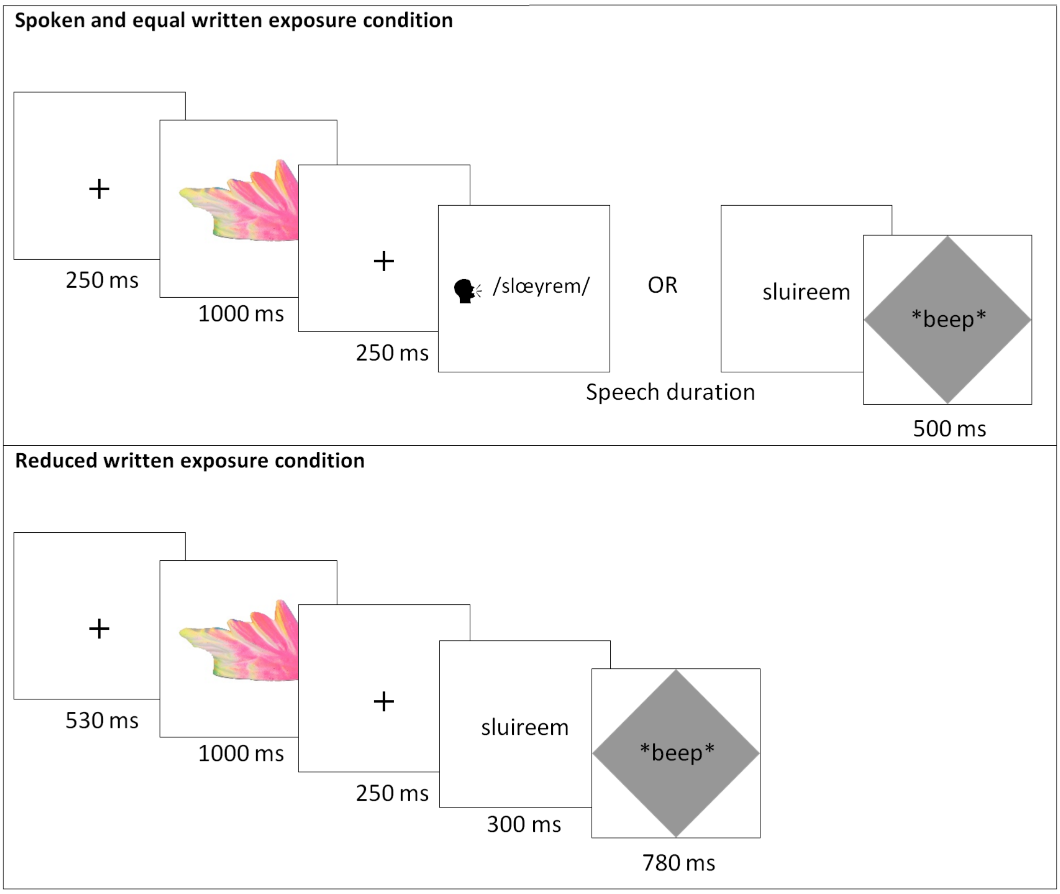


Figure 6.2: Trial example of the picture-word form pair in the spoken and equal written exposure (top) and reduced written exposure conditions (bottom) of the training task of Experiment 1.

not according to what they had learned in the training phase. The trial structure (Figure 6.3) was as follows. Participants saw a fixation cross for 250 ms, a picture for 1000 ms, another fixation cross for 250, then read or heard a word. Then they saw a blank screen where the participant had to push a button on the button box to state whether the picture and word form matched according to what they had learned during the training (right button) or mismatched (left button). The response time was limited to 2 seconds to ensure that the reaction times were not affected by strategic effects. When the reaction time exceeded 2 seconds after word onset, a red hourglass was shown and a beep was played for 500 ms to indicate that the response had been too slow. After a response or two seconds if the participant had not responded, a new trial started.

Written words were presented equally long as the speech duration of their spoken counterpart. On half of the trials, the picture matched the word form and on half of the trials the picture was a foil (i.e., one of the other 23 pictures) and did not match the word form. The foil pictures had several constraints with regard to their relationship to the mismatched word forms. Phonologically and orthographically, the word form that was paired with the foil picture during the training phase had a different onset than the target word form during the test. Moreover, the target word form on the test and the word form that corresponded with the foil picture had a Levenshtein's distance of four or higher. The order of the trials was semi-randomized and bound by several order constraints. First, the next (foil) picture or word could not be the previous ten word forms or (foil) pictures. Also, 50 percent of the target words were shown with a correct picture first before they appeared with a foil picture, and the other half of the word forms were first seen with a foil picture before being presented with the correct picture. Similar to the training task subsequent word forms were not phonologically or orthographically similar (Levenshtein's distance below four and different onset). Accuracy and response time were recorded during this task.

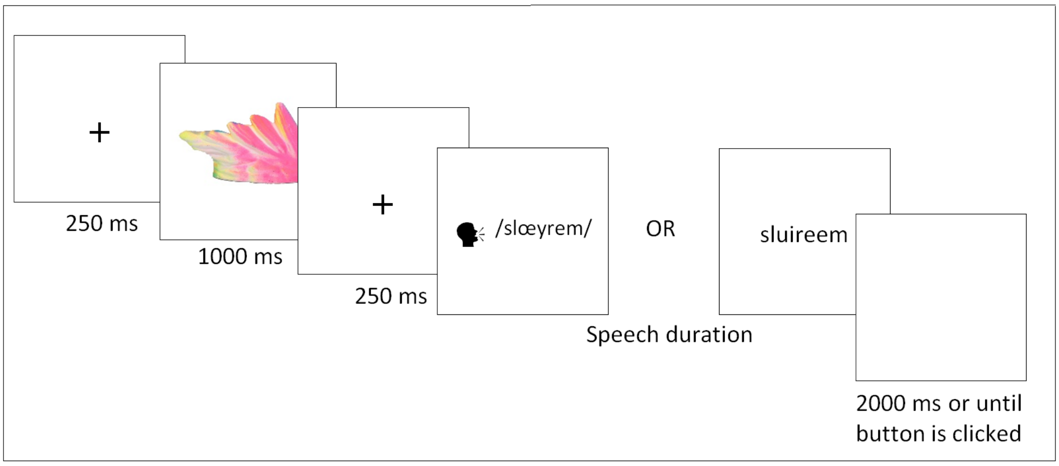


Figure 6.3: Trial example of picture-word form pair on the test in Experiment 1.

Reading ability. Two Dutch word reading tasks were used to measure reading ability. The tasks were computerized, rather than administered with pen and paper, which is the canonical procedure. Word reading ability was assessed with the Eén-Minuut-Test [One-Minute-Test] (Brus & Voeten, 1973) which uses existing Dutch words. For nonword reading ability the Klepel task (van den Bos et al., 1994) was used, which consists of nonwords. For both tasks, participants were are instructed to read aloud a list of 116 words, divided in four columns of 29 words, as fast and clear as possible from the top left to the bottom right. These words become progressively more difficult to read in terms of phoneme complexity and syllable length (range: 1 – 5 syllables). Participants' aloud reading was recorded for one (EMT) or two minutes (Klepel). These recordings were later used to count the number of errors and the total number of words read in order to calculate the participants' reading scores. The final score was the number of errors (incorrectly pronounced words) was subtracted from the number of words read within one minute.

Receptive vocabulary. Participants' receptive vocabulary size was assessed using a digitized version of the Peabody Picture Vocabulary Test (Dunn & Dunn, 1997; Schlichting, 2005, for the Dutch translation). On each trial, participants first previewed four numbered line drawings on their screen. When they were ready, they pressed the Return key on their keyboard to hear the word. They had to indicate which of the pictures best corresponded to the meaning of the spoken word by typing the corresponding number (1, 2, 3, or 4). Following the standard protocol for the test, items were presented in blocks of twelve items, with blocks increasing in difficulty. The starting level was 13, the best level participants could attain was 17. The test ended when a participant made nine or more errors within one block. Participants took, on average, twelve minutes to complete the test (range: 8 to 15 minutes). The participants' raw score (the serial number of their last item minus the number of errors made during the test) was standardized to correct for participants' age and transformed into a percentile score as provided by the manual.

Word retyping task. Participants in the reduced written training condition performed an additional task to assess whether they could read the pseudowords sufficiently when presented for only 300 ms. One-hundred twenty pseudowords generated for the training task, excluding the 24 words of the training task, were used in this word retyping task. Half of the words were shown for 300 ms and the other half for 860 ms, which was the average exposure duration of a word in the spoken and equal written training condition. Additionally, a mask was added to ensure that any ongoing visual processing of the written word was stopped after the word had disappeared from the screen. A practice task of eight additional trials with newly generated pseudowords preceded the retyping task. Due to a programming error, this task was only administered to the participants in the reduced written training condition and the written test condition and not to the participants in the reduced written training and spoken test condition.

Questionnaire. The questionnaire tested whether the participants knew any of the learned words or had seen any of the pictures before, and if so which ones.

Furthermore, it was assessed whether they had performed the filler task, reading ability tasks and the vocabulary tasks before. It was also asked whether they had used any strategies to learn the picture-word form pairs.

6.2.2 Results

Trials were removed from further analysis if the participant did not react within 2 seconds and if RTs were below 300 ms (Burke et al., 2017). Two participants were removed from the analyses because half of their responses were too slow or too fast. Applying the same criterion to the remaining participants, led to 3.18% of data exclusion. Misses and false alarms were counted as errors.

Four two-way ANOVA's with the four individual differences variables (nonverbal IQ, vocabulary, word- and nonword reading) as dependent variables and training and testing modality as independent variables indicated that participants in the four conditions did not differ in their nonverbal IQ, vocabulary size, word reading and nonword reading ability (see Tables 6.3 and 6.4).

We calculated d-prime (D') to estimate each participants' sensitivity to take into account participants' response biases to matching or mismatching trials. D' was calculated in R in the psycho package (Makowski, 2018) in R version 3.6.2 (R Development Core Team, 2008). A log-linear correction was applied. This overcomes D' estimation issues for data where proportions of hits and false alarms are close or equal to 0 or 1 (Hautus, 1995; Stanislaw & Todorov, 1999). Figure 6.4 depicts participants' sensitivity on the matching test split out by condition. Descriptive statistics are displayed in Table 6.3. The effect of training modality and test modality on D' was analysed with a two-way ANOVA. Full model output is displayed in Table 6.4.

The analysis indicated no effect of training modality ($p = .08$) and no effect of test modality ($p = .20$). A significant interaction effect between training and test modality was observed ($p = .01$). Simple main effect analyses of training type indicated that for the spoken test (Figure 6.5, left side), participants in the equal written training learned more words than participants in the other two training conditions ($F(2,43) = 5.14$, $p = .01$). For the written test condition (Figure 6.5, right side), however, there was no difference in learning between training conditions ($F(2,40) = 1.51$, $p = .23$).

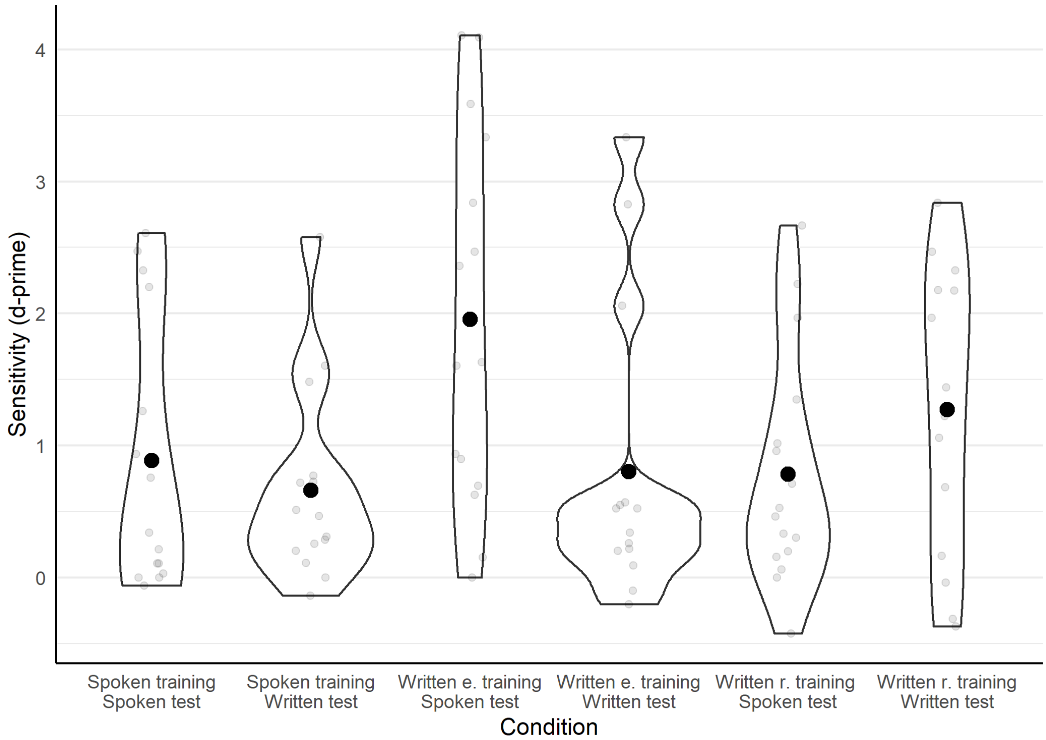


Figure 6.4: Sensitivity during matching test of Experiment 1 by training/test modality condition. Written e. training = Written equal training, Written r. training = Written reduced training.

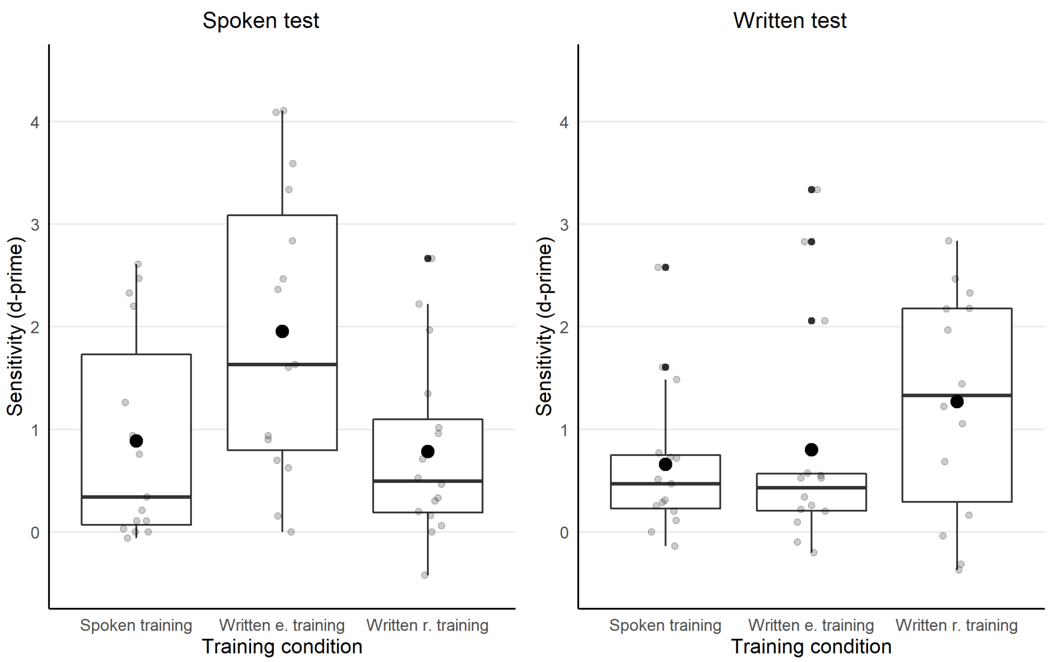


*Figure 6.5: Simple main effects of training modality for the spoken test (left) and written test (right).*Written e. training = Written equal training, Written r. training = Written reduced training.

Table 6.3: Descriptive statistics of proportion mismatching (MM) and matching trials (M), sensitivity (d-prime) and the individual differences variables of Experiment 1.

	SS		SW		WeS		WeW		WrS		WrW	
	M (SD)	Range	M (SD)	Range	M (SD)	Range	M (SD)	Range	M (SD)	Range	M (SD)	Range
MM	.74 (.44)	.43, .96	.68 (.47)	.30, 1	.83 (.38)	.50, 1	.66 (.47)	.38, 1	.69 (.46)	.32, 1	.82 (.39)	.43, 1
M	.56 (.50)	.14, .87	.55 (.50)	.21, .88	.76 (.43)	.50, 1	.59 (.49	.38, .92	.58 (.49)	0, .88	.61 (.49)	0, .92
D'	0.89 (1.02)	0.06, 2.61	0.66 (0.72)	0.14, 2.58	1.96 (1.41)	0, 4.11	0.80 (1.11)	0.20, 3.34	0.78 (0.87)	0.42, 2.66	1.27 (1.10)	0.37, 2.84
IQ	21.78 (5.22)	16, 33	24.00 (4.61)	17, 32	22.80 (7.09)	6, 35	20.57 (5.14)	13, 32	22.88 (5.21)	14, 32	22.64 (6.43)	14, 32
Voc	56.13 (24.63)	1, 88	55.87 (18.75)	23, 77	65.73 (28.02)	5, 97	51.64 (26.20)	5, 87	58.44 (18.90)	14, 79	60.29 (22.01)	27, 94
WR	95.47 (11.35)	73, 113	88.40 (16.53)	57, 109	92.67 (17.69)	53, 115	86.29 (20.38)	51, 116	91.75 (13.98)	71, 115	96.86 (17.02)	64, 116
nWR	63.67 (12.47)	40, 83	63.33 (12.98)	40, 86	67.93 (10.94)	51, 88	69.67 (11.65)	50, 86	66.44 (12.14)	49, 89	69.07 (11.02)	56, 92

Note. M = accuracy proportion matching trials, MM = accuracy proportion mismatching trials, D' = sensitivity (d-prime), IQ = nonverbal IQ, Voc = vocabulary, WR = word reading, nWR = nonword reading, SS = spoken training-spoken test, SW = spoken training-written test, WeS = equal written training-spoken test, WeW = equal written training-written test, WrS = reduced written training-spoken test, WrW = reduced written training-written test.

Table 6.4: Model output Experiment 1.

	Training modality	Test modality	Training modality * Test modality
D’	F(2,83) = 2.62, p = .08	F(1,83) = 1.66, p = .20	**F(2,83) = 4.48, p = .01**
IQ	F(2,83) = 0.39, p = .68	F(1,83) = 0.01, p = .94	F(2,83) = 1.09, p = .34
Voc	F(2,83) = 0.18, p = .84	F(1,83) = 0.68, p = .41	F(2,83) = 1.01, p = .37
WR	F(2,83) = 0.57, p = .57	F(1,83) = 0.63, p = .43	F(2,83) = 1.31, p = .28
nWR	F(2,82) = 1.30, p = .28	F(1,82) = 0.11, p = .74	F(2,82) = 0.13, p = .88

Note. Significant effects are displayed in bold. D’ = sensitivity (d-prime), IQ = nonverbal IQ, Voc = vocabulary, WR = word reading, nWR = nonword reading.

The data from the word retyping task were analysed using a linear mixed-effects model (lme4 package: Bates, Machler, Bolker, & Walker, 2015). In the linear mixed-effect model accuracy of retyping was used as dependent variable, and exposure duration (300 or 860 ms) and word length were added as experimental fixed effects. The random effects structure included a random intercept by participant and by word. Exposure duration was treatment contrast coded, with 300 ms being the reference level.

Analyses revealed no effect of exposure duration (estimate = 0.53 SE = 1.81, z = 0.29, p = .77), which means that participants could identify and retype the words shown for 300 and 860 ms equally accurately. A word length effect was found, where longer words were retyped with less accuracy (estimate = -0.54, SE = 0.15, z = -3.58, $p < .001$). The interaction between exposure duration and word length was not significant (estimate = 0.04, SE = 0.24, z = 0.17, p = .86).

6.2.3 Conclusions Experiment 1

Experiment 1 investigated modality effects on novel word learning in a controlled, experimental setting applying a novel paradigm. We observed no evidence for a main effect of training modality. An interaction indicated that participants who had taken part in the written training were more accurate on the spoken matching test. This crossmodal effect is unexpected, since Tulving and Thomson (1973)’s encoding specificity principle would predict that accuracy would be highest for the groups that performed the training and test in the same modality.

Interestingly, the written learning benefit observed in the equal written exposure training-spoken test group disappeared when reducing exposure to the written stimuli. This suggests that the learning benefit observed for the equal written training – spoken test group is caused by reading being faster than the process of unfolding speech. That is, when presented with the same information (in the case of the present experiments bisyllabic words), the process of information intake in the written modality is faster compared to the information intake processes in the spoken modality. This allows readers to move faster to the next piece of information, whereas the listener is bound by the articulation speed of the listener. This processing speed benefit results in more time and opportunities for multiple exposure in the written modality, ergo, more efficient learning, if exposure time to written and spoken information is equal.

Many theories of learning capitalize on the importance of consolidation for efficient storage and retrieval of novel information. For example, the Complementary Learning Systems (CLS) theory (McClelland, McNaughton, & O'Reilly, 1995), specifically adapted to word learning by M. H. Davis and Gaskell (2009) describes how a word is at first encoded in the hippocampal episodic memory, and, through interleaved reactivation during a period of consolidation, gradually integrated within the mental lexicon. Indeed, many studies report that lexical integration of novel word forms only occurs after a 24 hour delay that includes a period of sleep (Bakker et al., 2014; Bakker-Marshall et al., 2018; Bowers, Davis, & Hanley, 2005; Clay, Bowers, Davis, & Hanley, 2007; Dumay & Gaskell, 2007, 2012; Dumay, Gaskell, & Feng, 2004; Tamminen & Gaskell, 2013; Tham, Lindsay, & Gaskell, 2015; van der Ven et al., 2015). Lexical integration increases as time progresses (Bakker et al., 2014; Dumay & Gaskell, 2012; Gaskell & Dumay, 2003; Tamminen & Gaskell, 2008). Also, explicit memory, measured with recall and recognition tasks, have been found to improve after a night of sleep (Bakker et al., 2014; Dumay & Gaskell, 2007, 2012; Dumay et al., 2004; Tamminen, Payne, Stickgold, Wamsley, & Gaskell, 2010) and to be stable over time (Bakker, Takashima, van Hell, Janzen, & McQueen, 2015a, 2015b; van der Ven et al., 2015).

In Experiment 2, we aimed to investigate modality effects in novel word learning as consolidation progresses. To this end, we used the same word learning paradigm as in Experiment 1, with two additional test sessions the day and a week after training and the first test. Previous literature reported contradicting findings regarding the direction of the modality effect after a period of consolidation (Bakker et al., 2014; van der Ven et al., 2015), and as such we could not formulate a clear hypothesis concerning the direction of the modality effect.

6.3 Experiment 2

6.3.1 Methods

Participants

In total, 60 participants (M = 23.21 years, SD = 2.27, 49 female), recruited from the participant database of the Max Planck Institute for Psycholinguistics, participated in this experiment. All participants were right-handed and had no language, sight or hearing disorders. Participants were compensated with € 15,- for their participation.

Design

The study had a 2x2x3 mixed design with modality during training and modality during testing as between-subject factors. Words were either presented in the written or spoken modality during training. The test was provided in either the written or the spoken modality. The within-subjects factor was time: participants were tested on Day 1, Day 2 and Day 8.

Materials

The materials in experiment were the same as in Experiment 1.

Tasks

Most tasks were the same as in Experiment 1, except for the matching test and an additional pseudoword reading task. During the training, exposure duration was equalized: written words were presented for the duration of their spoken counterparts. The reduced written training condition was removed as the learning benefit found in Experiment 1 only occurred in the equal written training condition.

Test phase. The matching test was almost the same as the matching test in Experiment 1. Given that two participants from Experiment 1 had to be excluded due to too fast and/or too slow responses, we extended the response time from two to five seconds to reduce data loss.

On Day 2 and Day 8, participants performed the test online in the internet browser of their personal computers. First, their audio devices were checked with an audio test, where they heard a word and had to click the correct word out of on one of four written options. Then, they received the instruction of the matching test both in auditory and written format. The instructions were the same as the instructions for the matching test on Day 1. The matching test was the same as the matching test on Day 1, except that the word forms were paired with different foil pictures on the different days. As such, the mismatching pairs were pairs never seen before. After the final test on Day 8, participants were debriefed with an explanation of the experiment. Accuracy and reaction times were recorded during the matching task on Day 1. In the online sessions on Day 2 and 8 only accuracy was recorded.

On Day 8, the test task was performed twice. The first 'regular' run was administered in the same modality as the test of Day 1 and 2, but during the second run, the word forms were presented in the other modality. Thus, if participants were tested in the spoken modality on Day 1, 2, and 8, during their crossmodal test on Day 8 the word forms were presented written and vice versa. The order of the trials and word form – foil picture pairs during the crossmodal test on Day 8 were the same as that of the regular test on Day 8. By using the same mismatching trials in both tests, the trial order could be kept the same in both tests, which ensured that the time between the trials with the same (matching or mismatching) pairing was the same for all trials.

Reading ability. In addition to the word reading ability (Eén-Minuut-Test [one-minute-test, Brus and Voeten (1973)) and nonword reading ability (Klepel Test, van den Bos et al. (1994)) tasks, a pseudoword reading ability task was added to the battery to check whether the Klepel was was a good indicator of people's ability to read the pseudowords in our word learning paradigm.

In this task, the word lists contained the Dutch-like pseudowords that were generated for the training task, but not part of the final set of 24 words that were used in the training task. Again, participants' aloud reading was recorded for two minutes and these recordings were used to calculate the number of correctly read words in one minute. The correlation between this pseudoword reading ability task and the Klepel nonword reading ability task was high ($r = .91$, $p < .001$), indicating that the Klepel nonword reading ability task was a good indicator of people's ability to read the transparent, bisyllabic Dutch-like pseudowords used in the training task.

Questionnaire. One question was added to the questionnaire in which it was asked whether participants preferred learning from speech or from reading text, or did not have a preference.

Procedure

The experiment consisted of three different sessions run on Day 1, Day 2 and Day 8 (Figure 6.6). The session on Day 1, was the same as Experiment 1. On Day 2, participants performed the matching test at home. Through email, they received a link to an online web-version of the matching test. On Day 8, exactly a week after the training phase on Day 1, participants again received a link to complete the matching test online. Participants first completed the regular run of the matching test, which was in the same modality as all previous tests this participant performed. Next, they participated in the crossmodal matching test: if participants had performed all test sessions in the spoken test condition, they now received the written test and vice versa.

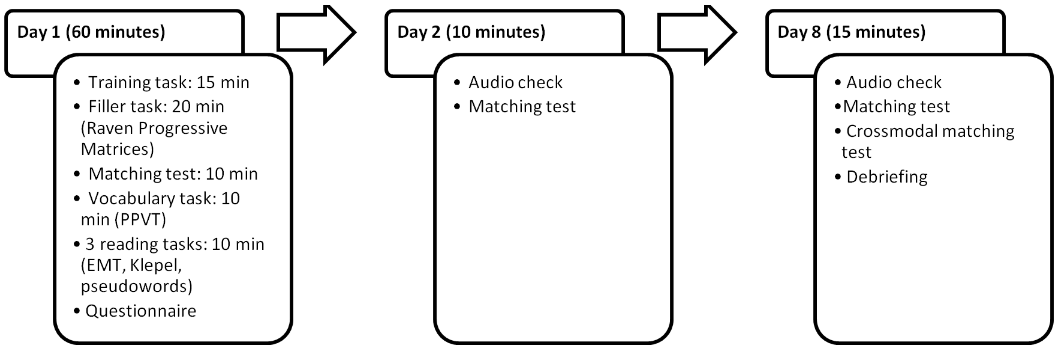


Figure 6.6: Experimental procedure of Experiment 2.

6.3.2 Results

Data from one participant were removed because the training task was administered twice due to a computer error during testing. Moreover, data on Day 8 were removed for one participant, because of an administrative error (they performed the test in the wrong modality). Six participants did not perform the test on Day 2 and six other participants did not perform the test on Day 8. For Day 1, trials were removed if RTs were below 300 ms or if the participant did not react within five seconds, which resulted in the removal of 2.01% of the trials. Trials on Day 2 and Day 8 were removed if participants did not react within five seconds (Day 2: 1.27% of the data, Day 8: 0.43%, Day 8 crossmodal: 2.29%).

Misses and false alarms were counted as errors (Day 1: 33.20%, Day 2: 36.89%, Day 8: 38.33%, Day 8 crossmodal: 35.64%).

Four two-way ANOVA's with the four individual differences variables (nonverbal IQ, vocabulary, word- and nonword reading ability) as dependent variables and training and testing modality as independent variables indicated that participants in the four conditions did not differ in their nonverbal IQ, word reading and nonword reading ability (see Tables 6.5 and 6.7). For vocabulary size, a significant effect of training modality indicated that participants in the spoken training modality groups had on average a slightly larger vocabulary than participants in the written training modality groups ($p = .02$).

As in Experiment 1, log-linear corrected d-prime (D') was calculated for each participant individually. Figure 6.7 depicts participants' sensitivity on the matching test split out by condition and test session. Descriptive statistics are displayed in Table 6.5. The effect of test session, training modality and test modality on D' was then analysed with a three-way ANOVA. Full model output is displayed in Table 6.6. The analysis indicated no effect of test session, training modality or test modality, nor any interactions between these variables.

Thus, we did not observe the interaction effect between training and test modality observed in Experiment 1. To investigate this further, we performed a two-way ANOVA with training and test modality as independent variables using on the data of the test session on Day 1. The analysis, displayed in Table 6.7, indeed indicated no effects of training modality, test modality or an interaction between these variables.

Table 6.5: Descriptive statistics of proportion mismatching (MM) and matching trials (M) and sensitivity (d-prime) for each test session, and the individual differences variables of Experiment 2.

	SS		SW		WS		WW	
Day 1								
	M (SD)	Range	M (SD)	Range	M (SD)	Range	M (SD)	Range
MM	.74 (.44)	.38, 1	.70 (.46)	.33, .96	.74 (.44)	.54, 1	.75 (.44)	.21, .92
M	.61 (.49)	.04, .92	.61 (.49)	.29, .88	.52 (.50)	.29, .71	.68 (.47)	.42, .83
D’	1.00 (0.80)	-0.21, 2.08	0.88 (1.08)	-0.73, 2.45	0.74 (0.61)	-0.10, 2.21	1.10 (0.60)	-0.17, 2.02
Day 2								
	SS		SW		WS		WW	
	M (SD)	Range	M (SD)	Range	M (SD)	Range	M (SD)	Range
MM	.74 (.44)	.40, 1	.70 (.46)	.42, .88	.76 (.43)	.46, .96	.75 (.43)	.42, .96
M	.55 (.50)	.04, .90	.54 (.50)	.33, .83	.49 (.50)	.25, .75	.52 (.50)	.29, .75
D’	0.87 (1.09)	-0.60, 2.47	0.64 (0.79)	-0.40, 1.94	0.71 (0.53)	-0.12, 1.66	0.77 (0.71)	-0.73, 2.08
Day 8								
	SS		SW		WS		WW	
	M (SD)	Range	M (SD)	Range	M (SD)	Range	M (SD)	Range
MM	.73 (.45)	.42, .1	.78 (.42)	.39, .96	.78 (.42)	.58, .96	.74 (.44)	.48, .96
M	.48 (.50)	0, .75	.55 (.50)	.24, .88	.43 (.50)	.13, .75	.45 (.43)	.25, .83
D’	0.56 (0.71)	-0.47, 2.20	0.98 (0.85)	-0.68, 2.36	0.61 (0.66)	-0.67, 1.39	0.62 (0.84)	-0.43, 2.20
Day 8 crossmodal								
	SS		SW		WS		WW	
	M (SD)	Range	M (SD)	Range	M (SD)	Range	M (SD)	Range
MM	.75 (.43)	.55, 1	.79 (.41)	.33, .96	.75 (.43)	.55, 1	.80 (.40)	.58, 1
M	.50 (.50)	0, .88	.61 (.49)	.25, .96	.47 (.50)	.21, .79	.50 (.50)	.25, .71
D’	0.66 (0.60)	-0.10, 1.72	1.23 (0.99)	-0.72, 2.64	0.66 (0.55)	0.12, 1.85	1.00 (0.75)	0, 2.58
IQ	23.40 (6.53)	10, 32	23.07 (4.83)	15, 30	25.13 (4.17)	15, 31	21.00 (6.23)	11, 31
Voc	61.87 (20.37)	27, 99	62.27 (17.66)	37, 92	46.67 (27.87)	4, 84	50.40 (21.12)	12, 82
WR	91.93 (14.61)	59, 115	88.27 (19.24)	53, 111	90.00 (17.15)	69, 115	93.60 (16.23)	52, 115
nWR	69.67 (11.96)	41, 89	64.27 (10.29)	49, 77	63.40 (11.49)	47, 85	64.87 (13.18)	44, 95

Note. M = accuracy proportion matching trials, MM = accuracy proportion mismatching trials, D’ = sensitivity (d-prime), IQ = nonverbal IQ, Voc = vocabulary, WR = word reading, nWR = nonword reading. SS = spoken training-spoken test, SW = spoken training-written test, WS = written training-spoken test, WW = written training-written test.

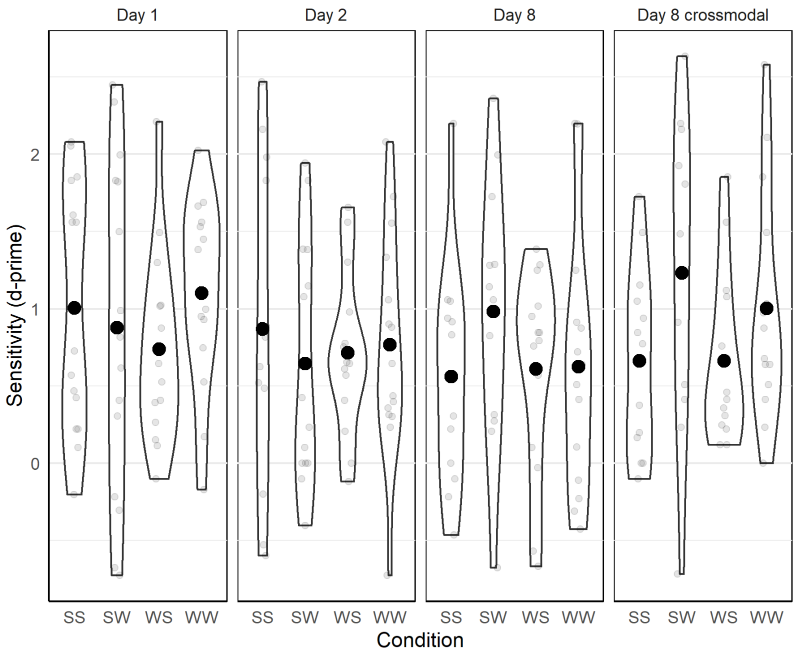


Figure 6.7: Sensitivity during matching test of Experiment 2 by test session and training/test modality condition. SS = spoken training-spoken test, SW = spoken training-written test, WS = written training-spoken test, WW = written training-written test.

Table 6.6: Model output Experiment 2 combining data from all test sessions.

	D'
Test session	F(3,204) = 1.23, p = .30
Training modality	F(1,204) = 0.41, p = .53
Test modality	F(1,204) = 2.68, p = .10
Test session * Training modality	F(3,204) = 0.12, p = .95
Test session * Test modality	F(3,204) = 1.06, p = .37
Training modality * Test modality	F(1,204) = 0.05, p = .82
Test session * Training modality * Test modality	F(3,204) = 1.04, p = .38

Note. Significant effects are displayed in bold. D' = sensitivity (d-prime).

Table 6.7: Model output Experiment 2 for test session on Day 1 (D') and the individual differences measures.

	Training modality	Test modality	Training modality * Test modality
D' Day 1	F(1,56) = 0.01, p = .92	F(1,56) = 0.33, p = .57	F(1,56) = 1.44, p = .24
IQ	F(1,56) = 0.01, p = .91	F(1,56) = 2.45, p = .12	F(1,56) = 0.18, p = .19
Voc	**F(1,56) = 5.64, p = .02**	F(1,56) = 0.13, p = .72	F(1,56) = 0.09, p = .77
WR	F(1,56) = 0.15, p = .70	F(1,56) = 0.00, p = .99	F(1,56) = 0.69, p = .41
nWR	F(1,56) = 0.87, p =. 36	F(1,56) = 0.42, p = .52	F(1,56) = 1.28, p = .26

Note. Significant effects are displayed in bold. D' = sensitivity (d-prime), IQ = nonverbal IQ, Voc = vocabulary, WR = word reading, nWR = nonword reading.

6.3.3 Conclusions Experiment 2

The aims of Experiment 2 were to investigate whether a modality effect on learning efficiency would be influenced by an extended period of consolidation. To this end, we augmented Experiment 1 with two additional online test sessions on Day 2 and Day 8.

Our results did not show evidence of a main effect of training modality, or an interaction of training modality and test session. This indicates that training modality did not influence word learning efficiency on any of the test sessions. The analysis on the data of Day 1 also revealed no evidence for an effect of training modality. Thus, we did not replicate the written learning benefit for the written training – spoken test condition observed in Experiment 1. This is surprising, given that the training phase on Day 1 of Experiment 2 was the same the training phase of Experiment 1.

One reason for the lack of a modality effect of training phase may be the small size of the effect. This would also explain previous reports of null-effects (Dean et al., 1988; Nelson et al., 2005). Indeed, the effect size of training, as calculated with the anova_stats function of the package sjstats (Lüdecke, 2019), was medium in Experiment 1 ($\eta^2 = 0.05$) and small for the test session on Day 1 of Experiment 2 of ($\eta^2 = 0.002$) (J. Cohen (1988): small: 0.01, medium: 0.06, large: 0.14). Small effects require large sample sizes in order to be detected. For Experiment 3 we therefore increased the number of participants per condition. Since our main interest is the effect of modality at training, we decided to simplify the design by removing the test modality manipulation. Thus, participants in Experiment 3 were trained in either the spoken or written modality, and performed the test in the same modality as the training (cf. Tulving & Thomson, 1973). The maximal group size that could be realised within the time constraints of the study was 60 per condition. With two modality conditions, this resulted in 120 participants in total. Compared to the previous two experiments, we thereby quadrupled the number of participants per condition.

6.4 Experiment 3

6.4.1 Methods

Participants

One-hundred twenty-three participants (M = 22.81 years, SD = 2.58, 100 female) were recruited from the participant database of the Max Planck Institute for Psycholinguistics. All participants were right-handed and had no language, sight or hearing disorders. Participants were compensated with € 10,- for their participation.

Design

Experiment 3 only concerned the within-modality conditions. The two cross-modal conditions were dropped. This resulted in two conditions: spoken training-spoken test, versus written training-written test. Participants were pseudo-randomly assigned to a condition.

Materials

The materials in the experiment, both the pictorial meaning and the word forms, were the same as in Experiment 1.

Tasks

Tasks were the same as in Experiment 1, except that during the matching task, participants had five seconds to respond, as in Experiment 2.

Procedure

The procedure of Experiment 3 was the same as the procedure of Experiment 1.

6.4.2 Results

As in the previous experiments, trials were removed if participants did not respond within five seconds, or if RTs were below 300 ms (1.74%). Misses and false alarms were counted as errors (36.12%). Four two-way ANOVA's with the four individual differences variables (nonverbal IQ, vocabulary, word- and nonword reading ability) as dependent variables and modality condition as independent variable indicated that participants in the four conditions did not differ in their nonverbal IQ, vocabulary size, word reading and nonword reading ability (see Tables 6.8 and 6.9).

As in Experiment 1, log-linear corrected d-prime (D') was calculated for each participant individually. Figure 6.8 depicts participants' sensitivity on the matching test split out by modality condition and test session. Descriptive statistics are displayed in Table 6.8. The effect of modality condition was then analysed with a one-way ANOVA. Full model output is displayed in Table 6.8. The analysis indicated no effect of modality condition.

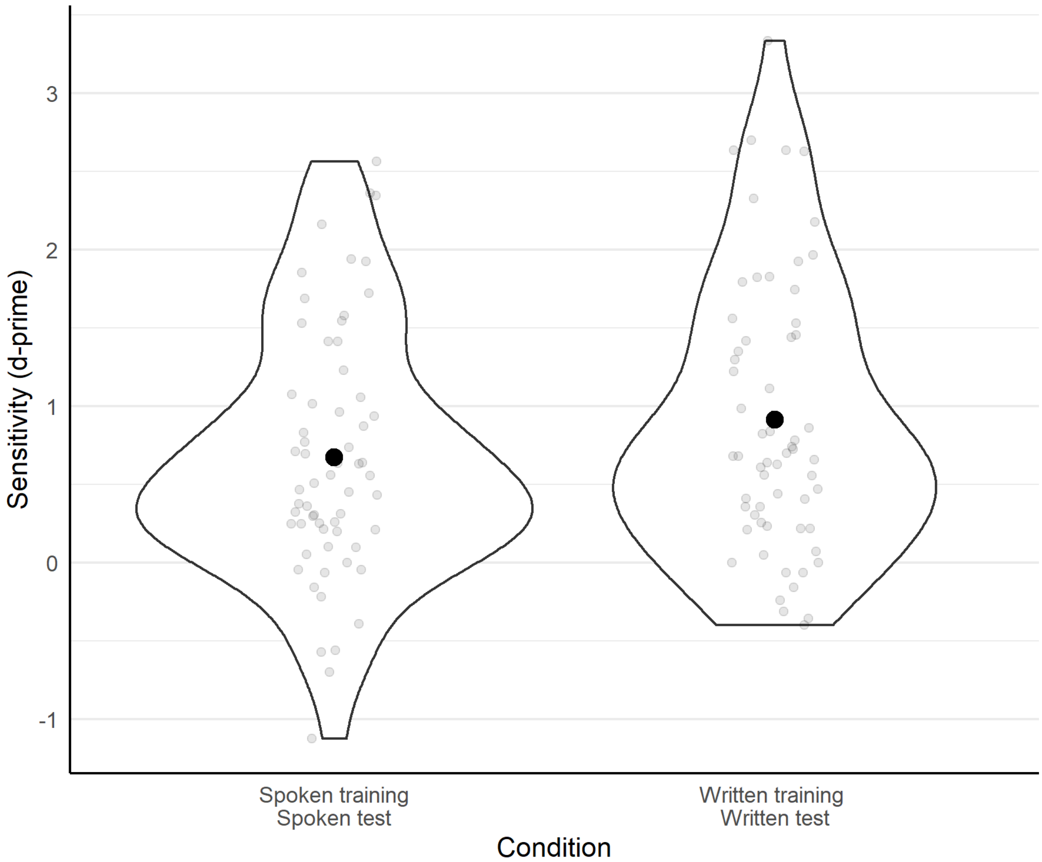


Figure 6.8: Sensitivity during matching test of Experiment 3 by modality condition.

Table 6.8: Descriptive statistics of proportion mismatching (MM) and matching trials (M) and sensitivity (d-prime) for each test session, and the individual differences variables of Experiment 3.

	SS		WW	
	M (SD)	Range	M (SD)	Range
MM	.75 (.45)	.33, 1	.71 (.46)	.33, 1
M	.53 (.50)	0, -.92	.62 (.49)	.15, 1
D'	0.67 (0.80)	-1.12, 2.56	0.91 (0.88)	-0.40, 3.34
IQ	21.95 (4.01)	15, 32	20.79 (4.94)	9, 31
Voc	48.92 (23.90)	4, 93	55.84 (23.65)	2, 95
WR	92.51 (15.08)	59, 116	93.80 (15.03)	60, 166
nWR	65.10 (10.98)	38, 89	64.15 (11.98)	38, 86

Note. M = accuracy proportion matching trials, MM = accuracy proportion mismatching trials, D' = sensitivity (d-prime), IQ = nonverbal IQ, Voc = vocabulary, WR = word reading, nWR = nonword reading, SS = spoken training-spoken test, WW = written training-written test.

Table 6.9: Model output Experiment 3.

D'	$F(1,121) = 2.53$, $p = .12$
IQ	$F(1,121) = 2.06$, $p = .15$
Voc	$F(1,121) = 2.60$, $p = .11$
WR	$F(1,121) = 0.22$, $p = .64$
nWR	$F(1,121) = 0.21$, $p = .65$

Note. D' = sensitivity (d-prime), IQ = nonverbal IQ, Voc = vocabulary, WR = word reading, nWR = nonword reading.

6.4.3 Conclusions Experiment 3

In spite of having a much larger sample than in Experiment 1, we did not find a modality effect on word learning efficacy. Thus, it is likely that the effect of modality on novel word learning is either non-existent or so small that its meaning is negligible.

6.5 Cross-experimental analyses

Traditionally, psycholinguistic research has focused on examining general linguistic processes that are shared among all humans. Striking is, however, that despite these general processes there tends to be large variation in humans linguistic ability. This notion has led to more interest in individual differences in humans' linguistic and cognitive skills (Dąbrowska, 2018; Kidd et al., 2018). To investigate the influence of various language and cognitive skills (nonverbal IQ, vocabulary size, word- and nonword reading) on word learning, several additional analyses were performed on the pooled data of all three experiments (for Experiment 2 only the data of the test session on Day 1 were used). Table 6.10 depicts the correlations between d-prime and the individual differences measures.

To investigate the effect of nonverbal IQ, vocabulary size, word- and nonword reading further, a linear model was run with log-linear corrected D' as the dependent variable and the four individual measures and experiment number (Experiment 1 being the reference level) as independent variables. Model output is displayed in Table 6.11. Nonverbal intelligence contributed significantly to word learning (D'), as did vocabulary size such that participants with higher intelligence and larger vocabularies learned more words.

Table 6.10: Correlations between sensitivity (d-prime) and the individual differences measures across all experiments.

	D'	Nonverbal IQ	Vocabulary	Word reading	Nonword reading
D'	1				
Nonverbal IQ	.25**	1			
Vocabulary	.24**	.27**	1		
Word reading	.08	.05	.16*	1	
Nonword reading	.16*	.04	.23**	.54**	1

Note. * p < .05, ** p < .001.

Table 6.11: Model output from the analysis run on the combined data from all experiments.

	Estimate	SE	t-value	p-value
Intercept	-0.67	0.43	-1.55	0.12
Experiment 2	-0.12	0.15	-0.83	0.41
Experiment 3	-0.19	0.13	-1.52	0.13
Nonverbal IQ	**0.04**	**0.01**	**3.33**	**0.001**
Vocabulary size	**0.01**	**0.002**	**2.37**	**0.02**
Word reading	-0.001	0.004	-0.20	0.84
Nonword reading	0.01	0.01	1.72	0.09

Note. Significant effects are displayed in bold.

6.6 General discussion

Novel words can be learned through encounters in at least two distinct modalities: either by listening to speech or by reading written text. Several theoretical frameworks make contradicting predictions with regard to the modality effect in novel word learning. Also, there is contradicting experimental evidence reporting both a spoken and written learning benefit, and some studies finding no difference at all. These studies differ substantially in methodology, which complicates comparison and interpretation of their findings. Therefore, the present study aimed to investigate modality effects on novel word learning in a new, controlled experimental setting.

In this new paradigm, we used an implicit training task where participants had to learn 24 Dutch-like, fully transparent pseudowords, presented in either the written or spoken modality, and an accompanying meaning in the form of a picture of a non-existing object. Exposure to the written words was equal to the duration of the recordings of their spoken counterparts (M = 860 ms). After a 20-minute period of consolidation in the form of a purely visual task, with no orthographical or phonological input, participants were tested on their knowledge of the learned word forms and meanings. In this matching task, participants had to decide whether a picture matched with a previously learned word or not. Here, word forms were presented in either the written or spoken modality, depending on which condition the participant was assigned to, in order to control for Tulving and Thomson (1973)'s encoding specificity principle.

In Experiment 1, we found no main effect of training modality. An interaction with test modality indicated that participants in the written training and spoken test condition were more accurate on the matching test than participants in the other groups. However, to our knowledge, there is no experimental work that corroborates this finding, nor theoretical frameworks that could explain this result.

In Experiment 2, we aimed to investigate the modality effect with regard to an important aspect of many theories of word learning, namely consolidation. Many studies report the necessity of at least 24 hours of consolidation for successful integration of a novel lexical entry (Clay et al., 2007; Dumay & Gaskell, 2007; Tham et al., 2015; van der Ven et al., 2015). Therefore, we augmented our paradigm with two additional online test sessions a day and a week after the initial learning and test phase. We did not find evidence for a modality effect during any of the test session. This, and several null-effects reported in previous studies (Dean et al., 1988; Nelson et al., 2005) led us to believe that the effect of training modality is likely to be small and requires a large sample size to be detected.

For Experiment 3 we therefore quadrupled the number of participants per condition, testing n = 123 participants in total. Since we were mainly interested in the modality effect during the training phase rather than the test phase, we omitted the test modality manipulation. This yielded two conditions: participants were either trained and tested in the written modality, or trained and tested in the spoken modality. Experiment 3 did not produce any evidence for an effect of modality.

Based on the results of these three experiments, we can assume that there is no meaningful difference regarding the efficiency with which spoken and written words are learned. The presence of the learning benefit for the written training – spoken test group in Experiment 1 may have been the result of random noise. Perhaps by coincidence participants in this group were better learners in general, had higher memory capacity, were more familiar to the word learning paradigms or behavioural experiments in general, or were benefitted by any other characteristic we could not control for.

One important aspect of our word learning paradigm was the modality manipulation of the matching test to control for Tulving and Thomson (1973)'s encoding specificity principle, which states that learning is most efficient if the conditions during the test match the conditions during the training phase. Unexpectedly, our results did not produce evidence for this principle, since the participants whose training and test occurred in the same modality did not show superior performance. Instead, participants in the cross-modal conditions were consistently able to recognize words presented in a modality in which they had never encounter this word before. This suggests that participants were able to form crossmodal representations of adequate quality to recognize a novel word in the other modality without ever having encountered the words in this modality. Theories of word reading (Coltheart et al., 2001) and reading acquisition (Ehri, 1995; Shankweiler, 1999; Share, 1995) indeed describe how readers can create phonological representations of words encountered in the written modality through recoding. Previous literature has reported that this recoding may be not as efficient in the spoken modality, resulting in less precise or incomplete orthographic representations of words encountered in the spoken form (Johnston et al., 2004). Our results, however, may suggest that crossmodal recoding may be equally efficient in both the written and spoken modality, at least for highly literate adults with a transparent native language.

The conclusion that there is no meaningful difference in the efficiency of spoken and written word learning is not necessarily at odds with theoretical frameworks that predicted learning to be more efficient in either the written or spoken modality. To reiterate, the statistical learning frameworks (Frost et al., 2015) predicted a spoken benefit as a result of auditory cortices being more efficient in processing sequential information, such as new words. Theories word reading

(Coltheart et al., 2001) and reading acquisition (Ehri, 1995; Shankweiler, 1999; Share, 1995) predicted a written learning benefit as a result of more efficient recoding in the written modality, allowing for the storage of high quality phonological as well as orthographic representations. Aside from these theoretical accounts, there are other aspects of these modalities that, hypothetically, could yield a learning benefit. For example, the less fleeting and less obtrusive nature of the written modality is thought to result in less continuous attention requirements, and leads to greater availability of memory and attentional resources for learning (van der Ven et al., 2015). Moreover, as addressed in Experiment 2, information processing occurs faster in the written modality, which allows for more exposure and thus more learning opportunities. Finally, some researchers argue that the spoken modality is the evolutionary and developmental primary modality and may therefore process information more efficient in general (Bakker et al., 2014).

Our results suggests that none of these advantages result in an apparent learning benefit in one modality over the other. Learning in each modality comes with distinct advantages and disadvantages and it may be the case that these advantages and disadvantages balance each other out. For example, even though recoding may not be as efficient in the spoken modality compared to the written modality, and speech being more obtrusive than written text, the sequential processing advantage and the evolutionary advantage in the spoken modality may help overcome these disadvantages. Thus, the modality-specific advantages described in the literature may be very much present, but none of these advantages may be strong enough to provide a learning benefit in one modality over the other.

The findings of our study provide an interesting insight into the development of modality-specific learning mechanisms, as the lack of a modality effect suggests that learning is equally efficient in the written and spoken modality. Particularly for the written modality, this is a remarkable conclusion, since the invention of writing is only 5500 years old. Even though written information intake is, in terms of evolution, a relatively novel feature of the human information processing system, our findings indicate that within approximately 250 human generations, the brain has evolved to optimize information processing in the written modality.

Evidence even suggests that humans have not only evolved to process written materials relatively fast in the timescale of evolution, but that the emergence of written processing systems also influences spoken language processing. Speech processing areas are activated during reading, and phonological processing areas show increased activation for spoken stimuli in literate compared to illiterate individuals (Dehaene et al., 2010). Moreover, enhanced literacy seems to sharpen phonological representations (Huettig & Pickering, 2019), which seems to improve phonological and speech processing (Cheung & Chen, 2004; Morais et al., 1986).

From an individual differences perspective, our experiments showed that both nonverbal IQ and vocabulary contribute to word learning in adults. With regard to nonverbal IQ, increased speed of processing (Fernald et al., 2006; Marchman & Fernald, 2008; McMurray, Horst, & Samuelson, 2012) and working memory capacity (Edwards & Anderson, 2014; Hansson, Forsberg, Lofqvist, Maki-Torkko, & Sahlen, 2004; Lukács, Racsmány, & Pléh, 2000) are likely candidates that may improve word learning. Faster processing allows individuals to allocate more time and cognitive resources to the next piece of information (Fernald et al., 2006). Also, larger working memory capacity may improve word learning efficiency by reducing the number of exposures necessary to retain the word (Leonard et al., 2007).

Regarding the relationship between vocabulary and word learning, many usage-based approaches assume that large vocabularies are characterized by precise and stable lexical representations as a result of extensive language experience (e.g., Dąbrowska, 2012; Lieven, 2016). These high quality lexical representations ameliorate lexical integration of novel words (Sailor, 2013; Storkel, Bontempo, & Pak, 2014) and facilitate activation of the correct lexical item by minimizing activation of competitors (Andrews & Hersch, 2010). Moreover, sharpened lexical representations enhance novel word learning by improving language processing speed, consequently allowing for faster and more efficient novel word processing and learning (Fernald et al., 2006; Hurtado et al., 2008; Jones & Rowland, 2017; Weisleder & Fernald, 2013).

The findings of our study open up new avenues for future research. First, an investigation of the modality effect in populations different from the population in our study is warranted. Our sample consisted of university students, which is a highly literate population familiarized with learning from written materials. It may be possible that modality does affect word learning in individuals with lower reading abilities and individuals who have less experience with learning from written instruction. Related to this, our study yielded a null-effect of modality using stimulus materials created with the aim to be as transparent as possible. Our finding may therefore not generalize to more opaque languages, where grapheme-phoneme correspondences are less reliable.

Another avenue for future research pertains to the development of written and spoken word learning. Children are able to learn from spoken information from birth, but learning in the written modality only commences after the child learns to read, typically around the age of six. Our results indicate that highly literate adults, learn equally efficient from written and spoken learning materials. The question remains how written word learning develops and seemingly 'catches-up' regarding their efficiency compared to spoken word learning. Moreover, it is unclear what level of reading is required for written learning mechanisms to become equally efficient as their spoken counterpart. Investigating these topics may give new insights into the development of children's word learning and how to mediate when this development goes awry.

With our aim to overcome several methodological issues present in previous literature, our learning paradigm was heavily controlled. These controls, although necessary to understand the fundamentals of the modality effect in word learning, caused our learning environment to deviate largely from a normal learning environment. Therefore, future research could move to a more ecologically valid learning environment. Examples may be studying word types other than concrete nouns or creating a richer learning context by using worded meaning, sentences or text where participants have to infer a word's meaning. Insights accumulated using such ecologically valid learning environments may result in a multitude of practical implications for general education.

To conclude, the present study suggested that there is no difference in the efficiency of written and spoken learning mechanisms. Previous literature has described various modality-specific advantages in the written and spoken modality, but it seems that none of these advantages results in an absolute learning benefit in one modality over the other. Instead, it seems that learning efficiency is mostly influenced by individual differences in linguistic and cognitive abilities. Our results suggest that, although humans typically begin as aural-only word learners, their word learning ultimately develops to be equally efficient in the spoken and written modality.

6.7 Appendix I - Pseudowords

Pseudoword	Audio duration (ms)	Phonemes	Graphemes	Letters
euhom	796	4	4	5
asheer	664	5	5	6
rollang	682	5	5	7
juslerf	822	7	7	7
sarweek	805	6	6	7
rafjaak	778	6	6	7
wolwier	926	6	6	7
faamruil	960	6	6	8
luiswaas	986	6	6	8
waamhies	921	6	6	8
roofkoor	953	6	6	8
mesrier	936	6	6	7
mimkras	990	7	7	7
slomkiel	990	7	7	8
slekwoes	993	7	7	8
flersmal	877	8	8	8
walmlaf	830	7	7	7
worshiel	888	7	7	8
kusfus	841	6	6	6
warfres	901	7	7	7
makles	858	6	6	6
rolsjuk	761	7	7	7
klaarleek	843	7	7	9
krokwok	721	7	7	7

7 General discussion

With the invention of writing came a new modality in which language could be used. This led to both a new processing pipeline, specialized in processing written language, and a new type of representation in which language could be mentally represented, namely orthography. This additional representational type allows literate individuals to represent the same linguistic content in two ways: phonologically and orthographically. The Lexical Quality Hypothesis (LQH, Perfetti, 2007) assumes that different types of representations of the same lexical entry can differ in their lexical quality (i.e., the level of precision and completeness). Moreover, it assumes that high quality representations are processed more efficiently than low quality representations.

The first main question of this dissertation was derived from the Lexical Quality Hypothesis: is there a difference in processing efficiency of the orthographic and phonological representation of the same word? The second main question was connected to the observation that individuals differ with respect to their language abilities and language experience. Language experience is assumed to sharpen lexical representations (Brysbaert, Lagrou, & Stevens, 2016; Castles, 1999; Castles et al., 2007; Diependaele et al., 2013), which in turn improves language processing efficiency. This mechanism may be stronger for the written modality, as written language is syntactically more complex and more varied in its vocabulary than day-to-day, conversational spoken language. Written language therefore provides a richer context in which lexical representations can be specified. Moreover, one can take in more written than spoken language within the same time frame, as processing written language is generally faster than processing spoken language. While spoken language unfolds over time, written language, at least at the word level, is available instantaneously. Thus, the second main question of this dissertation was: are individuals who are more experienced with written language more efficient at word-level language processing compared to individuals who are less experienced?

7.1 Summary of findings

Chapter 2 examined how word recognition is influenced by presentation modality and experience with written language. Participants, sampled from the university student population, took part in a lexical decision task. Words were presented in the written, spoken or audio-visual modality (auditory and visual presentation coincided). The words differed in difficulty, with some words being more and others being less known to the general population. Participants' experience with written language was assessed using a receptive vocabulary task and a task measuring print exposure (Author Recognition Test). No evidence was observed for a modality effect on word recognition accuracy: words presented in the written, spoken or audio-visual modality were recognized equally well. Experience with written language, however, did influence word recognition, as experienced individuals recognized difficult words more accurately than less experienced individuals. The results obtained in this chapter suggest that in a rather homogeneous group of individuals (in terms of their experience with written language), presentation modality does not affect word recognition accuracy. However, despite the homogeneity of the participant sample, word recognition accuracy was observed to vary as a function of written language experience.

Chapter 3 continued studying the effect of presentation modality and experience with written language on word recognition. In the previous chapter, there was no evidence for an effect of presentation modality, but findings suggested that even in the relative homogeneous population of university students, variation with regard to experience with written language influenced word recognition accuracy. Chapter 3 examined a more heterogeneous sample with respect to experience with written language, as a sample with a larger range of reading experience may increase sensitivity to detect effects of reading experience and its interaction with presentation modality. Participants in this study were therefore sampled from both the vocational and university student population. Recruiting participants with a more diverse educational background would broaden the sample's range of experience with written language. The lexical decision task was slightly adapted for the new, more diverse group of participants. Moreover, an additional measure of experience with written language was administered in the form of a reading behaviour questionnaire. In addition to accuracy, response times were recorded to investigate not only the success but also speed of word recognition.

The results from the preceding chapter were replicated in that there was no difference across the three presentation modalities conditions in terms of recognition accuracy. Experience with written language did influence word recognition accuracy, as experienced readers were more accurately recognizing words of high to medium difficulty. With regard to reaction times, a modality effect was observed. Responses to written and audio-visually presented words were faster than responses to spoken words. This likely reflects the fact that processing written information is faster than processing spoken information, because spoken information unfolds over time whereas written information is present all at once. Interestingly, an interaction was observed between presentation modality and word difficulty (i.e., the degree to which a word was known to the general population). Easy, more well-known words were generally recognized faster than more difficult words, but this effect was larger in the written compared to spoken modality. In addition, experience with written language also influenced recognition speed: easy, more well-known words in particular were recognized faster by experienced individuals compared to individuals with less experience with written language. In sum, the findings of this chapter indicate that word recognition accuracy is not influenced by presentation modality, but that the speed with which these words are recognized is influenced by presentation modality. With regard to experience with written language, both the success of word recognition and the speed of accessing lexical representations are improved in individuals who are experienced with written language.

Chapter 4 explored the relationship between written language experience and spoken language processing in more depth. Specifically, the study aimed to improve our understanding of the degree to which variation in individuals' experience with written language explained variation in their spoken language processing, while accounting for individual differences in several general cognitive abilities that are likely to be involved in spoken language processing (nonverbal intelligence, processing speed). Spoken word comprehension and word production were examined. In the analyses, a latent-variable rather than single-variable approach was adopted in order to obtain a purer measurement of the broad, multifaceted constructs of word production, spoken word comprehension, experience with written language, and processing speed. Only nonverbal intelligence was assessed with a single task. Results showed that experience with written language contributed to word production and word comprehension, even when accounting for individual differences in nonverbal intelligence and processing speed. The findings of this chapter highlight the profound influence of experience with written language on spoken word processing.

Chapter 5 presented a systematic review that explored the state of the current literature with regard to the relationship between written language experience and both written and spoken word recognition. Literature was discussed separately for different types of lexical representations (semantic, phonology, orthography, syntax) and different levels of representation (lexical, morpheme/syllable, sublexical). Moreover, studies conducted with adults as well as studies conducted with children were included in the review to examine the scope of the literature with regard to the development of word recognition. The current literature provides clear evidence that word recognition is facilitated by experience with written language as a result of refining the quality of semantic, phonological and orthographic representations and improving processing efficiency at the lexical, morpheme/syllable and sublexical level. Patterns of findings were very similar for spoken and written word recognition, indicating that the facilitatory effects of written language experience transfer to spoken language processing. The literature review identified several gaps in the literature, such as the relationship between written language experience and spoken word recognition in adults, and the influence of written language experience on word-level syntactic processing. Moreover, the current literature does not provide conclusive evidence as to at what stages written language may influence word recognition, and how the relationship between word recognition and written language experience develops as children age.

Chapter 6 examined how learning novel words is influenced by presentation modality. Participants, sampled from the university student population, took part in a newly designed, implicit and fast-paced word learning task where they had to learn 24 associations between Dutch-like pseudowords and pictures of non-existing objects. After a 20-minute period of consolidation, participants were tested on their learning using a matching task where they had to decide whether a word and picture matched according to what they had learned before, or not. In all three experiments, no evidence was observed for an effect of presentation modality on word learning, as accuracy on the matching test did not differ for participants who learned the words in their written form compared to participants who learned the words in the spoken form. Interestingly, participants were able to recognize the words in the spoken or written modality, even if they had only encountered the word in the other modality during the training phase. The findings from this study indicate that adult speakers of Dutch, who are relatively experienced with written language compared to the general population, 1) show no learning advantage for words presented in one modality over the other, and 2) are able to create crossmodal representations of sufficient quality that allow for successful recognition of the word in the other modality.

7.2 Influence of presentation modality on word learning and word processing

This dissertation sought to explore how presentation modality influences word recognition and word learning. With regard to word recognition, Chapters 2 and 3 provided evidence that the accuracy with which words are recognized is not influenced by the modality in which words are presented: written and spoken words were recognized with the same accuracy. However, Chapter 3 showed that presentation modality does influence the speed with which words are recognized. Easy, common words were recognized faster in the written than in the spoken modality.

With regard to word learning, Chapter 6 showed across three experiments that novel word learning was not influenced by presentation modality. Note that in the paradigm used in this chapter, exposure duration between the two modalities was either equal, where written and spoken words were presented for the same amount of time, or equated, where written words were presented for a shorter duration to account for the fact that reading is faster than listening to speech. Exposure duration was controlled in this way, because the written compared to spoken modality has two characteristics that allow for relatively more exposure within a given time frame. First, processing written information is faster than processing spoken information, which allows for relatively more information intake. Second, unlike spoken information, written information does not disappear as time progresses, which allows the reader to re-expose themselves multiple times to the same information. Chapter 6 showed that when controlling for these inherent differences related to exposure between the two modalities, there is no difference in learning written or spoken words, suggesting that the learning mechanisms that operate on written or spoken information are equally efficient.

The combined evidence from these three chapters paint an interesting picture. It seems that in highly literate adults, the initial creation of a new lexical representation is not affected by presentation modality (when accounting for the inherent difference in exposure time between the two modalities). However, as a word becomes integrated in the mental lexicon over many encounters, the quality of orthographic and phonological lexical representations may diverge, resulting in an effect of presentation modality on the speed with which (commonly known) words are recognized. Since in these word recognition studies, several word-level aspects known to induce speed benefits, such as word length or word duration, were controlled for, these factors could not explain this written speed benefit. It seems that there are other forces at play that may particularly improve the sharpening of lexical representations in the written modality.

First, written language provides a rich context that may help sharpening semantic lexical representations. Written text, in comparison to spoken text, is characterized by a larger, more diverse vocabulary (Cunningham & Stanovich, 1998; Hayes & Ahrens, 1988), a higher degree of structure (Kroll, 1977) and more complex grammar phrases (Biber, 1991; Roland et al., 2007; Scott, 2008). There is evidence that suggests that a richer context allows for the creation of more precisely defined lexical representations. For example, words learned in a rich context versus words learned in a less rich and uninformative context show earlier and stronger signs of semantic integration (Frishkoff, Perfetti, & Collins-Thompson, 2010; Mestres-Missé, Rodriguez-Fornells, & Münte, 2007). In these studies, participants learned novel words in sentence contexts. For some words, the meaning of the novel word could be easily derived from the sentence, but for other words, the sentence did not provide meaningful information. Mestres-Missé et al. (2007) found that ERPs (N400, indicative of semantic surprisal) to the novel words learned in a rich context were similar to ERPs to known words, whereas response words learned in an uninformative context were more like responses to unknown words. Frishkoff et al. (2010) found a stronger ERP response indicative of difficulty of retrieval ($N300_{FT7}$), for words learned in an uninformative context compared to words learned in a rich, informative context, suggesting that words learned in an uninformative context were more difficult to retrieve than words learned in an informative context. Rich, informative contexts are thought to particularly help sharpen semantic representations: by encountering the word in such an informative context, one can specify the representation by adding very subtle semantic details that are inferred from the context (Frishkoff, Collins-Thompson, Perfetti, & Callan, 2008).

An explanation as to why orthographic representations may be more strongly sharpened in the written modality is that written language, particularly in printed form, is very regular. Due to font standardizations, letter tokens within a font are identical and easily recognizable across fonts. Visually speaking, written forms are therefore easy to decipher and to predict based on the reader's previous experience with this font (Huettig & Pickering, 2019). On the other hand, speech is prone to a lot of variance across and even within speakers (see J. H. L. Hansen & Bořil, 2018, for an overview of causes of intra- and inter-speaker variance). Moreover, speech processing is susceptible to environmental noise, whereas such noise is less present in written language (except for perhaps the occasional typo or ink smudge). This variance and increased noise may result in processing being more effortful and/or prone to errors in the spoken but not written modality, and may explain why subsequent sharpening of orthographical representations is more efficient in the written modality compared to sharpening phonological representations in the spoken modality.

Another explanation for more efficient sharpening of lexical representations in the written compared to spoken modality is related to a difference in the degree of crossmodal processing in the written and spoken modality. Crossmodal recoding is a mechanism prominently present in processing written language, being an integral part of many theories of word reading (Coltheart et al., 2001) and reading acquisition (Ehri, 1995; Shankweiler, 1999; Share, 1995). Upon encountering a written word, readers are thought to mentally recode and automatically activate the crossmodal, in this case, phonological, representation. Although connectionist theories of word recognition (Grainger & Ferrand, 1994; McClelland & Rumelhart, 1981; Seidenberg & McClelland, 1989) assume that such crossmodal automatic activation is present for both written and spoken word recognition, evidence for the same mechanism at work during spoken word recognition is sparse. For example, although crossmodal activation seems to take place in the spoken modality, it may not be automatic (Damian & Bowers, 2010; Pattamadilok, Kolinsky, Ventura, Radeau, & Morais, 2007). Thus, these crossmodal recoding/activation processes may help sharpen crossmodal representations in the written modality, but less so in the spoken modality. Note that language transparency might determine the efficiency of crossmodal recoding in the spoken modality. It seems that in opaque languages, such as English, crossmodal recoding results in only roughly rather than precisely specified crossmodal representations when learning novel words in the spoken modality (Johnston et al., 2004). In contrast, Chapter 6 showed that highly proficient language users were able to crossmodally recode pseudowords learned in the spoken modality, so that they could efficiently recognize words in their written form during a later test, even though they had never seen the written forms before. The pseudowords used in that study were fully transparent, as opaque grapheme-phoneme correspondences were avoided when creating the words. It may be the case that in transparent languages, crossmodal recoding may aid in sharpening of lexical representations in the spoken modality, but not or to a lesser extent in opaque languages.

In conclusion, while initial word learning was found to be unaffected by modality, speed of word recognition was modulated by modality, as easier, well-known words were recognized faster in the written compared to spoken modality. This seems to suggest that word learning mechanisms are not influenced by presentation modality, as representations of adequate quality are created in both modalities, but that through multiple, modality-specific encounters, different representations of the same word may start differ to in their lexical quality. Potential factors that may result in more specified representations, resulting in a recognition speed benefit in the written modality, may be that written language provides a very rich context, in terms of grammar and vocabulary, in which sematic representations are more likely to be specified. Moreover, writing is more invariant and less noisy than speech, allowing for more effortless processing and sharpening of orthographic representations in the written compared to phonological representations in the spoken modality. Finally, language processing in the written modality is characterized by automatic and efficient crossmodal recoding/activation, and it is thought that this mechanism is less efficient and automatic in the spoken modality, particularly in opaque languages.

7.3 Influence of experience with written language on language processing

This dissertation also aimed to examine how experience with written language influences the efficiency (i.e., accuracy and speed) of word-level language processing. Chapter 2, 3 and the systematic review in Chapter 5 showed that word recognition accuracy and speed were influenced by experience with written language, with experienced individuals recognizing words more accurately and faster. Chapters 2 and 3 showed that experienced individuals were especially more accurate at recognizing difficult, less common words and faster at recognizing easy, more well-known words. These findings suggest that experience with written language increases word knowledge and the size of the mental lexicon, resulting in a higher word recognition accuracy (Hurtado et al., 2008; Monaghan et al., 2017). At the same time, experience with written language seems to increase processing speed of words that have high quality lexical representations (i.e., commonly encountered words).

Theories such as the lexical entrenchment hypothesis (Brysbaert, Lagrou, & Stevens, 2016; Diependaele et al., 2013) and lexical tuning hypothesis (Castles, 1999; Castles et al., 2007) have indeed argued that language experience fine-tunes lexical representations, which improves the processing mechanisms that operate on these lexical representations. According to the lexical tuning hypothesis, lexical representations become more precisely tuned to ensure efficient word recognition in an ever-expanding mental lexicon. The lexical entrenchment hypothesis argues that experience with language lowers the activation threshold of lexical representations, so that words are faster and easier accessed. Moreover, it argues that experience with language increases the quality of lexical representations so that there is less competition of lexical competitors during the word recognition process, allowing for swifter retrieval of the correct word. However, in these two accounts, the precise mechanisms that cause lexical representations to be tuned or sharpened with increased language experience are not well-specified. Also, these accounts are not specific to the written modality. One could argue that the sharpening mechanisms specific to the written modality, as described in the previous paragraphs, may increase in strength as individuals are more exposed to language in the written modality. Experienced written language users encounter more (written) language embedded in a rich context, allowing them to sharpen their semantic lexical representations even more. Avid, compared to less avid readers encounter written language more often, which is less invariant and less noisy than speech, thereby efficiently sharpening their orthographic representations. Finally, experienced readers will crossmodally activate/recode language while reading, thereby also sharpening their crossmodal representations. In other words, the benefits associated with experience with written language will amplify as individuals expose themselves more often to written language.

Moreover, Chapter 4 demonstrated that, by accounting for the influence of general cognitive skills in the analyses, the influence of experience with written language on language processing ability exists independent of the influence of general cognitive abilities. This finding is in line with emergentist approaches to language (Dąbrowska, 2018; Lieven, 2016), which argue that differences between individuals with respect to their language abilities are not only the result of variation with respect to individuals' intrinsic neurocognitive learning mechanisms, but also variation in individuals' experience with language.

7.4 Overlap between word comprehension and word production

An intriguing finding from Chapter 4 was that experience with written language not only contributes to word recognition/comprehension skills, but also to word production skills. The beneficial effects associated with written language experience thus seems to transfer from the domain of word comprehension to the domain of word production. An explanation for this observation may be that word comprehension and word production share some of the cognitive architecture for language. Indeed, several researchers proposed an integrated cognitive architecture for language where comprehension and production are two different 'language tasks' involving one, unitary language system that consists of shared representations and processing operations (Chater, McCauley, & Christiansen, 2015; McQueen & Meyer, 2019). Naturally, each 'language task' is distinct and as such some processing mechanisms may be exclusively used for one language task (e.g., planning an articulatory score during speech production), whereas other processes are shared (e.g., activating lexical representations). Crucially, comprehension and production are assumed to draw upon the same set of linguistic representations and processing mechanisms, rather than each having their own separate cognitive architecture. The findings of Chapter 4 indeed provide evidence that, due to this shared cognitive architecture, the facilitatory effects of experience with (written) language on word recognition/comprehension transfer to the 'language task' word production.

7.5 Recommendations for future research

The findings reported in this dissertation have addressed some of the questions related to the influence of presentation modality and experience with written language on language processing. However, as is only natural to the scientific method, answers to questions result in new questions. There are many avenues for future research to build upon the present work.

First, factors that improve sharpening of lexical representations in the written compared to spoken modality may be further explored. Spoken, everyday conversational language is thought to be different from language used in books and other written texts. Of course, not all spoken language is conversational: lectures and speeches may be more structured than everyday speech, and contain a larger variety with regard to vocabulary. Audiobooks are also examples of written language transposed to the spoken modality. If the beneficial effect of written language experience on language processing efficiency were due to the different use of written and spoken language, then one would expect that these benefits would also occur for spoken language that mimics written language in content and structure. At the same time, with the invention of email and chat-apps, conversational spoken language is transposed to the written modality. Since this informal written language likely lacks the complex grammatical structures found in formal written texts, and likely has a smaller breath of vocabulary, one would expect that the benefits associated with written language experience does not extend to this type of written language. Thus, the influence of listening to audiobooks as compared to written books, and the use of written conversational language or ‘chat speak’ on language processing and ability may be a fruitful avenue for further research to uncover the precise mechanisms that fine-tune lexical representations in the written/spoken modality.

In this dissertation it is argued that crossmodal recoding or automatic crossmodal activation during written and spoken language processing may be another factor that improves sharpening of lexical representations in the written compared to spoken modality. Evidence suggests that crossmodal activation is more automatic and efficient in the written compared to spoken modality. However, studies that directly compared crossmodal activation in the written and spoken modality are scarce. In order to make such direct comparisons, studies could investigate the orthographic and phonological consistency effects in spoken and written word recognition. These effects refer to the fact that words containing elements that can be spelled in multiple ways (/ɛr/ in care-hair-wear) or elements that can be pronounced in multiple ways (/rid/ and /rɛd/ for “read”) are recognized more slowly and/or less accurately than words containing elements with consistent grapheme-phoneme correspondences. It is thought that the recognition disadvantage is the result of increased competition due to the activation of the crossmodal representations. Directly comparing consistency effects in the spoken and written modality may improve our understanding of the degree of crossmodal activation in each modality.

In addition, neuropsychological methods can also be used to investigate crossmodal recoding/activation in the written and spoken modality. For example, fMRI may be used to see whether areas traditionally associated with spoken or written language processing are active when stimuli are presented in the other modality (e.g., Chiarello, Vaden Jr., & Eckert, 2018). ERPs can be examined to see whether ERP components thought to be related to phonological or orthographic processing are present when encountering stimuli in the other modality (e.g., Perre, Pattamadilok, Montant, & Ziegler, 2009). Finally, TMS at certain areas in the brain, known to be involved in modality-specific processing, can be used to interfere with either phonological or orthographic processing after being presented with written or spoken language (e.g., Pattamadilok, Knierim, Duncan, & Devlin, 2010). By exploring the degree of crossmodal recoding/activation in the written and spoken modality, we may uncover whether differences in processing mechanism result in different degrees of sharpening crossmodal lexical representations in the written compared to spoken modality.

A third topic that may be explored further pertains to language modalities other than speech or writing. Sign language is visual, like writing, but, similar to speech, bound to time and place. Future studies could compare word processing and word learning for written, spoken and sign language to uncover the connections between orthographic, phonological and signed forms in the mental lexicon. Moreover, sign language users show large variation with regard to their experience with sign language, with some users being signers from birth, whereas other users may have only learned sign language relatively recently in adulthood. This allows the influence of experience with (sign) language on (sign) language processing to be studied with much nuances. Also, sign language users tend to vary in the extent to which they have been exposed to written and spoken language, which allows for careful examination of crossmodal influences on language processing.

A fourth avenue for further research would be to examine the effect of written language experience on language beyond the word-level. This dissertation showed that experience with written language allowed for faster and more accurate *word* recognition. The question is whether this also transfers to larger units of language. There is some evidence suggesting that experienced written language users are better at predicting upcoming language in sentence contexts (Favier et al., 2020; Huettig & Pickering, 2019). Moreover, reading-related behaviours are associated with improving written text comprehension (de Naeghel, van Keer, Vansteenkiste, & Rosseel, 2012; Torppa et al., 2020). Evidence suggests that children who are more experienced with language (as approximated with a vocabulary test) show increased logical reasoning (Segers & Verhoeven, 2016) and inference making abilities (Cain & Oakhill, 2014; Daugaard, Cain, & Elbro, 2017; Language & Reading Research Consortium, Currie, & Muijselaar, 2019), which are fundamental skills required for the construction of a coherent mental model of a text. Given that written and spoken text comprehension abilities partly overlap (Wolf, Muijselaar, Boonstra, & de Bree, 2019), it may be the case that the influence of reading-related behaviours also transfer to spoken language comprehension. Thus, more research beyond the word level will improve our understanding of the influence of written language experience on higher level language skills.

A final future direction may be to examine the influence of experience with written language on language processing over the course of development. Children learn to read around the age of six (Seymour et al., 2003; Vaessen et al., 2010) and from that moment onwards, they develop skills and abilities related to reading. Not only do reading skills tend to be relatively stable over development (Hulslander, Olson, Willcutt, & Wadsworth, 2010), also reading habits do not tend to differ much as a child grows up (Wigfield & Guthrie, 1997). Thus, children who are good readers and read a lot at an early age, are likely still proficient and frequent readers at a later age. The relationship between written language experience and language processing abilities therefore seems to take root from the moment a child learns to read, and may develop in a very stable manner. It may therefore be worthwhile to examine not only which internal and external factors place a child on a certain developmental pathway, but also which interventions can be implemented if this pathway is not optimal for later outcomes.

7.6 Conclusion

Humans encounter and represent language in, amongst others, the written and spoken modality. This dissertation explored whether presentation modality and experience with written language influences language processing. With regard to presentation modality, the results suggest that the modality in which new words are learned does not influence the efficiency with which lexical representations are created, but that after multiple exposures, lexical quality of orthographic and phonological lexical representations may diverge, resulting in faster recognition of well-known words in the written compared to spoken modality. With regard to experience with written language, the present findings indicate that individuals who are more experienced with written language tend to process language faster and more accurately, which is likely due to the fact that extensive exposure improves the quality of lexical representations. This dissertation highlights how a seemingly simple invention such as a writing system, which is essentially only an arbitrary system to visualize language sounds, has a profound influence on human's cognitive architecture for language.

References

Acha, J., & Perea, M. (2008). The effects of length and transposed-letter similarity in lexical decision: Evidence with beginning, intermediate, and adult readers. *British Journal of Psychology*, *99*(2), 245-264. doi: 10.1348/000712607X224478

Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods*, *40*(1), 278-289. doi: 10.3758/brm.40.1.278

Allen, L., Cipielewski, J., & Stanovich, K. E. (1992). Multiple indicators of children's reading habits and attitudes: Construct validity and cognitive correlates. *Journal of Educational Psychology*, *84*(4), 489-503.

Alvarez, C. J., Carreiras, M., & De Vega, M. (2000). Syllable-frequency effect in visual word recognition: Evidence of sequential-type processing. *Psicológica*, *21*(2), 341-374.

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin and Review*, *4*(4), 439-461. doi: 10.3758/Bf03214334

Andrews, S. (2015). Individual differences among skilled readers: The role of lexical quality. In A. Pollatsek & R. Treiman (Eds.), *The Oxford handbook of reading* (p. 129-148). New York, NY, US: Oxford University Press.

Andrews, S., & Hersch, J. (2010). Lexical precision in skilled readers: Individual differences in masked neighbor priming. *Journal of Experimental Psychology: General*, *139*(2), 299-318. doi: 10.1037/a0018366

Andringa, S., & Godfroid, A. (2020). Sampling bias and the problem of generalizability in applied linguistics. *Annual Review of Applied Linguistics*, *40*, 134-142. doi: 10.1017/S0267190520000033

Andringa, S., Olsthoorn, N., van Beuningen, C., Schoonen, R., & Hulstijn, J. (2012). Determinants of success in native and non-native listening comprehension: An individual differences approach. *Language Learning*, *62*(s2), 49-78. doi: 10.1111/j.1467-9922.2012.00706.x

Arnett, J. J. (2016). *The neglected 95 percent: Why American psychology needs to become less American*. Washington, DC, US: American Psychological Association. doi: 10.1037/14805-008

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database (release 2): Linguistic Data Consortium*. Philadelphia, PA, USA: University of Pennsylvania.

Bakker, I., Takashima, A., van Hell, J. G., Janzen, G., & McQueen, J. M. (2014). Competition from unseen or unheard novel words: Lexical consolidation across modalities. *Journal of Memory and Language*, *73*, 116-130. doi: 10.1016/j.jml.2014.03.002

Bakker, I., Takashima, A., van Hell, J. G., Janzen, G., & McQueen, J. M. (2015a). Changes in theta and beta oscillations as signatures of novel word consolidation. *Journal of Cognitive Neuroscience*, *27*(7), 1286-1297. doi: 10.1162/jocn_a_00801

Bakker, I., Takashima, A., van Hell, J. G., Janzen, G., & McQueen, J. M. (2015b). Tracking lexical consolidation with erps: Lexical and semantic-priming effects on N400 and LPC responses to newly-learned words. *Neuropsychologia*, *79*(Pt A), 33-41. doi: 10.1016/j.neuropsychologia.2015.10.020

Bakker-Marshall, I., Takashima, A., Schoffelen, J. M., van Hell, J. G., Janzen, G., & McQueen, J. M. (2018). Theta-band oscillations in the middle temporal gyrus reflect novel word consolidation. *Journal of Cognitive Neuroscience*, *30*(5), 621-633. doi: 10.1162/jocn_a_01240

Balass, M., Nelson, J. R., & Perfetti, C. A. (2010). Word learning: An ERP investigation of word experience effects on recognition and word processing. *Contemporary Educational Psychology*, *35*(2), 126-140. doi: 10.1016/j.cedpsych.2010.04.001

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133*(2), 283-316. doi: 10.1037/0096-3445.133.2.283

Bane, M. (2008). Quantifying and measuring morphological complexity. In *Proceedings of the 26th West Coast Conference on Formal Linguistics* (p. 69-76). Cascadilla Proceedings Project.

Bates, D., Machler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1-48. doi: 10.18637/jss.v067.i01

Bent, T., Baese-Berk, M., Borrie, S. A., & McKee, M. (2016). Individual differences in the perception of regional, nonnative, and disordered speech varieties. *Journal of the Acoustical Society of America*, *140*(5), 3775-3786.

doi: 10.1121/1.4966677

Bergen, B., & Chan Lau, T. T. (2012). Writing direction affects how people map space onto time. *Frontiers in Psychology*, *3*(109). doi: 10.3389/fpsyg.2012.00109

Beyersmann, E., Castles, A., & Coltheart, M. (2012). Morphological processing during visual word recognition in developing readers: Evidence from masked priming. *Quarterly Journal of Experimental Psychology*, *65*(7), 1306-1326. doi: 10.1080/17470218.2012.656661

Beyersmann, E., Grainger, J., Casalis, S., & Ziegler, J. C. (2015). Effects of reading proficiency on embedded stem priming in primary school children. *Journal of Experimental Child Psychology*, *139*, 115-126. doi: 10.1016/j.jecp.2015.06.001

Beyersmann, E., Grainger, J., & Castles, A. (2019). Embedded stems as a bootstrapping mechanism for morphological parsing during reading development. *Journal of Experimental Child Psychology*, *182*, 196-210. doi: 10.1016/j.jecp.2019.01.010

Biber, D. (1991). *Variation across speech and writing*. Cambridge University Press.

Bock, K., & Levelt, W. J. M. (1994). *Language production: Grammatical encoding*. Academic Press.

Boersma, P., & Weenink, D. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5*(9/10), 341-345.

Bowers, J. S., Davis, C. J., & Hanley, D. A. (2005). Interfering neighbours: the impact of novel word learning on the identification of visually similar words. *Cognition*, *97*(3), B45-B54. doi: 10.1016/j.cognition.2005.02.002

Bramao, I., Mendonca, A., Faísca, L., Ingvar, M., Petersson, K. M., & Reis, A. (2007). The impact of reading and writing skills on a visuo-motor integration task: A comparison between illiterate and literate subjects. *Journal of the International Neuropsychological Society*, *13*(2), 359-364.

Brus, B. T., & Voeten, M. (1973). *Een-minuuttest, Vorm A en B, verantwoording en handleiding [One-minute-test, Form A and B, justification and manual]*. Nijmegen, Netherlands: Berkhout.

Brysbaert, M., Lagrou, E., & Stevens, M. (2016). Visual word recognition in a second language: A test of the lexical entrenchment hypothesis with lexical decision times. *Bilingualism: Language and Cognition*, *20*(3), 530-548. doi: 10.1017/S1366728916000353

Brysbaert, M., Mandera, P., McCormick, S. F., & Keuleers, E. (2019). Word prevalence norms for 62,000 English lemmas. *Behavior Research Methods*, *51*(2), 467-479.

Brysbaert, M., New, B., & Keuleers, E. (2012). Adding part-of-speech information to the SUBTLEX-US word frequencies. *Behavior Research Methods*, *44*(4), 991-997. doi: 10.3758/s13428-012-0190-4

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, *1*(1), 9. doi: 10.5334/joc.10

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant's age. *Frontiers in Psychology*, *7*(1116). doi: 10.3389/fpsyg.2016.01116

Brysbaert, M., Sui, L., Dirix, N., & Hintz, F. (2020). Dutch author recognition test. *Journal of Cognition*, *3*(1). doi: 10.5334/joc.95

Burani, C., Marcolini, S., & Stella, G. (2002). How early does morpholexical reading develop in readers of a shallow orthography? *Brain and Language*, *81*(1), 568-586. doi: 10.1006/brln.2001.2548

Burke, D., Linder, S., Hirsch, J., Dey, T., Kana, D., Ringenbach, S., ... Alberts, J. (2017). Characterizing information processing with a mobile device: Measurement of simple and choice reaction time. *Assessment*, *24*(7), 885-895. doi: 10.1177/1073191116633752

Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1-28. doi: 10.18637/jss.v080.i01

Cain, K., & Oakhill, J. (2014). Reading comprehension and vocabulary: Is vocabulary more important for some aspects of comprehension? *L'Année Psychologique*, *114*(4), 647-662. doi: 10.4074/s0003503314004035

Carreiras, M., Seghier, M. L., Baquero, S., Estévez, A., Lozano, A., Devlin, J. T., & Price, C. J. (2009). An anatomical signature for literacy. *Nature*, *461*(7266), 983-986. doi: 10.1038/nature08461

Castles, A. (1999). Neighbourhood effects on masked form priming in developing readers. *Language and Cognitive Processes*, *14*(2), 201-224. doi: 10.1080/016909699386347

Castles, A., Davis, C., Cavalot, P., & Forster, K. (2007). Tracking the acquisition of orthographic skills in developing readers: Masked priming effects. *Journal of Experimental Child Psychology*, *97*(3), 165-182. doi: 10.1016/j.jecp.2007.01.006

Castro-Caldas, A., Miranda, P. C., Carmo, I., Reis, A., Leote, F., Ribeiro, C., & Ducla-Soares, E. (1999). Influence of learning to read and write on the morphology of the corpus callosum. *European Journal of Neurology*, *6*(1), 23-28. doi: 10.1046/j.1468-1331.1999.610023.x

Chateau, D., & Jared, D. (2000). Exposure to print and word recognition process. *Memory and Cognition*, *28*(1), 143-153. doi: 10.3758/BF03211582

Chater, N., McCauley, S., & Christiansen, M. (2015). Language as skill: Intertwining comprehension and production. *Journal of Memory and Language*, *89*, 244-254. doi: 10.1016/j.jml.2015.11.004

Cheung, H., & Chen, H. C. (2004). Early orthographic experience modifies both phonological awareness and on-line speech processing. *Language and Cognitive Processes*, *19*(1), 1-28. doi: 10.1080/01690960344000071

Chiarello, C., Vaden Jr., K. I., & Eckert, M. A. (2018). Orthographic influence on spoken word identification: Behavioral and fMRI evidence. *Neuropsychologia*, *111*, 103-111.

Clay, F., Bowers, J. S., Davis, C. J., & Hanley, D. A. (2007). Teaching adults new words: The role of practice and consolidation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(5), 970-976. doi: 10.1037/0278-7393.33.5.970

Coady, J. A., & Evans, J. L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *International Journal of Language and Communication Disorders*, *43*(1), 1-40. doi: 10.1080/13682820601116485

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Abingdon, UK: Routledge.

Cohen, L., & Dehaene, S. (2004). Specialization within the ventral stream: the case for the visual word form area. *Neuroimage*, *22*(1), 466-476. doi: 10.1016/j.neuroimage.2003.12.049

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M.-A., & Michel, F. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, *123*(2), 291-307. doi: 10.1093/brain/123.2.291

Cohen-Shikora, E. R., & Balota, D. A. (2016). Visual word recognition across the adult lifespan. *Psychology and Aging*, *31*(5), 488-502. doi: 10.1037/pag0000100

Colombo, L., Sulpizio, S., & Peressotti, F. (2017). Serial mechanism in transposed letters effects: A developmental study. *Journal of Experimental Child Psychology*, *161*, 46-62. doi: 10.1016/j.jecp.2017.04.002

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204-256. doi: 10.1037/0033-295x.108.1.204

Connine, C. M., Mullennix, J., Shernoff, E., & Yelen, J. (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(6), 1084-1096. doi: 10.1037//0278-7393.16.6.1084

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(1), 24-39. doi: 10.1037/0278-7393.31.1.24

Conway, C. M., Pisoni, D. B., & Kronenberger, W. G. (2009). The importance of sound for cognitive sequencing abilities: The auditory scaffolding hypothesis. *Current Directions in Psychological Science*, *18*(5), 275-279. doi: 10.1111/j.1467-8721.2009.01651.x

Cunningham, A. E., & Stanovich, K. E. (1990). Assessing print exposure and orthographic processing skill in children: A quick measure of reading experience. *Journal of Educational Psychology*, *82*(4), 733-740.

Cunningham, A. E., & Stanovich, K. E. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of Educational Psychology*, *83*(2), 264-274. doi: 10.1037/0022-0663.83.2.264

Cunningham, A. E., & Stanovich, K. E. (1998). What reading does for the mind. *American Educator*, *22*, 8-17.

D'Alessio, M. J., Wilson, M. A., & Jaichenco, V. (2019). Morphological de-com-pos-it-ion helps recognize low-er frequency words in typically developing Spanish-speaking children. *Journal of Psycholinguistic Research*, *48*(6), 1407-1428. doi: 10.1007/s10936-019-09665-8

Damian, M. F., & Bowers, J. S. (2010). Orthographic effects in rhyme monitoring tasks: Are they automatic? *European Journal of Cognitive Psychology*, *22*(1), 106-116.

Daugaard, H. T., Cain, K., & Elbro, C. (2017). From words to text: Inference making mediates the role of vocabulary in children's reading comprehension. *Reading and Writing*, *30*(8), 1773-1788. doi: 10.1007/

s11145-017-9752-2

Davis, C., Castles, A., & Iakovidis, E. (1998). Masked homophone and pseudohomophone priming in children and adults. *Language and Cognitive Processes*, *13*(6), 625-651. doi: 10.1080/016909698386401

Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: Neural and behavioural evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1536), 3773-3800. doi: 10.1098/rstb.2009.0111

Dawson, N., Rastle, K., & Ricketts, J. (2018). Morphological effects in visual word recognition: Children, adolescents, and adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(4), 645-654. doi: 10.1037/xlm0000485

Dean, R. S., Yekovich, F. R., & Gray, J. W. (1988). The effect of modality on long-term recognition memory. *Contemporary Educational Psychology*, *13*(2), 102-115. doi: 10.1016/0361-476x(88)90010-0

Deary, I. J. (2001). Human intelligence differences: A recent history. *Trends in Cognitive Sciences*, *5*(3), 127-130. doi: 10.1016/S1364-6613(00)01621-1

Deary, I. J., Liewald, D., & Nissan, J. (2011). A free, easy-to-use, computer-based simple and four-choice reaction time programme: The Deary-Liewald reaction time task. *Behavior Research Methods*, *43*(1), 258-268. doi: 10.3758/s13428-010-0024-1

de Groot, F., Koelewijn, T., Huettig, F., & Olivers, C. N. L. (2015). A stimulus set of words and pictures matched for visual and semantic similarity. *Journal of Cognitive Psychology*, *28*(1), 1-15. doi: 10.1080/20445911.2015.1101119

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, *122*(3), 371-396. doi: 10.1037/0096-3445.122.3.371

Dehaene, S., Cohen, L., Morais, J., & Kolinsky, R. (2015). Illiterate to literate: Behavioural and cerebral changes induced by reading acquisition. *Nature Reviews Neuroscience*, *16*(4), 234-244. doi: 10.1038/nrn3924

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., Jobert, A., ... Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science*, *330*(6009), 1359-1364. doi: 10.1126/science.1194140

de Naeghel, J., van Keer, H., Vansteenkiste, M., & Rosseel, Y. (2012). The relation between elementary students' recreational and academic reading motivation, reading frequency, engagement, and comprehension: A self-determination theory perspective. *Journal of Educational Psychology*, *104*(4), 1006-1021.

Diependaele, K., Lemhöfer, K., & Brysbaert, M. (2013). The word frequency effect in first- and second-language word recognition: A lexical entrenchment account. *Quarterly Journal of Experimental Psychology*, *66*(5), 843-863. doi: 10.1080/17470218.2012.720994

Dobel, C., Enriquez-Geppert, S., Zwitserlood, P., & Bölte, J. (2014). Literacy shapes thought: The case of event representation in different cultures. *Frontiers in Psychology*, *5*, 290. doi: 10.3389/fpsyg.2014.00290

Domahs, U., Plag, I., & Carroll, R. (2014). Word stress assignment in German, English and Dutch: Quantity-sensitivity and extrametricality revisited. *The Journal of Comparative Germanic Linguistics*, *17*(1), 59-96. doi: 10.1007/s10828-014-9063-9

Dood, C., Gubbels, J., & Segers, P. C. J. (2020). *PISA-2018 de verdieping: Leesplezier, zelfbeeld bij het lezen, leesgedrag en leesvaardigheid en de relatie daartussen*. Nijmegen: Expertisecentrum Nederlands.

Doran, H., Bates, D., Bliese, P., & Dowling, M. (2007). Estimating the multilevel Rasch model: With the lme4 package. *Journal of Statistical Software*, *20*(2), 1-18.

Ducrot, S., Pynte, J., Ghio, A., & Lété, B. (2013). Visual and linguistic determinants of the eyes' initial fixation position in reading development. *Acta Psychologica*, *142*(3), 287-298. doi: 10.1016/j.actpsy.2013.01.013

Dumay, N., & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science*, *18*(1), 35-9. doi: 10.1111/j.1467-9280.2007.01845.x

Dumay, N., & Gaskell, M. G. (2012). Overnight lexical consolidation revealed by speech segmentation. *Cognition*, *123*(1), 119-132. doi: 10.1016/j.cognition.2011.12.009

Dumay, N., Gaskell, M. G., & Feng, X. (2004). A day in the life of a spoken word. In *Proceedings of the twenty-sixth annual meeting of the Cognitive Science Society* (p. 339-344). Mahwah, NJ: Erlbaum.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test (3rd ed.)*. Circle Pines, MN: American Guidance Service.

Dąbrowska, E. (2012). Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism*, *2*(3), 219-253. doi: 10.1075/lab.2.3.01dab

Dąbrowska, E. (2018). Experience, aptitude and individual differences in native language ultimate attainment. *Cognition*, *178*, 222-235. doi: 10.1016/j.cognition.2018.05.018

Edwards, L., & Anderson, S. (2014). The association between visual, nonverbal cognitive abilities and speech, phonological processing, vocabulary and reading outcomes in children with cochlear implants. *Ear and Hearing*, *35*(3), 366-374. doi: 10.1097/AUD.0000000000000012

Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading*, *18*(2), 116-125. doi: 10.1111/j.1467-9817.1995.tb00077.x

Elbro, C. (1996). Early linguistic abilities and reading development: A review and a hypothesis. *Reading and Writing*, *8*(6), 453-485. doi: 10.1007/Bf00577023

Elman, J. L. (2004). An alternative view of the mental lexicon. *Trends in Cognitive Sciences*, *8*(7), 301-306.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191. doi: 10.3758/bf03193146

Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. R. (1999). Individual differences in information-processing rate and amount: Implications for group differences in response latency. *Psychological Bulletin*, *125*(6), 777-799. doi: 10.1037/0033-2909.125.6.777

Favier, S., & Huettig, F. (2021). Long-term written language experience affects grammaticality judgements and usage but not priming of spoken sentences. *Quarterly Journal of Experimental Psychology*, *74*(8), 1378-1395. doi: 10.1177/17470218211005228

Favier, S., Meyer, A. S., & Huettig, F. (2020). Literacy can enhance syntactic prediction in spoken language processing. *Journal of Experimental Psychology: General*.

Federmeier, K. D., McLennan, D. B., De Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology*, *39*(2), 133-146. doi: 10.1017/S0048577202001373

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, *42*(1), 98-116. doi: 10.1037/

0012-1649.42.1.98

Ferrand, L., Meot, A., Spinelli, E., New, B., Pallier, C., Bonin, P., ... Grainger, J. (2018). MEGALEX: A megastudy of visual and auditory word recognition. *Behavior Research Methods*, *50*(3), 1285-1307. doi: 10.3758/s13428-017-0943-1

Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Meot, A., ... Pallier, C. (2010). The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods*, *42*(2), 488-496. doi: 10.3758/BRM.42.2.488

Fleischhauer, E., Bruns, G., & Grosche, M. (2021). Morphological decomposition supports word recognition in primary school children learning to read: Evidence from masked priming of German derived words. *Journal of Research in Reading*, *44*(1), 90-109. doi: 10.1111/1467-9817.12340

Foss, D. J., & Blank, M. A. (1980). Identifying the speech codes. *Cognitive Psychology*, *12*(1), 1-31.

Frishkoff, G. A., Collins-Thompson, K., Perfetti, C. A., & Callan, J. (2008). Measuring incremental changes in word knowledge: Experimental validation and implications for learning and assessment. *Behavior Research Methods*, *40*(4), 907-925.

Frishkoff, G. A., Perfetti, C. A., & Collins-Thompson, K. (2010). Lexical quality in the brain: ERP evidence for robust word learning from context. *Developmental Neuropsychology*, *35*(4), 376-403.

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, *19*(3), 117-25. doi: 10.1016/j.tics.2014.12.010

Gallik, J. D. (1999). Do they read for pleasure? Recreational reading habits of college students. *Journal of Adolescent and Adult Literacy*, *42*(6), 480-488.

Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition*, *89*(2), 105-132. doi: 10.1016/s0010-0277(03)00070-2

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y. S. (2008). A weakly informative default prior distribution for logistic and other regression models. *Annals of Applied Statistics*, *2*(4), 1360-1383. doi: 10.1214/08-Aoas191

Gijsel, M. A. R., Ormel, E. A., Hermans, D., Verhoeven, L., & Bosman, A. M. T. (2011). Semantic categorization and reading skill across dutch primary grades: Development yes, relationship no. *Journal of Child Language*, *38*(2), 356-379. doi: 10.1017/S0305000909990420

Goldinger, S. D. (1996). Auditory lexical decision. *Language and Cognitive Processes*, *11*(6), 559-568.

Grainger, J., & Ferrand, L. (1994). Phonology and orthography in visual word recognition: Effects of masked homophone primes. *Journal of Memory and Language*, *33*(2), 218-233.

Grainger, J., Lété, B., Bertand, D., Dufau, S., & Ziegler, J. C. (2012). Evidence for multiple routes in learning to read. *Cognition*, *123*(2), 280-292. doi: 10.1016/j.cognition.2012.01.003

Grainger, J., & Ziegler, J. C. (2011). A dual-route approach to orthographic processing. *Frontiers in Psychology*, *2*, 2-54. doi: 10.3389/fpsyg.2011.00054

Gubbels, J., van Langen, A., Maassen, N., & Meelissen, M. (2019). *Resultaten PISA-2018 in vogelvlucht* (Tech. Rep.). Universiteit Twente. doi: 10.3990/1.9789036549226

Hale, S. (1990). A global developmental trend in cognitive processing speed. *Child Development*, *61*(3), 653-663. doi: 10.1111/j.1467-8624.1990.tb02809.x

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). doi: 10.3389/fnhum.2012.00113

Hansen, J. H. L., & Bořil, H. (2018). On the issues of intra-speaker variability and realism in speech, speaker, and language recognition tasks. *Speech Communication*, *101*, 94-108. doi: 10.1016/j.specom.2018.05.004

Hansson, K., Forsberg, J., Lofqvist, A., Maki-Torkko, E., & Sahlen, B. (2004). Working memory and novel word learning in children with hearing impairment and children with specific language impairment. *International Journal of Language and Communication Disorders*, *39*(3), 401-422. doi: 10.1080/13682820410001669887

Hasenäcker, J., Beyersmann, E., & Schroeder, S. (2020). Morphological priming in children: Disentangling the effects of school-grade and reading skill. *Scientific Studies of Reading*, *24*(6), 484-499. doi: 10.1080/10888438.2020.1729768

Hasenäcker, J., & Schroeder, S. (2017b). Syllables and morphemes in German reading development: Evidence from second graders, fourth graders, and adults. *Applied Psycholinguistics*, *38*(3), 733-753. doi: 10.1017/S0142716416000412

Hasenäcker, J., Schröter, P., & Schroeder, S. (2017a). Investigating developmental trajectories of morphemes as reading units in German. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(7), 1093-1108. doi: 10.1037/xlm0000353

Hasenäcker, J., Verra, L., & Schroeder, S. (2019). Comparing length and frequency effects in children across modalities. *Quarterly Journal of Experimental Psychology*, *72*(7), 1682-1691. doi: 10.1177/1747021818805063

Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of d-prime. *Behavior Research Methods, Instruments, and Computers*, *27*(1), 46-51.

Hayes, D. P., & Ahrens, M. G. (1988). Vocabulary simplification for children: A special case of 'motherese'? *Journal of Child Language*, *15*(2), 395-410. doi: 10.1017/s0305000900012411

Hintz, F., Dijkhuis, M., van 't Hoff, V., McQueen, J. M., & Meyer, A. S. (2020). A behavioural dataset for studying individual differences in language skills. *Scientific Data*(1), 1-18.

Hintz, F., Jongman, S. R., Dijkhuis, M., van 't Hoff, V., McQueen, J. M., & Meyer, A. S. (2020). Shared lexical access processes in speaking and listening? An individual differences study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(6), 1048-1063. doi: 10.1037/xlm0000768

Horowitz-Kraus, T., & Hutton, J. S. (2015). From emergent literacy to reading: How learning to read changes a child's brain. *Acta Paediatrica*, *104*(7), 648-656. doi: 10.1111/apa.13018

Horst, J. S., & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) Database: A collection of novel images for use in experimental research. *Behavior Research Methods*, *48*(4), 1393-1409. doi: 10.3758/s13428-015-0647-3

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, *31*(1), 80-93. doi: 10.1080/23273798.2015.1047459

Huettig, F., & Pickering, M. J. (2019). Literacy advantages beyond reading: Prediction of spoken language. *Trends in Cognitive Sciences*, *23*(6), 464-475. doi: 10.1016/j.tics.2019.03.008

Hulslander, J., Olson, R. K., Willcutt, E. G., & Wadsworth, S. J. (2010). Longitudinal stability of reading-related skills and their prediction of reading development. *Scientific Studies of Reading*, *14*(2), 111-136.

Hurtado, N., Marchman, V. A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science*, *11*(6), F31-F39. doi: 10.1111/j.1467-7687.2008.00768.x

IBM Corp. (2011). IBM SPSS Statistics for Windows, Version 20.0.

James, A. N., Fraundorf, S. H., Lee, E. K., & Watson, D. G. (2018). Individual differences in syntactic processing: Is there evidence for reader-text interactions? *Journal of Memory and Language*, *102*, 155-181. doi: 10.1016/j.jml.2018.05.006

Janse, E., & Jesse, A. (2014). Working memory affects older adults' use of context in spoken-word recognition. *Quarterly Journal of Experimental Psychology*, *67*(9), 1842-1862. doi: 10.1080/17470218.2013.879391

JASP Team. (2020). JASP (Version 0.14).

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford: Oxford University Press.

Johnston, M., McKague, M., & Pratt, C. (2004). Evidence for an automatic orthographic code in the processing of visually novel word forms. *Language and Cognitive Processes*, *19*(2), 273-317. doi: 10.1080/01690960344000189

Jones, G., & Rowland, C. F. (2017). Diversity not quantity in caregiver speech: Using computational modeling to isolate the effects of the quantity and the diversity of the input on vocabulary growth. *Cognitive Psychology*, *98*, 1-21. doi: 10.1016/j.cogpsych.2017.07.002

Jongman, S. R., Khoe, Y. H., & Hintz, F. (2020). Vocabulary size influences spontaneous speech in native language users: Validating the use of automatic speech recognition in individual differences research. *Language and Speech*, *64*(1), 35-51. doi: 10.1177/0023830920911079

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, *80*(4), 437-447. doi: 10.1037/0022-0663.80.4.437

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31-36.

Kennedy, A. T. (1996). *Individual differences in adult word recognition* (Unpublished doctoral dissertation).

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*(3), 627-633. doi: 10.3758/BRM.42.3.627

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for

Dutch word frequency based on film subtitles. *Behavior Research Methods*, *42*(3), 643-650. doi: 10.3758/BRM.42.3.643

Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice effects in large-scale visual word recognition studies: A lexical decision study on 14,000 Dutch mono- and disyllabic words and nonwords. *Frontiers in Psychology*, *1*, 174. doi: 10.3389/fpsyg.2010.00174

Keuleers, E., Stevens, M., Mandera, P., & Brysbaert, M. (2015). Word knowledge in the crowd: Measuring vocabulary size and word prevalence in a massive online experiment. *Quarterly Journal of Experimental Psychology*, *68*(8), 1665-1692. doi: 10.1080/17470218.2015.1022560

Kezilas, Y., McKague, M., Kohnen, S., Badcock, N. A., & Castles, A. (2017). Disentangling the developmental trajectories of letter position and letter identity coding using masked priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(2), 250-258. doi: 10.1037/xlm0000293

Kidd, E., Bidgood, A., Donnelly, S., Durrant, S., Peter, M. S., & Rowland, C. F. (2020). Individual differences in first language acquisition and their theoretical implications. *Trends in Language Acquisition Research*(27), 189-219. doi: 10.1075/tilar.27.09kid

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual differences in language acquisition and processing. *Trends in Cognitive Sciences*, *22*(2), 154-169. doi: 10.1016/j.tics.2017.11.006

Kolinsky, R. (2015). How learning to read influences language and cognition. In *The oxford handbook of reading* (p. 377-393). New York, NY, US: Oxford University Press.

Kolinsky, R., Verhaeghe, A., Fernandes, T., Mengarda, E. J., Grimm-Cabral, L., & Morais, J. (2011). Enantiomorphy through the looking glass: Literacy effects on mirror-image discrimination. *Journal of Experimental Psychology: General*, *140*(2), 210-238. doi: 10.1037/a0022168

Kordes, J., Feenstra, H., Partchev, I., Feskens, R., & de Graaf, A. (2012). *Leesmotivatie, leesgedrag en leesvaardigheid van Nederlandse 15-jarigen. Aanvullende analyses op basis van PISA-2009* (Tech. Rep.). Cito.

Kosmidis, M. H., Tsapkini, K., & Folia, V. (2006). Lexical processing in illiteracy: Effect of literacy or education? *Journal Devoted to the Study of the Nervous System and Behavior*, *42*(7), 1021-1027.

Kroll, B. E. (1977). Combining ideas in written and spoken English: A look at subordination and coordination. In E. Keenan & T. Bennet (Eds.), *Discourse across time and space.* California, LA: University of Southern California.

Language & Reading Research Consortium, T., Currie, N. K., & Muijselaar, M. M. L. (2019). Inference making in young children: The concurrent and longitudinal contributions of verbal working memory and vocabulary. *Journal of Educational Psychology*, *111*(8), 1416-1431. doi: 10.1037/edu0000342

Lee, H., Seong, E., Choi, W., & Lowder, M. W. (2019). Development and assessment of the Korean author recognition test. *Quarterly Journal of Experimental Psychology*, *72*(7), 1837-1846. doi: 10.1177/1747021818814461

Leonard, L. B., Ellis Weismer, S., Miller, C. A., Francis, D. J., Tomblin, J. B., & Kail, R. V. (2007). Speed of processing, working memory, and language impairment in children. *Journal of Speech, Language and Hearing Research*, *50*(2), 408-428. doi: 10.1044/1092-4388(2007/029)

Levelt, W. J. M. (2001). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences*, *98*(23), 13464-13471. doi: 10.1073/pnas.231459498

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1-38.

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, *10*(8), 707-710.

Lewellen, M. J., Goldinger, S. D., Pisoni, D. B., & Greene, B. G. (1993). Lexical familiarity and processing efficiency: Individual differences in naming, lexical decision, and semantic categorization. *Journal of Experimental Psychology: General*, *122*(3), 316-330. doi: 10.1037/0096-3445.122.3.316

Liang, F., Paulo, R., Molina, G., Clyde, M. A., & Berger, J. O. (2008). Mixtures of g priors for Bayesian variable selection. *Journal of the American Statistical Association*, *103*(481), 410-423.

Lieven, E. (2016). Usage-based approaches to language development: Where do we go from here? *Language and Cognition*, *8*(3), 346-368. doi: 10.1017/langcog.2016.16

Lin, C. Y., Wang, M., Newman, R. S., & Li, C. (2018). The development of stress sensitivity and its contribution to word reading in school-aged children. *Journal of research in reading*, *41*(2), 259-277. doi: 10.1111/1467-9817.12094

Lopez Zunini, R. A., Baart, M., Samuel, A. G., & Armstrong, B. C. (2020). Lexical access versus lexical decision processes for auditory, visual, and audiovisual items: Insights from behavioral and neural measures. *Neuropsychologia*, *137*, 107305. doi: 10.1016/j.neuropsychologia.2019.107305

Lukács, A., Racsmány, M., & Pléh, C. (2000). Vocabulary and morphological patterns in Hungarian children with Williams syndrome: A preliminary report. *Acta Linguistica Hungarica*, *47*(1/3), 243-269. doi: 10.1023/a:1015603723980

Luque, J. L., Alvarez, C. J., Bordoy, S., Gimenez, A., Lopez-Perez, P. J., & Lopez-Zamora, M. (2020). Inhibitory effect of positional syllable frequency in Spanish 2nd and 4th grade readers. *Applied Psycholinguistics*, *42*(1), 1-17.

Lynch, S. M. (2007). *Introduction to applied Bayesian statistics and estimation for social scientists*. New York: Springer.

Lázaro, M., Acha, J., de la Rosa, S., García, S., & Sainz, J. (2017). Exploring the derivative suffix frequency effect in Spanish speaking children. *Reading and Writing*, *30*(1), 163-185. doi: 10.1007/s11145-016-9668-2

Lüdecke, D. (2019). sjstats: Statistical functions for regression models (version 0.17.8). , *10*. doi: 10.5281/zenodo.1284472

Mahr, T., & Edwards, J. (2018). Using language input and lexical processing to predict vocabulary size. *Developmental Science*, *21*(6), e12685. doi: 10.1111/desc.12685

Mainz, N. (2018). *Chapter 5: Origins of individual differences in word learning: An exploration of cognitive and environmental effects in adult native speakers. in: Vocabulary knowledge and learning: Individual differences in adult native speakers* (Unpublished doctoral dissertation).

Mainz, N., Shao, Z., Brysbaert, M., & Meyer, A. S. (2017). Vocabulary knowledge predicts lexical processing: Evidence from a group of participants with diverse educational backgrounds. *Frontiers in Psychology*, *8*(1164). doi: 10.3389/fpsyg.2017.01164

Makowski, D. (2018). The psycho package: An efficient and publishing-oriented workflow for psychological science. *Journal of Open Source Software*, *3*(22), 470.

Mani, N., & Huettig, F. (2014). Word reading skill predicts anticipation of upcoming spoken language input: A study of children developing proficiency in reading. *Journal of Experimental Child Psychology*, *126*, 264-279. doi: 10.1016/j.jecp.2014.05.004

Mar, R. A., & Rain, M. (2015). Narrative fiction and expository nonfiction differentially predict verbal ability. *Scientific Studies of Reading*, *19*(6), 419-433. doi: 10.1080/10888438.2015.1069296

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *11*(3), F9-F16. doi: 10.1111/j.1467-7687.2008.00671.x

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. *PLoS One*, *7*(8), e43230. doi: 10.1371/journal.pone.0043230

Marslen-Wilson, W. D., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, *8*(1), 1-71. doi: 10.1016/0010-0277(80)90015-3

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*(1), 29-63. doi: 10.1016/0010-0285(78)90018-X

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1-86.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419-457. doi: 10.1037/0033-295X.102.3.419

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. an account of basic findings. *Psychological Review*, *88*(5), 375-407. doi: 10.1037/0033-295x.88.5.375

McCutchen, D., Logan, B., & Biangardi-Orpe, U. (2009). Making meaning: Children's sensitivity to morphological information during word reading. *Reading Research Quarterly*, *44*(4), 360-376. doi: 10.1598/RRQ.44.4.4

McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review*, *119*(4), 831-877. doi: 10.1037/a0029872

McQueen, J. M., & Meyer, A. S. (2019). Key issues and future directions: Towards a comprehensive cognitive architecture for language use. In P. Hagoort (Ed.), *Human language: From genes and brain to behavior* (p. 85-96). Cambridge (MA): MIT Press.

Mestres-Missé, A., Rodriguez-Fornells, A., & Münte, T. F. (2007). Watching the brain during meaning acquisition. *Cerebral Cortex*, *17*(8), 1858-1866.

Michael, M. (2008). *Inference and context: Individual differences in interpretation and integration* (Unpublished doctoral dissertation).

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager,

T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49-100. doi: 10.1006/cogp.1999.0734

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), e1000097. doi: 10.1371/journal.pmed.1000097

Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, *137*(2), 267-96. doi: 10.1037/a0021890

Monaghan, P., Chang, Y. N., Welbourne, S., & Brysbaert, M. (2017). Exploring the relations between word frequency, language exposure, and bilingualism in a computational model of reading. *Journal of Memory and Language*, *93*, 1-21. doi: 10.1016/j.jml.2016.08.003

Montag, J. L., & MacDonald, M. C. (2015). Text exposure predicts spoken production of complex sentences in 8- and 12-year-old children and adults. *Journal of Experimental Psychology: General*, *144*(2), 447-468. doi: 10 .1037/xge0000054

Moore, M., & Gordon, P. C. (2015). Reading ability and print exposure: Item response theory analysis of the author recognition test. *Behavior Research Methods*, *47*(4), 1095-1109. doi: 10.3758/s13428-014-0534-3

Morais, J., Bertelson, P., Cary, L., & Alegria, J. (1986). Literacy training and speech segmentation. *Cognition*, *24*(1-2), 45-64. doi: 10.1016/ 0010-0277(86)90004-1

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, *7*(4), 323-331. doi: 10.1016/0010-0277(79)90020-9

Morais, J., & Kolinsky, R. (2002). Literacy effects on language and cognition. In *Psychology at the turn of the millennium, Vol. 1: Cognitive, biological, and health perspectives* (p. 507-530). Hove, UK: Psychology Press/Taylor and Francis.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, *76*(2), 165-178.

Moscoso del Prado, F. (2011). The mirage of morphological complexity. In *Proceedings of the annual meeting of the Cognitive Science Society* (Vol. 33).

Nelson, J. R., Balass, M., & Perfetti, C. A. (2005). Differences between written and spoken input in learning new words. *Written Language and Literacy*, *8*(2), 25-44. doi: 10.1075/wll.8.2.04nel

Neurobehavioral Systems. (2017). Presentation.

Newman, J. E., & Dell, G. S. (1978). The phonological nature of phoneme monitoring: A critique of some ambiguity studies. *Journal of Verbal Learning and Verbal Behavior*, *17*(3), 359-374. doi: 10.1016/S0022-5371(78)90228-1

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, *52*(3), 189-234.

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, *23*, 299-325.

Pattamadilok, C., Knierim, I. N., Duncan, K. J. K., & Devlin, J. T. (2010). How does learning to read affect speech perception? *Journal of Neuroscience*, *30*(25), 8435-8444.

Pattamadilok, C., Kolinsky, R., Ventura, P., Radeau, M., & Morais, J. (2007). Orthographic representations in spoken word priming: No early automatic activation. *Language and Speech*, *50*(4), 505-531.

Pattamadilok, C., Morais, J., De Vylder, O., Ventura, P., & Kolinsky, R. (2009). The orthographic consistency effect in the recognition of French spoken words: An early developmental shift from sublexical to lexical orthographic activation. *Applied Psycholinguistics*, *30*(3), 441-462.

Payne, B. R., Gao, X., Noh, S. R., Anderson, C. J., & Stine-Morrow, E. A. (2012). The effects of print exposure on sentence processing and memory in older adults: Evidence for efficiency and reserve. *Aging, Neuropsychology, and Cognition*, *19*(1-2), 122-149. doi: 10.1080/13825585.2011.628376

Pegado, F., Nakamura, K., Braga, L. W., Ventura, P., Nunes Filho, G., Pallier, C., ... Dehaene, S. (2014). Literacy breaks mirror invariance for visual stimuli: A behavioral study with adult illiterates. *Journal of Experimental Psychology: General*, *143*(2), 887-894. doi: 10.1037/a0033198

Perdijk, K., Schreuder, R., Baayen, R. H., & Verhoeven, L. (2012). Effects of morphological family size for young readers. *British Journal of Developmental Psychology*, *30*(3), 432-445. doi: 10.1111/j.2044-835X.2011.02053.x

Perfetti, C. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (p. 145–174). Hillsdale, NJ: Lawrence Erlbaum.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*(4), 357-383. doi: 10.1080/10888430701530730

Perfetti, C. (2011). 8. phonology is critical in reading. In S. Brady, D. Braze, & C. Fowler (Eds.), *Explaining individual differences in reading: Theory and* *evidence* (p. 153-175). New York, NY: Psychology Press.

Perre, L., Pattamadilok, C., Montant, M., & Ziegler, J. C. (2009). Orthographic effects in spoken language: On-line activation or phonological restructuring? *Brain Research*, *1275*, 73-80. doi: 10.1016/j.brainres.2009.04.018

Petersson, K. M., Silva, C., Castro-Caldas, A., Ingvar, M., & Reis, A. (2007). Literacy: A cultural influence on functional left–right differences in the inferior parietal cortex. *European Journal of Neuroscience*, *26*(3), 791-799. doi: 10.1111/j.1460-9568.2007.05701.x

Pfost, M., Dörfler, T., & Artelt, C. (2013). Students' extracurricular reading behavior and the development of vocabulary and reading comprehension. *Learning and Individual Differences*, *26*, 89-102. doi: 10.1016/j.lindif.2013.04.008

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(4), 329-347. doi: 10.1017/s0140525x12001495

Polse, L. R., & Reilly, J. S. (2015). Orthographic and semantic processing in young readers. *Journal of Research in Reading*, *38*(1), 47-72. doi: 10.1111/j.1467-9817.2012.01544.x

Quémart, P., Casalis, S., & Colé, P. (2011). The role of form and meaning in the processing of written morphology: A priming study in French developing readers. *Journal of Experimental Child Psychology*, *109*(4), 478-496. doi: 10.1016/j.jecp.2011.02.008

Quémart, P., Casalis, S., & Duncan, L. G. (2012). Exploring the role of bases and suffixes when reading familiar and unfamiliar words: Evidence from French young readers. *Scientific Studies of Reading*, *16*(5), 424-442. doi: 10.1080/10888438.2011.584333

Quémart, P., Gonnerman, L. M., Downing, J., & Deacon, S. H. (2018). The development of morphological representations in young readers: A cross-modal priming study. *Developmental Science*, *21*(4), e12607. doi: 10.1111/desc.12607

R Development Core Team. (2008). R: A language and environment for statistical computing.

Rastle, K., & Davis, M. H. (2008). Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes*, *23*(7-8), 942-971.

Raven, J. C., Court, J. H., & Raven, J. E. (1998). *Raven's coloured progressive matrices*. San Antonio, TX: Harcourt Assessment.

Rayner, K., Pollatsek, A., Ashby, J., & Clifton Jr, C. (2012). *Psychology of reading, 2nd ed.* New York, NY, US: Psychology Press.

Robinson, C. W., & Sloutsky, V. M. (2007). Visual statistical learning: Getting some help from the auditory modality. In *Annual meeting of the Cognitive Science Society* (Vol. 29).

Rodriguez-Aranda, C., & Jakobsen, M. (2011). Differential contribution of cognitive and psychomotor functions to the age-related slowing of speech production. *Journal of the International Neuropsychological Society: JINS*, *17*(5), 807-821. doi: 10.1017/S1355617711000828

Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language*, *57*(3), 348-379. doi: 10.1016/j.jml.2007.03.002

Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language*, *47*(1), 172-196. doi: 10.1006/jmla.2001.2839

Sailor, K. M. (2013). Is vocabulary growth influenced by the relations among words in a language learner's vocabulary? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(5), 1657-1662. doi: 10.1037/a0032993

Samuels, S. J., LaBerge, D., & Bremer, C. D. (1978). Units of word recognition: Evidence for developmental changes. *Journal of Verbal Learning and Verbal Behavior*, *17*(6), 715-720. doi: 10.1016/S0022-5371(78)90433-4

Schiff, R., Raveh, M., & Fighel, A. (2012). The development of the Hebrew mental lexicon: When morphological representations become devoid of their meaning. *Scientific Studies of Reading*, *16*(5), 383-403. doi: 10.1080/10888438.2011.571327

Schild, U., Becker, A. B. C., & Friedrich, C. K. (2014). Processing of syllable stress is functionally different from phoneme processing and does not profit from literacy acquisition. *Frontiers in Psychology*, *5*, 530.

Schild, U., Roder, B., & Friedrich, C. K. (2011). Learning to read shapes the activation of neural lexical representations in the speech recognition pathway. *Developmental Cognitive Neuroscience*, *1*(2), 163-174. doi: 10.1016/j.dcn.2010.11.002

Schilling, H. H., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory and Cognition*, *26*(6), 1270-1281. doi: 10.3758/bf03201199

Schlichting, J. E. P. T. (2005). *Peabody picture vocabulary test-III-NL*. Amsterdam, Netherlands: Harcourt Test Publishers.

Scott, C. (2008). A case for the sentence in reading comprehension. *Language, Speech, and Hearing Services in Schools*, *40*(2), 184-191.

Sears, C. R., Siakaluk, P. D., Chow, V. C., & Buchanan, L. (2008). Is there an effect of print exposure on the word frequency effect and the neighborhood size effect? *Journal of Psycholinguistic Research*, *37*(4), 269-291. doi: 10.1007/s10936-008-9071-5

Segers, E., & Verhoeven, L. (2016). How logical reasoning mediates the relation between lexical quality and reading comprehension. *Reading and Writing*, *29*(4), 577-590. doi: 10.1007/s11145-015-9613-9

Segui, J., Frauenfelder, U., & Mehler, J. (1981). Phoneme monitoring, syllable monitoring and lexical access. *British Journal of Psychology*, *72*(4), 471-477. doi: 10.1111/j.2044-8295.1981.tb01776.x

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*(4), 523-568. doi: 10.1037/0033-295X.96.4.523

Sereno, S. C., Rayner, K., & Posner, M. I. (1998). Establishing a time-line of word recognition: Evidence from eye movements and event-related potentials. *NeuroReport*, *9*(10), 2195-2200.

Seymour, P. H., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*(Pt 2), 143-174. doi: 10.1348/000712603321661859

Shankweiler, D. (1999). Words to meanings. *Scientific Studies of Reading*, *3*(2), 112-127. doi: 10.1207/s1532799xssr0302_2

Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? predictors of verbal fluency performance in older adults. *Frontiers in Psychology*, *5*, 772. doi: 10.3389/fpsyg.2014.00772

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, *55*(2), 151-218. doi: 10.1016/0010-0277(94)00645-2

Smith, M. C. (1996). Differences in adults' reading practices and literacy proficiencies. *Reading Research Quarterly*, *31*(2), 196-219. doi: 10.1598/RRQ.31.2.5

Sprenger-Charolles, L., Siegel, L. S., Béchennec, D., & Serniclaes, W. (2003). Development of phonological and orthographic processing in reading aloud, in silent reading, and in spelling: A four-year longitudinal study. *Journal of Experimental Child Psychology*, *84*(3), 194-217. doi: 10.1016/

S0022-0965(03)00024-9

Stainthorp, R. (1997). A children's author recognition test: A useful tool in reading research. *Journal of Research in Reading*, *20*(2), 148-158.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, and Computers*, *31*(1), 137-149. doi: 10.3758/bf03207704

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, *24*(4), 402-433. doi: 10.2307/747605

Stanovich, K. E., West, R. F., & Harrison, M. R. (1995). Knowledge growth and maintenance across the life span: The role of print exposure. *Developmental Psychology*, *31*(5), 811-826. doi: 10.1037/0012-1649.31.5.811

Storkel, H. L., Bontempo, D. E., & Pak, N. S. (2014). Online learning from input versus offline memory evolution in adult word learning: effects of neighborhood density and phonologically related practice. *Journal of Speech, Language and Hearing Research*, *57*(5), 1708-1721. doi: 10.1044/2014_JSLHR-L-13-0150

Street, J. A., & Dąbrowska, E. (2010). More individual differences in language attainment: How much do adult native speakers of English know about passives and quantifiers? *Lingua*, *120*(8), 2080-2094. doi: 10.1016/j.lingua.2010.01.004

Sullivan, A., & Brown, M. (2015). Reading for pleasure and progress in vocabulary and mathematics. *British Educational Research Journal*, *41*(6), 971-991. doi: 10.1002/berj.3180

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251-296. doi: 10.1023/A:1022193728205

Tamminen, J., & Gaskell, M. G. (2008). Newly learned spoken words show long-term lexical competition effects. *Quarterly Journal of Experimental Psychology*, *61*(3), 361-371. doi: 10.1080/17470210701634545

Tamminen, J., & Gaskell, M. G. (2013). Novel word integration in the mental lexicon: Evidence from unmasked and masked semantic priming. *Quarterly Journal of Experimental Psychology*, *66*(5), 1001-1025. doi: 10.1080/17470218.2012.724694

Tamminen, J., Payne, J. D., Stickgold, R., Wamsley, E. J., & Gaskell, M. G. (2010). Sleep spindle activity is associated with the integration of new memories and existing knowledge. *Journal of Neuroscience*, *30*(43), 14356-14360.

doi: 10.1523/JNEUROSCI.3028-10.2010

Tham, E. K., Lindsay, S., & Gaskell, M. G. (2015). Markers of automaticity in sleep-associated consolidation of novel words. *Neuropsychologia*, *71*, 146-157. doi: 10.1016/j.neuropsychologia.2015.03.025

Thiebaut de Schotten, M., Cohen, L., Amemiya, E., Braga, L. W., & Dehaene, S. (2014). Learning to read improves the structure of the arcuate fasciculus. *Cerebral Cortex*, *24*(4), 989-995. doi: 10.1093/cercor/bhs383

Thompson, G. L. (2011). *Towards a process-based characterization of syllable effects in visual word recognition* (Unpublished doctoral dissertation).

Torppa, M., Niemi, P., Vasalampi, K., Lerkkanen, M., Tolvanen, A., & Poikkeus, A. (2020). Leisure reading (but not any kind) and reading comprehension support each other: A longitudinal study across grades 1 and 9. *Child Development*, *91*(3), 876-900.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*(5), 352-373. doi: 10.1037/h0020071

Turner, J. E., Valentine, T., & Ellis, A. W. (1998). Contrasting effects of age of acquisition and word frequency on auditory and visual lexical decision. *Memory and Cognition*, *26*(6), 1282-1291. doi: 10.3758/bf03201200

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2011). Variation in verbal fluency: A latent variable analysis of clustering, switching, and overall performance. *Quarterly Journal of Experimental Psychology*, *64*(3), 447-66. doi: 10.1080/17470218.2010.505292

Unsworth, S. J., & Pexman, P. M. (2003). The impact of reader skill on phonological processing in visual word recognition. *Quarterly Journal of Experimental Psychology Section A*, *56*(1), 63-81. doi: 10.1080/02724980244000206

Vaessen, A., Bertrand, D., Tóth, D., Csépe, V., Faísca, L., Reis, A., & Blomert, L. (2010). Cognitive development of fluent word reading does not qualitatively differ between transparent and opaque orthographies. *Journal of Educational Psychology*, *102*(4), 827-842.

van den Bos, K., Spelberg, H., Scheepsma, A., & De Vries, J. (1994). *De Klepel. Vorm A en B. Een test voor de leesvaardigheid van pseudowoorden. Verantwoording, handleiding, diagnostiek en behandeling [The Klepel. Form A and B. A test for reading ability of pseudowords. Justification, manual, diagnostics and treatment].* Nijmegen, The Netherlands: Berkhout.

van der Ven, F., Takashima, A., Segers, E., & Verhoeven, L. (2015). Learning word meanings: Overnight integration and study modality effects. *PLoS One*, *10*(5), e0124926. doi: 10.1371/journal.pone.0124926

van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176-1190. doi: 10.1080/17470218.2013.850521

Ventura, P., Morais, J., & Kolinsky, R. (2007). The development of the orthographic consistency effect in speech recognition: From sublexical to lexical involvement. *Cognition*, *105*(3), 547-576. doi: 10.1016/j.cognition.2006.12.005

Waldie, K. E., & Mosley, J. L. (2000). Developmental trends in right hemispheric participation in reading. *Neuropsychologia*, *38*(4), 462-474.

Watkins, M. W., & Edwards, V. A. (1992). Extracurricular reading and reading achievement: The rich stay rich and the poor don't read. *Reading Improvement*, *29*(4), 236-242.

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, *24*(11), 2143-2152. doi: 10.1177/0956797613488145

Welcome, S. E., & Trammel, E. R. (2019). ERPs reveal relationships between neural orthographic priming effects and reading skill. *Journal of Psychophysiology*, *33*(2), 76-84. doi: 10.1027/0269-8803/a000211

West, R. F., Stanovich, K. E., & Mitchell, H. R. (1993). Reading in the real world and its correlates. *Reading Research Quarterly*, *28*(1), 35-50. doi: 10.2307/747815

Wift, A., & Ander, S. (2017). Rumors of the demise of books greatly exaggerated. *Gallup*.

Wigfield, A., & Guthrie, J. T. (1997). Relations of children's motivation for reading to the amount and breadth or their reading. *Journal of Educational Psychology*, *89*(3), 420-432. doi: 10.1037/0022-0663.89.3.420

Wolf, M. C., Muijselaar, M. M. L., Boonstra, A. M., & de Bree, E. H. (2019). The relationship between reading and listening comprehension: Shared and modality-specific components. *Reading and Writing*, *32*(7), 1747-1767. doi: 10.1007/s11145-018-9924-8

Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language*, *60*(4), 502-529. doi: 10.1016/j.jml.2009.02.001

Yap, M. J., Balota, D. A., Sibley, D. E., & Ratcliff, R. (2012). Individual differences in visual word recognition: insights from the English Lexicon Project. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(1), 53-79. doi: 10.1037/a0024177

Zaidel, E., & Schweiger, A. (1984). On wrong hypotheses about the right hemisphere: Commentary on K. Patterson and D. Besner,"Is the right hemisphere literate?". *Cognitive Neuropsychology*, *1*(4), 351-364.

Ziegler, J. C., Bertrand, D., Lété, B., & Grainger, J. (2014). Orthographic and phonological contributions to reading development: Tracking developmental trajectories using masked priming. *Developmental Psychology*, *50*(4), 1026-1036. doi: 10.1037/a0035187

Nederlandse samenvatting

Taal kun je tegenkomen in verschillende vormen: bijvoorbeeld in gesproken vorm wanneer je met iemand praat, en in het geval van sommige talen in de vorm van gebaren. Door de uitvinding van het schrift konden mensen ook taal in de schriftelijke vorm weergeven. Deze uitvinding heeft ervoor gezorgd dat we taal zowel in gesproken vorm als in schriftelijke vorm kunnen tegenkomen en gebruiken. Psycholinguïstische theorieën nemen veelal aan dat taal in ons hoofd gerepresenteerd wordt in de vorm van een soort mentaal woordenboek: voor elk woord is er een lemma in ons mentale woordenboek, en onder dat lemma vind je alle specifieke informatie die iemand weet over dat ene woord: hoe het uitgesproken wordt, hoe je het schrijft, wat het betekent, hoe het grammaticaal gebruikt kan worden, et cetera. Logischerwijs zijn niet alle stukjes informatie even precies: je weet hoe je "bagatelliseren" uitspreekt, en ongeveer wat het betekent, maar hoe spel je dat ook al weer? Met eerste een –a en dan een –e of andersom? Met één g of dubbel g, en met één l of dubbel ll? Aan de andere kant weet je misschien hoe je "compromis" schrijft, maar hoe spreek je dat nou uit? Met of zonder –s aan het einde? Nu is het zo dat je woorden waarvan deze specifieke informatie wel heel precies is, makkelijker kunt verwerken, en woorden waarvan je kennis wat minder specifiek is, minder goed verwerkt. Dat zou betekenen dat als je iemand het woord 'bagatelliseren' hoort uitspreken, je het snel zal herkennen, maar als je het woord leest, duurt het misschien een fractie van een seconde langer voordat je het woord herkent. Andersom zou je "compromis" misschien in de geschreven vorm snel herkennen, maar ben je net iets langzamer als je "compromis" of "compromie" hoort. Dit was dan ook de eerste vraag van dit proefschrift: is er een verschil in het verwerken van de ene presentatievorm (gesproken) vergeleken met de andere presentatievorm (geschreven) van hetzelfde woord?

De tweede vraag hing samen met het feit dat mensen onderling erg verschillen in hoe vaak ze met taal, of specifieker geschreven taal, in aanraking komen. Sommige mensen zijn echte lezers, terwijl anderen niet van lezen houden. Het wordt gedacht dat hoe meer je met taal in aanraking komt, hoe beter je taalvaardigheden worden. Geschreven taal is op een bepaalde manier erg bijzonder, en verschilt van gesproken taal. Zo is geschreven taal grammaticaal complexer dan de gesproken taal die we gebruiken in gesprekken. Het kent ook een grotere verscheidenheid aan woorden: geschreven taal bevat doorgaans veel moeilijke woorden die je in normale gesprekken niet zult gebruiken. Daarnaast kunnen mensen heel snel lezen: veel sneller dan wanneer je naar dezelfde tekst zou luisteren. Dus als je een kwartiertje leest, zal je meer taal opnemen dan wanneer je in datzelfde kwartiertje naar taal luistert (bijvoorbeeld via een audioboek). De tweede vraag van dit proefschrift was daarom: kunnen mensen die veel in aanraking komen met geschreven taal daadwerkelijk taal ook efficiënter verwerken dan mensen die niet veel in aanraking komen met geschreven taal?

In Hoofdstuk 2 van dit proefschrift werd gekeken in hoeverre de presentatievorm invloed heeft op hoe goed mensen woorden herkennen. In dit experiment kregen participanten, allemaal universitaire studenten, geschreven, gesproken of geschreven én gesproken woorden voorgeschoteld en moesten ze aangeven of ze het woord kenden of niet. De woorden verschilden in moeilijkheid, met woorden die niet veel mensen kennen (bijvoorbeeld "polemologie") en woorden die veel mensen wel kennen ("onromantisch"). Het bleek dat de participanten de gesproken, geschreven en gesproken én geschreven woorden even goed herkenden. Het maakte dus niet uit in welke vorm het woord werd gepresenteerd. Daarnaast namen we een aantal taken af die maten in hoeverre de participanten in aanraking kwamen met geschreven taal. Het bleek dat participanten die veel in aanraking kwamen met geschreven taal, meer moeilijke woorden herkenden dan participanten die weinig ervaring hadden met geschreven taal. Dat betekent dus dat mensen die veel lezen, meer moeilijke woorden kennen dan mensen die weinig lezen.

Hoofdstuk 3 ging dieper in op de bevindingen van Hoofdstuk 2. Deze keer namen we het experiment af bij zowel universitaire studenten als mbo-studenten. Universitaire studenten volgen doorgaans een theoretische opleiding, terwijl mbo-studenten een praktische opleiding volgen. Om die reden kunnen we aannemen dat universitaire studenten meer ervaring hebben met geschreven taal dan mbo-studenten. Het bleek inderdaad, net zoals in het vorige hoofdstuk, dat participanten die veel in aanraking kwamen met geschreven taal, meer moeilijke woorden herkenden dan participanten die minder ervaring hadden met geschreven taal. Dit suggereert dat ervaring met geschreven taal ervoor zorgt dat iemand een grotere woordenschat heeft (en dus meer moeilijke woorden kent). Daarnaast was het zo dat de participanten die veel ervaring hadden met geschreven taal, over het algemeen makkelijke woorden sneller herkenden dan participanten die weinig in aanraking kwamen met geschreven taal. Dit betekent dat door ervaring met geschreven taal de kennis in het mentale woordenboek preciezer is, waardoor men woorden makkelijker (en dus sneller) herkent. Wat betreft presentatievorm was er, net zoals in het vorige hoofdstuk, geen verschil in hoe goed de participanten de woorden herkenden: gesproken, geschreven en gesproken én geschreven woorden werden even goed herkend. De presentatievorm beïnvloedde wel de snelheid waarmee participanten woorden herkenden: geschreven woorden werden sneller herkend dan gesproken woorden. Dit komt waarschijnlijk doordat mensen sneller kunnen lezen dan luisteren naar gesproken taal. Een interessante bevinding was dat met name makkelijke woorden sneller werden herkend in de geschreven vorm. Kennelijk is het zo dat de geschreven presentatievorm het makkelijker maakt om woorden die je al goed kent, te herkennen.

Hoofdstuk 4 borduurde verder op de vraag in hoeverre ervaring met geschreven taal een invloed heeft op iemands taalvaardigheden. De taalvaardigheden woordbegrip en woordproductie werden onderzocht. Woordproductie is het plannen en uiteindelijk uitspreken van woorden. Om deze vraag te onderzoeken, werd een grote dataset geanalyseerd. Deze dataset bevatte data van een grote groep participanten die allerlei verschillende taaltaken had uitgevoerd. Ook waren algemene cognitieve vaardigheden gemeten, zoals bijvoorbeeld verwerkingssnelheid en intelligentie. Uit de analyses bleek dat zelfs wanneer je verschillen in cognitieve vaardigheden in ogenschouw nam, participanten met veel ervaring met geschreven taal over het algemeen taalvaardiger waren dan participanten die minder in aanraking kwamen met geschreven taal. Dit betekent dat mensen die veel lezen, beter en sneller woorden begrijpen, en ook beter en sneller woorden kunnen produceren.

In Hoofdstuk 5 werd uitgebreid onderzocht op welke manier ervaring met geschreven taal het verwerken van taal beïnvloedt. Er werd een groot literatuuronderzoek uitgevoerd om te zien wat de huidige wetenschappelijke literatuur beschrijft over dit fenomeen. In het literatuuronderzoek werden studies met volwassenen en studies met kinderen meegenomen. Uit de literatuurstudie bleek ten eerste dat ervaring met geschreven taal alle stukjes informatie over een woord in het mentale woordenboek preciezer maakt: mensen die veel in aanraking waren gekomen met geschreven taal, bleken meer kennis te hebben over de klank, de schrijfwijze en de betekenis van een woord. Ten tweede bleek dat de informatie niet alleen preciezer was op het niveau van het hele woord, maar dat mensen met veel ervaring met geschreven taal ook preciezere kennis hadden over kleinere stukjes van woorden, zoals lettergrepen en letters/klanken in een woord. Ten derde bleek dat ervaring met geschreven taal niet alleen geschreven taalverwerking verbetert, maar ook gesproken taalverwerking. Enthousiaste lezers kunnen dus niet alleen goed geschreven taal begrijpen, maar ook goed *gesproken* taal begrijpen. Het literatuuronderzoek van Hoofdstuk 5 legde ook een aantal zaken bloot waarover slechts weinig bekend is in de huidige literatuur. Zo is het relatief onduidelijk hoe ervaring met geschreven taal de kennis in het mentale woordenboek wat betreft grammatica beïnvloedt: is het zo dat mensen die veel lezen ook preciezere grammaticale kennis hebben van de woorden die zij kennen? Daarnaast werd het duidelijk dat er weinig kennis is over de relatie tussen ervaring met geschreven taal en taalverwerking bij kinderen. Kinderen leren immers lezen rond een jaar of zes en vanaf dat moment kunnen ze ervaring opdoen met geschreven taal. Vanaf welk moment heeft deze ervaring met geschreven taal invloed op hun taalverwerking? En wat als kinderen niet goed kunnen lezen of niet van lezen houden? Dit soort vragen zullen in de toekomst onderzocht moeten worden.

In Hoofdstuk 6 werd weer teruggegrepen op de vraag hoe presentatievorm het verwerken van taal beïnvloedt. In dit hoofdstuk werd niet gekeken naar het verwerken van woorden die men al weet, maar het aanleren en later herkennen van *nieuwe* woorden. In dit experiment leerden participanten, allemaal universitaire studenten, nieuwe woorden zoals “floemhaafs” en “walmlaf”. Deze woorden waren geen echte Nederlandse woorden, maar volgen wel de regels van het Nederlands. De participanten leerden 24 van deze nepwoorden te associëren met 24 plaatjes van gekke, onbekende objecten. Sommige participanten kregen de woorden in gesproken vorm aangeboden via de koptelefoon, en andere participanten lazen de woorden op een computerscherm. Elk van de 24 woorden werd zeven keer aangeboden. Daarna volgde een test, waarin goede en foute combinaties van de plaatjes en woorden te zien waren. De participanten moesten aangeven of de combinatie goed of fout was, op basis van wat ze eerder hadden geleerd. In de test werden de woorden weer geschreven of gesproken aangeboden. Sommige participanten hoorden dus de woorden zoals ze die tijdens het leren ook hadden gehoord. Sommige participanten lazen de woorden zoals ze die tijdens het leren hadden gelezen. Als laatste was er een groep participanten die eerder de woorden hadden gehoord of gelezen, maar ze nu in de anderen vorm aangeboden kregen: als ze de woorden eerst hadden gehoord, kregen ze de woorden nu te zien, en als ze eerst de woorden hadden gelezen, kregen ze de woorden nu via de koptelefoon. Het bleek uiteindelijk dat participanten de nieuw geleerde woorden goed konden herkennen op de test, maar dat het niet uitmaakte of ze de woorden in de geschreven of gesproken vorm hadden geleerd. Opmerkelijk was dat de groep participanten die op de test de woorden in de andere vorm aangeboden kregen, de woorden wel goed konden herkennen, ondanks dat ze de woorden nog nooit in deze vorm hadden gezien of gehoord. Uit deze studie bleek dat Nederlandse, hoogopgeleide volwassenen even efficiënt de gesproken en geschreven vorm van een woord kunnen leren. Daarnaast bleek dat deze groep mensen een geschreven of gesproken mentale versie kunnen creëren van een woord, ook al zijn ze dit woord alleen in de andere vorm tegengekomen.

Samengevat: in dit proefschrift werden twee vragen behandeld die te maken hebben met het feit dat wij mensen, dankzij de uitvinding van het schrift, taal zowel in de gesproken als geschreven vorm kunnen tegenkomen. De eerste vraag was of er een verschil is in het verwerken van de ene presentatievorm (gesproken) vergeleken met de andere presentatievorm (geschreven) van hetzelfde woord. Het bleek dat presentatievorm geen invloed had op het leren van nieuwe woorden, maar dat het herkennen van bekende woorden wel beïnvloed werd door presentatievorm. Hoewel gesproken en geschreven woorden wel even goed werden herkend, was het herkennen van geschreven woorden sneller (met name wat betreft makkelijke woorden). De tweede vraag was of mensen die veel in aanraking komen met geschreven taal, taal daadwerkelijk ook efficiënter kunnen verwerken dan mensen die niet veel in aanraking komen met geschreven taal. Uit vrijwel alle uitgevoerde studies bleek dat mensen die veel ervaring hadden met geschreven taal, taal inderdaad efficiënter konden verwerken. Ze konden woorden beter en sneller herkennen, begrijpen en produceren. Een verklaring hiervoor zou zijn dat dankzij ervaring met geschreven taal, de kennis van woorden in het mentale woordenboek preciezer en gedetailleerder wordt. Deze gedetailleerde woordkennis vergemakkelijkt het verwerken en gebruiken van deze woorden. Dit proefschrift onderschrijft hoe een ogenschijnlijk simpele uitvinding zoals het schrift, in feite een visuele weergave van taalklanken, een enorme invloed heeft op ons taalvermogen.

English Summary

Language can be encountered in many forms. For example, one can encounter spoken language when having a conversation with someone, or encounter language in the form of signs when communicating in sign language. When humans invented writing about 5500 years ago, language could now also be encountered in written form. Psycholinguistic theories often assume that language is mentally represented in our brains in the form of a mental dictionary. Every word you know is stored under a mental lemma in your mental dictionary. This lemma contains all information you know about this word: how you pronounce it, how you spell it, what it means, how you use it grammatically, et cetera. Naturally, for a single word these different pieces of information are not equally well known. For example, you may know that Wednesday is pronounced as “wenzday”, but you may have to think twice when spelling it as “wed-nes-day”. Also, you may have encountered the word “chicanery” in a book, but how it’s pronounced or what it exactly means may be a mystery (“shi-KAY-nuh-ree” is a form of deception). It happens to be the case that processing words becomes easier when your mental lemma of the word contains a lot of detailed information. When this information is not so clearly defined, processing the word is more difficult. For example, you may be very fast at recognizing “wenzday” in a conversation, but recognition may be slightly slower when you read Wednesday. Likewise, you may quickly recognize “chicanery” while reading a novel, but when you hear someone pronounce it, you may have to think twice before understanding what they just said. The first question of this dissertation was related to this phenomenon: is processing a word presented in one form (written) different from processing the same word presented in a different form (spoken)?

The second question of this dissertation was related to the observation that humans tend to differ in how much they are exposed to language, in particular written language. Some people are very avid readers, whereas others don’t like to read at all. It is thought that your language skills improve as you encounter more language. Written language has some very distinct features that sets it apart from spoken language. For example, written language tends to be more structured and complex in terms of grammar. Also, written language is characterized by a large variety of words, whereas in conversational spoken language we tend to use a smaller set of simpler words. Reading is also generally much faster than listening to someone speaking. If you would read for 15 minutes, you would encounter more language than if you would listen to language in those same 15 minutes (for example by listening to an audiobook). The second question of this dissertation was therefore: is it the case that people who are very experienced with written language, also process language more efficiently than people who are less experienced with written language?

Chapter 2 of this dissertation investigated the influence of form of presentation on word recognition. Participants, all university students, were presented written, spoken or audio-visually presented words and had to indicate whether they knew the word or not. The words differed in their difficulty, as some words were generally well-known to people ("unromantic") whereas others were not ("polemology"). It was found that participants could recognize the written, spoken and audiovisually presented words equally well. It did not matter in which way the words were presented. In addition, several tasks were administered to measure to what extent the participants were experienced with written language. It was found that participants who were very experienced with written language knew more difficult words than participants who were less experienced with written language. This indicates that people who read a lot, tend to know more words than people who do not read much.

Chapter 3 examined the findings of Chapter 2 in more depth. This time, the experiment was administered in both university- and vocational education students. University courses are characterized by learning from books, whereas vocational education courses are more practically oriented. We therefore assumed that university students would be more experienced with written language than vocational education students. It was indeed the case, like in Chapter 2, that participants who were more experienced with written language knew more difficult words than participants who were less experienced with written language. This indicates that experience with written language increases one's vocabulary size. Furthermore, it was found that participants experienced with written language also recognized easy words faster than participants with less experience with written language. Thus, as a result of experience with written language, the mental knowledge of words increases, which allows for speedier recognition. With regard to form of presentation, there was no difference in how well participants recognized the written, spoken or audio-visually presented words. However, it turned out that participants were able to recognize words faster when these words were presented in their written forms. This is likely due to the fact that people can generally read faster than that they can listen to speech. An interesting finding was that easy words in particular were recognized faster in their written forms. It appears to be the case that presentation in the written form improves recognition of words you already know pretty well.

Chapter 4 continued investigating to what extent experience with written language influences language abilities. In particular, the language skills word comprehension and word production were examined. Word production refers to the planning and ultimately the pronunciation of words. In order to investigate this topic, a large dataset was analysed. This dataset contained data of a large group of participants who had performed all sorts of linguistic tasks. Tasks measuring cognitive abilities, such as processing speed and intelligence, were also administered. It was found that even if you control for differences between individuals with regard to their cognitive skills, participants who were experienced with written language scored higher on the language abilities tasks than participants who were less experienced with written language. This means that people who read a lot tend to understand language faster and more accurately, and are also able to produce language faster and more accurately.

Chapter 5 thoroughly examined how experienced with written language influences language processing. In a large literature review the scope of the current literature on this topic was assessed. Both studies conducted with adults, as well as studies conducted with children were included in the review. The review provided some very useful insights. First, it was found that experience with written language improves people's mental knowledge about language: participants who are very experienced with written language appear to have more precise knowledge about the pronunciation of words, the spelling of words, and the meaning of words. Second, it was found that this mental information is not only more detailed on the level of the whole word, but that people who are experienced with language also have more precise knowledge about smaller parts of words as well, such as the syllable or letters/sounds. Third, the review showed that experience with written language does not only improve processing written language, but also processing spoken language. This means that avid readers are not only better at understanding written language, but also spoken language. The literature review also uncovered several topics that are underrepresented in the current literature. It is for example not clear how experience with written language improves mental knowledge of grammar: do individuals who read a lot also have more precise grammatical knowledge of the words they know? Furthermore, there seemed to be little knowledge about the relationship between experience with written language and language processing in children. Children learn to read at the age of six and from that moment onwards, they can encounter written language. At what moment in their development does this experience with written language influence their language processing? And what if children have difficulty reading, or do not like to read? These questions must be examined further in future research.

Chapter 6 returned to the question of how form of presentation influences language processing. This chapter did not examine processing known words, but instead looked at learning and later recognizing new words. In this experiment, participants (all university students) learned new Dutch-like words. These words were not real Dutch words, but did follow the rules of Dutch so that they appeared Dutch-like to native speakers of Dutch. Participants learned to associate 24 words with 24 pictures of strange, unknown objects. Some participants learned the words in the spoken forms, whereas other participants learned the words in their written forms. Each word/picture combination was shown seven times. After that, there was a test in which participants were shown correct and incorrect word/picture combinations, and participants had to indicate whether the combination was correct or not based on what they had learned previously. During the test, words were again presented in the written form for some participants and in the spoken form for other participants. Thus, some participants learned and were later tested on the written forms and some participants learned and were later tested on the spoken forms. Some participants learned the words in the written/spoken form, but were later confronted with the other form, which they had never encountered before. It was found that all participants learned the words quite well, but that it did not matter whether they learned the words in the spoken or written form. Remarkable was that the participants who learned the words in one form, but were later confronted during the test with the other form, were able to recognize the words very well, even though they had never encountered the words in this form before. This study showed that Dutch, highly educated adults are able to learn new words equally efficient when they encounter them in the written or spoken form. Furthermore, this group of people is able to create a written or spoken mental version of the word, even if they had only ever encountered this word in the other form.

To conclude: this dissertation investigated two topics related to the fact that humans can encounter language in the written and spoken form, thanks to the invention of writing. The first question was whether there was a difference in processing written versus spoken words. It was found that form of presentation did not influence learning new words, but that recognition of known words was influenced by form of presentation. Although written and spoken words were equally well recognized, recognition was faster for words presented in the written form (especially easy words). The second question was whether people who are very experienced with written language process language more efficiently than people who are less experienced with written language. All studies provided evidence that people experienced with written language were indeed better at processing language. They were able to recognize words faster and more accurately, and also comprehend and produce language more efficiently. An explanation for this finding may be that experience with written language improves the mental knowledge of words in the mental dictionary. This detailed knowledge may improve processing and using words. This dissertation highlights how a seemingly simple invention like writing, in fact only a visual representation of language sounds, has such an enormous influence on our linguistic abilities.

Acknowledgements

Lev Vygostky said: "It is through others that we become ourselves"[1]. In other words, personal growth is the result of interactions with other people. They provide learning challenges just outside your current comfort zone, and guidance when performing these difficult challenges. Of course, development still requires a tremendous amount of motivation on the learner's part. But only through interaction with others, we can go beyond our current level of development and reach our full potential.

A PhD, a learning challenge like many others you encounter during your life, is always accompanied by a larger group of people that supported and guided you. A PhD is also a distinct phase in your life, one that you probably will never go through again, and therefore finishing a PhD always comes with many I-will-miss-yous and good-byes. Fortunately, you can dedicate a whole chapter in your dissertation to this.

Let's start with my supervisory team from the past four years: Alastair, Florian, Antje and Caro. Alastair, I always looked forward to our weekly chats in your office, where we would discuss difficult, scientific topics: you had me consider every possible outcome and its consequence while designing an experiment, which taught me to anticipate and think ahead. We also talked a lot about light-hearted subjects such as our hobbies, holiday destinations, and the many cultural or linguistic differences we observed at our multicultural workplace. The difference in meaning between a meeting at 'half two' and 'half twee' turned out to be particularly troubling for us! Florian, you entered the supervisory team like a fresh breeze, providing a new line of research (individual differences) that seamlessly merged with the already existing topic (modality effects) of my dissertation. You brought with you a lot of data, which was already collected by you and your team, allowing us to venture into new topics and explore new hypotheses. Also, you taught me about the inner workings of the publishing world: I learned to write manuscripts that match the premise of a particular journal, and learned to deal with desk rejections and (difficult) reviewers. Antje, you

[1]Vygotsky, L. S. (1987). The genesis of higher mental functions. In R. Reiber (Ed.), *The history of the development of higher mental functions* (Vol. 4, pp. 97-120). New York: Plennum.

kept a watchful eye on the quality of our research: your comments on hypotheses, research designs and manuscripts were insightful and always on point. I learned to become more precise at reasoning and writing arguments, a very helpful skill I will use and value the rest of my life. Caro, you showed me that working in science is not only about research, but also about showing the world that the research we do is important. You taught me to be persuasive in my writing by always circling back to 'this-study-is-important-because...'. If the writer is not convinced about the significance of their study, then the reader will not be convinced either! We shared the belief that we should not only convince fellow researchers about the importance of our work, but that we should take it much further. In your role as Managing Director you supported me setting up our official science communication blog 'MPI TalkLing', a first in the history of the Max Planck Society!

Now, onto two large bodies of support: my two departments within the MPI. I was one of the few people in multiple departments and I always enjoyed observing the differences between the two departments, that each had their own spirit. I will just go over the departments alphabetically, since it is quite impossible to pick a favourite and discuss that one first.

Alphabetically, the Language Development Department comes first. When I started my PhD in September 2017, this department was brand new. Within four years, LaDD evolved from a cosy and small department to a cosy and large department. LaDD is filled to the brim with enthusiastic and kind people, who have many shared interests, including baking (our department table near Caro's office was always full of home-made 'baksels'). You will all be missed: Julia, Katja, Melis, Shanthi, Yevheniy, Ingeborg, Christina, Tineke, Middy, Raquel, Rebecca, Rowena, Andrew, Evan, Seamus, Patricia, Nienke, Ilse, Jeroen, Pim and Caro, and of course all research assistants, student assistants, trainees, interns and guest researchers, and everyone from the First Language Acquisition group/ Centre for Language Studies / Baby and Child Research Center of the Radboud.

Next is the Psychology of Language Department. Home of a brilliant and friendly group of researchers. Whether it was helping you writing complicated analysis scripts or grabbing a coffee together: every activity was welcomed with the same enthusiasm. I will not forget you: Johanne, Nina, Merel M., Amie, Limor, Sara, Miguel, Joe, Greta, Eirini, Saoradh, Federica, Fan, Aitor, Jieying, Sophie, Jeroen, Giulio, Suzanne, Philip, Falk, Andrea, Hans Rutger, Caitlin, Annelies, Laurel, Florian, Marjolijn, Vera, Alastair, Cecilia, Ava, Elli, Orhun, Ronny, Ruth and Antje, and of course all research assistants, student assistants, trainees, interns and guest researchers.

During my PhD I took part in many internal MPI projects. For two years I was one of the PhD representatives of our institute. I really enjoyed working closely the other representatives: Limor, Merel P. Julia E. Julia M., Teun, Fenja, Shanthi, Melis, and in particular Merel M. with whom I travelled to the PhDnet Meeting in Tübingen in 2018.

The IMPRS Conference 2020 was another side project I was involved with. The conference committee already started planning a year in advance, but two months before the conference, COVID-19 kicked in, and we had to transfer the whole event to an online environment. This was only possible due to our amazing team: Katja, Chén, Melis, Adrian Adrian, Joery, Federica, Julia E., Sophie and Kevin.

MPI TalkLing, our science communication blog, is the side project I take most pride in. Originally conceptualized in June 2018 with Kevin and Julia E., it took flight when Marjolein came on board. After two years of preparation, it was launched in November 2020. I was editor-in-chief of an amazing team and leading the blog was one of the things I enjoyed most during my time at the MPI. Marjolein, Kevin and Maurice, you made it possible that this website aired. Caro saw the blog's potential and importance for our MPI and supported it enthusiastically from the beginning. All members of the MPI TalkLing blog (Francie, Alessio, Dilay, Federica, Francesca, Greta, Julia E., Julia M., Laurel, Marina, Naomi, Natalia, Natascha, Rowan, Sophie S., John, Melis, Guillermo, Sara, Cecilia, Adam, Lynn, Ava, Dennis, Caitlin, Inge, Cielke Elly, Annelies, Barbara, Bianca, Fenja, Ronny): your amazing and interesting content made and will continue to make this blog a great success. Eva and Julia von der F., I am very happy that I could hand over the official editor-in-chief-pen to you both, and I am confident that you will continue to lead this blog "met verve".

At the MPI there were many other people that have not been mentioned that I greatly enjoyed working with. More senior research staff and IMPRS'ers alike, I thoroughly enjoyed your advice on my projects, chats in the hallway and the occasional coffees.

The MPI does not only employ research staff, but also a great support staff that should not be forgotten. Kevin, you coordinate the IMPRS with so much dedication and spirit. TG staff, no single project at the MPI can be done without your help. Meggie and Karin, you run the library like no other. Operations staff, you ensure that everything runs smoothly and hiccups are resolved even before we may notice them. Receptionists and secretaries, your help and friendly faces made working at the MPI a delight.

Surely, I should not forget my paranymphs Julia E. and Naomi. Julia, één van mijn eerste kamergenootjes. We konden altijd praten over werkdingen en ook allerlei andere zaken. Je bent vrolijk en empathisch, en lijkt altijd te weten hoe anderen zich voelen. Ook een echte regelaar die ervoor zorgde dat alle evenementen die wij samen organiseerden perfect verliepen. Naomi, een zonnestraal op een regenachtige dag. Vrolijk en enthousiast, altijd overal voor in. Maar ook serieus, als we over moeilijke zaken praatten (met een kopje koffie in de hand). Ik ga jullie missen!

Buiten het MPI waren er ook een groot aantal mensen die een rol speelden de afgelopen vier jaar. Ten eerste, mijn scriptiebegeleiders van de Universiteit van Amsterdam: Peter, Madelon, Marloes and Elise. Terwijl ik werkte bij het MPI hebben we ook nog contact gehouden en publiceerden we een mooi artikel gebaseerd op mijn masterscriptie. Elise, in het laatste half jaar kon ik daarnaast rekenen op jouw steun en luisterend oor, en dat zal mij altijd bijblijven.

Nu treden we buiten de werksfeer en wil ik mijn vrienden van de middelbare school en studie onder de aandacht brengen. De afgelopen jaren zag ik ons allemaal groeien, van studenten naar Grote Mensen met Volwassen Banen. Iedereen van het Utrechts Studenten Koor en Orkest (USKO), wat hebben we mooie muziek gemaakt in het binnen en buitenland de afgelopen jaren. Annet, mijn voormalig huisgenootje, jouw gezelligheid en vrolijkheid maken elke dag een goede dag. Sanne en Liza, door onze gezellige kletsmiddagen de afgelopen jaren word ik blij als ik aan jullie denk.

De schoonfamilie mag natuurlijk ook niet ontbreken in dit lijstje: de Lanooy-Kuijperclan stond altijd klaar. Jullie liefde en betrokkenheid is altijd aanwezig. Elk bezoek was altijd weer een kleine vakantie, ook al was het maar een middag.

Wouter, de afgelopen 5 jaar zat je dan wel aan de andere kant van de Atlantische Oceaan, maar uit het oog is niet uit het hart. Als oudere broer had jij al gauw door dat je kleine zusje veel in haar mars had (“ze kan heus wel naar het gymnasium”). Ik ben blij dat je weer terug bent aan déze kant van de oceaan, samen met Jenny, en dat we elkaar weer vaker kunnen zien.

Lieve Jan Pieter, JP. Samen zijn we het avontuur in Arnhem/Nijmegen aangegaan en ik had het niet anders kunnen of willen doen. Elke dag kon ik weer fijn thuiskomen bij jou en toen we thuiswerkten, werden we zelfs collega’s met een vaste koffie- en lunchpauze. Jij ben de ster aan mijn hemel, de dauw op de bladeren, de vogels in de ochtend: jij geeft het leven extra glans!

Dan als laatste: papa en mama. Jullie hebben mij precies gegeven wat je nodig hebt om een promotietraject succesvol te doorlopen. Analytisch en kritisch aan de ene kant, praktisch en georganiseerd aan de andere kant. Het grote geheel zien enerzijds, oog voor detail anderzijds. Jullie hebben mij een brede interesse meegegeven, nieuwsgierigheid naar van alles en nog wat, doorzettingsvermogen en zelfvertrouwen. Ik hou van jullie!

Curriculum Vitae

Merel Wolf was born on July 13, 1994 in Amsterdam, the Netherlands. From an early age, she has been interested in many different topics, such as (natural) history, geography, biology, zoology and psychology. Basically any question about how the world works, how its inhabitants live, and how they all came to be and continue to grow, sparks her curiosity. Growing up in Utrecht, she decided to pursue a Bachelor in Psychology at the University of Utrecht in 2012. She followed the track 'Child and Youth Psychology', did a minor in 'Learning Disorders' and took part in the multidisciplinary Honours Program 'Descartes College'. She wrote her Bachelor's thesis in collaboration with the University of Amsterdam on aloud and silent reading in children. She then decided to continue her studies in 2015 at the University of Amsterdam with the Research Master Child Development and Education. There, she wrote two Master's theses: the first about reading and listening comprehension in children and the second about eye movements during aloud and silent reading in children. While studying, she worked as a student assistant on a two-year project that investigated perceived stress in young, starting high school teachers. After graduating in 2017 she was awarded a four-year fellowship by the International Max Planck Research School (IMPRS) for Language Sciences to pursue a PhD at the Max Planck Institute (MPI) in Nijmegen. During her PhD, she divided her time between the Language Development Department and Psychology of Language Department. She also invested time in many side-projects, such as being one of the institute's PhD representatives, organizing events such as the IMPRS Conference 2020, supervising high school students with their final year research project (profielwerkstuk), and last, but certainly not least, initiating and being editor-in-chief of the MPI TalkLing science blog. She currently works as a researcher for Expertisecentrum Beroepsonderwijs (ECBO) [Center of Expertise for Vocational Education] in Den Bosch.

Publications

Wolf, M. C., Meyer, A. S., Rowland, C. F., & Hintz, F. (2021). The effects of input modality, word difficulty and individual differences in reading experience on word recognition accuracy. *Collabra: Psychology. 7*(1): 24919. doi:10.1525/collabra.24919

Wolf, M. C., Smith, A. C., Meyer, A. S., & Rowland, C. F. (2019). Modality effects in vocabulary acquisition. In A. Goel, C. Seifert, & C. Freksa (Eds.), *Proceedings of the 41st Annual Meeting of the Cognitive Science Society (CogSci 2019)* (pp. 1212-1218). Austin, TX: Cognitive Science Society.

Wolf, M. C., Muijselaar, M. M. L., Boonstra, A. M., & De Bree, E. H. (2019). The relationship between reading and listening comprehension: Shared and modality-specific components. *Reading and Writing, 32*(7), 1747-1767. doi:10.1007/s11145-018-9924-8.

Wolf, M. C. (2015). Het verschil tussen hardop en stillezen wat betreft leessnelheid en tekstbegrip en de invloed hierop van fonologisch bewustzijn, benoemsnelheid en visuele aandachtsspanne. *Student Undergraduate Research E-journal, 1*(1), 261-264.