Moderating predictive language processing: The role of age and reading skill

Claudia Gerards
SNR 2012633

Master’s Thesis
Communication and Information Sciences
Specialization Communication and Cognition

Department Communication and Cognition
School of Humanities and Digital Sciences
Tilburg University, Tilburg

Supervisor: Dr. J. A. A. Engelen
Second reader: Dr. S. Bögels

March 2022

Abstract

Despite their potentially large influence, factors that moderate linguistic anticipatory processing are not yet well-understood. Recent findings suggest that the detrimental effects of high age on predictive processing may be mitigated by advanced language skills. In an online experiment, 76 Dutch speaking adults of varying ages and reading skill levels performed a maze task. Highly constraining stories (e.g., "It is Sunday morning. The whole religious family goes, as always, to…") were presented as sequences of choices between distractors and correct continuations. Stories ended with either a predictable or an unpredictable gender-marked article and noun ("the church" vs. "the house of prayer"). Reading times on the articles were recorded; predictive processing was indexed by reading times that were relatively longer on unpredictable articles and shorter on predictable articles. Reading skill was measured using an adaptation of the Woodcock-Johnson reading fluency task and a lexical decision task. Processing speed and memory capacity were assessed with four cognitive tasks. A 2x2x2 mixed ANOVA revealed three main effects on article reading time (unpredictable > predictable articles, middle-aged > younger adults, less skilled > highly skilled readers). Apart from an unexpected three-way interaction (age group interacted with article predictability in the hypothesized manner, but only among highly skilled readers), no convincing evidence was found for serious deterioration in predictive language processing related to poor reading skill or higher age. We encourage future studies to better investigate who does and does not systematically predict upcoming language as well as whether prediction is necessary for comprehension.

Keywords: prediction, language, age, reading skill, maze task, sentence processing

Index

Preface

Language has always fascinated me. The incentive, I believe, came from my parents, who both speak a variety of languages very well. My mother in particular is an avid reader and lives by the motto that money spent on books is never wasted money. Unsurprisingly, I was excited about being assigned this thesis topic. Still, linguistic prediction is not the most straightforward topic and I must admit it took me some weeks to come to a basic understanding of what I was going to research, let alone how.

I aspired to step outside the typical university student population that most scientific research has been conducted on, because I suspected that the majority of people with whom I interacted on a daily basis would respond differently from what the literature on predictive language processing hypothesized. Does *everybody* predict language? Does prediction depend on one's age or reading skill? What if reading skill can preserve cognitive functioning later in life? Is money spent on books really never wasted money?

Although it was challenging and even frustrating at times, I have truly enjoyed the process of delving into the matter, conducting the experiment, and reporting the results. A sincere thank you goes out to my thesis supervisor. Throughout this project, he has been incredibly helpful, patient, and kind. From my point of view, our teamwork was very rewarding. I would also like to thank my friend for helping me write the code needed to run the experiment. Although we wanted to throw in the towel a number of times, it felt like a huge achievement to get the code right after many long nights.

Moderating predictive language processing: The role of age and reading skill

The notion that language comprehenders systematically predict upcoming linguistic content has reawakened over the past two decades. In fact, a large body of psycholinguistic experimental evidence suggests that anticipatory language processing, defined as the "preactivation […] of linguistic input before it is encountered by the language comprehender" (Huettig, 2015, p. 122), is one of the main reasons why interlocutors are able to communicate so effortlessly, accurately, and efficiently via the verbal medium. Predictions are said to serve the retrieval of semantic, lexical, and other information that people associate with words that are likely to follow (Federmeier et al., 2010). This, in turn, facilitates processing as soon as the expected targets are encountered. To illustrate, when people view a visual scene that contains a boy, a cake, and some toys, and they hear either "The boy will move the …" or "The boy will eat the …", they tend to shift their eye movements to the cake much sooner (i.e., before the acoustic onset of "cake") when they hear "eat" than when they hear "move". The cake represents the only edible object (Altman & Kamide, 1999), and so, listeners plausibly preactivated its mental representation more than those of the other objects (Altmann & Kamide, 2007). This way, eye movements may demonstrate that people had predicted upcoming linguistic input.

Whereas cues for (e.g., Frisson et al., 2005) and contents of (e.g., Federmeier et al., 2002) predictions have been investigated extensively, the mechanisms that underlie linguistic anticipatory processes have remained a relatively understudied topic. Moreover, moderators such as age and literacy, as well as their potential interactions, are not yet well-understood, even though they are likely to influence these processes among adults (Huettig, 2015). Apart from research on them being scarce in general, many psycholinguistic experiments have been built on WEIRD (i.e., Western, educated, industrialized, rich, and democratic, see Henrich et al., 2010) samples, mostly American undergraduates, who make up only a minority of the world's population and who are rather homogenous in terms of age and literacy. As a consequence, the generalizability of study results to older or less literate populations who are more psychologically common remains limited (Huettig, 2015; Pickering & Gambi, 2018).

Whereas different facets of literacy, such as reading skill or speaking ability, have been shown to positively relate to language-mediated anticipatory eye moments (Huettig & Brouwer, 2015; Huettig & Pickering, 2019; Mani & Huettig, 204), the influence of higher age appears to be more varied. Age is often seen as a proxy for experience, with higher age reflecting more life experience. Nevertheless, high age has also been associated with cognitive decline (Federmeier et al., 2002, 2010) and processing speed costs (Huettig & Janse, 2016; Steen-Baker et al., 2017). Consequently, the effects of lifelong experience on predictive language processing may depend on whether certain cognitive differences have been accounted for (Huettig & Janse, 2012; Ramscar et al., 2014). In terms of age-literacy interactions, it is possible that the negative effects of higher age on anticipatory linguistic processes, if existent at all, can be buffered by sophisticated language skills (Huettig & Janse, 2012; Mulder & Hulstijn, 2011; Steen-Baker et al., 2017). This study attempts to test this hypothesis.

Results obtained thus far do not provide unequivocal answers to theoretical questions such as whether higher age supports or undermines predictive language processing and how literacy may compensate for potential age-related deterioration in such processing. The more abstract question of whether anticipation of linguistic content is universal, thus, remains partially unanswered. Although it is generally accepted that language comprehension involves some kind of prediction (Kuperberg & Jaeger, 2016), it may be true that prediction is "not a fundamental principle of language processing and the human mind" (Huettig, 2015, p. 131). If so, interlocutors who systematically anticipate upcoming word candidates may show no communicative advantages over those who do not. In other words, there need not be a strong relationship between prediction and comprehension.

Due to problems with conventional methods used to investigate predictive language processing (e.g., spillover effects in the case of self-paced reading; Mitchell, 1984), the current study employs an online maze task. In this task, sentences appear as sequences of choices between two words, only one of which results in a correct sentence continuation (Witzel & Foster, 2014). This application of the maze task is relatively new, and so, the first goal of this study is to reproduce previous findings concerning the independent effects of age and literacy on anticipatory language processing. The second goal is to explore any meaningful age-literacy interactions; being less lab-bound than, for example, eye-tracking experiments (Boyce et al., 2020), the maze task allows for a significant number of non-WEIRD participants to be involved, thereby enabling this exploration. Ultimately, this study contributes to the growing body of research that re-evaluates whether prediction is in fact a fundamental principle of language processing (James, 1890), or, whether *everybody* predicts *all the time*.

Theoretical frame

Empirical evidence for predictive language processing

Experiments that exploit various contextual constraints on the word forms that precede high-cloze target words have been crucial in the process of collecting evidence for probabilistic language prediction (Huettig, 2015; Husband, 2021). In the upcoming paragraphs, general findings from the field will be presented. Traditionally, a few different experimental methodologies have been used to study linguistic anticipation processes online. Their strengths and weaknesses will be discussed as well.

Self-paced reading

In experiments that follow the self-paced moving window paradigm (Aaronson & Ferres, 2018; Aaronson & Scarborough, 1977; Just et al., 1982), participants read sentences that appear word-by-word on a screen, whereby they press a key to read each next word. Latencies between these key presses are recorded; as they relate positively to processing load (Kieras, 1978), they can be used to study the effects of, for example, syntactic ambiguity (e.g., MacDonald, 1994). Under certain circumstances, however, self-paced reading times can demonstrate that participants had preactivated linguistic input before they encountered it. Evidence of predictive processing then takes the form of numerical reading time (dis)advantages in the case of (un)predictable words. If a specific noun is highly expected due to the global constraints set in the broader discourse preceding it (i.e., when it is a high-cloze noun), readers are likely to strongly anticipate it. When they encounter a word that does not match the morphological or phonological features of the expected head noun, this will result in longer reading times on that word.

In line with this, Brothers et al. (2017) found that self-paced readers who encountered "an" when "necklace" was highly anticipated slowed down as soon as they read the incongruent determiner; participants who read the congruent "a" did not. The fact that this delay appears before the onset of the noun indicates that the reader must have predicted the subsequent noun (Otten & Van Berkum, 2009). Brothers et al. (2017) also demonstrated that top-down discursive factors on the message-level of a text representation form a greater conceptual basis for prediction during reading than do prime words. The extent to which readers anticipate upcoming lexical items from context, thus, appears to be influenced by global factors. This is corroborated by findings from eye-tracking studies (e.g., Otten & Van Berkum, 2008) and Brothers et al.'s (2017) first experiment, which measured the amplitude of two neural signatures of lexical prediction (i.e., the N400 and the frontal PNP).

Still, traditional self-paced reading is characterized by some methodological problems. First, only the total time spent on processing a single word is measured. This renders it difficult to distinguish early (nonconscious) stages of processing, which do not require any attention or cognitive control, from later (conscious) stages of processing, which may involve processes such as retrospective prediction verification (Luke & Christianson, 2013). Doing so with more continuous measures, such as electrophysiological indicators, is easier (for a review, see

Bendixen et al., 2012). Second, because prediction effects are likely to manifest themselves in both the target words and the words that follow (i.e., the spillover region), it remains difficult to segregate focal effects to specific words (Mitchell, 1984). As a consequence, spillover effects may problematize the interpretation of results. Finally, Hintz et al. (2015) witnessed prediction effects disappearing when participants engaged in self-paced reading only, as opposed to when the task set involved production, in this case naming objects. This is in line with the prediction-by-production theory (see, e.g., Martin et al., 2018; Schomers et al., 2015; Silbert et al., 2014), according to which language generation, which self-paced reading experiments typically lack, is a central component of language prediction (Bock, 1990; Horton, 2005).

Eye-tracking

Another way to study real-time language processing in language comprehension and production is by tracking participants' eye movements (Rayner, 1998). Participants' task is to either read text off a screen, or, in experiments that follow the visual world paradigm, to look at a screen and concurrently listen to utterances that comment upon the depicted scenes, which contain target objects and distractor objects. The activation of an object's mental representation increases the likelihood of a listener fixating that object's location (Altmann & Kamide, 2007), and so, language-mediated anticipatory eye movements can provide insight into when and how linguistic information that is presented auditorily is integrated with information retrieved from the display (Huettig et al., 2011).

In general, eye-tracking studies have yielded convincing evidence for predictive language processing (e.g., Curcic et al., 2019; Rommers et al., 2015, 2013; Staub & Clifton, 2006)[1]. On the one hand, reading studies have shown that word predictability influences the duration of fixations, whereby unpredictable words are read slower than highly anticipated ones (e.g., Ehrlich & Rayner, 1981; Rayner et al., 2011). Similarly, Kliegl et al. (2004) found low predictability to be strongly related to second-pass reading. In a review study, Staub (2015) concluded that "predictability effects in reading result from graded activation of potentially many words" (p. 311) and not the specific prediction of one or two words. This activation, in turn, facilitates very early (nonconscious) stages of lexical processing.

On the other hand, visual world studies have shown that listeners continuously make use of various cues to predict upcoming events when they are exposed to auditory input (for a review of the visual world paradigm, see Huettig et al., 2011). Participants have been observed

[1] It should still be noted that some eye-tracking studies found only small anticipation effects (e.g., Huettig & Guerra, 2019) and that effect sizes often depend on individual differences, for example in cognitive efficiency (Huettig & Janse, 2012) or reading skill (Steen-Baker et al., 2017).

to use selectional information conveyed by verbs (Altmann & Kamide, 1999) and case-marking (Kamide et al., 2003) to predict upcoming themes. In Altmann and Kamide's (1999) study, this manifested itself in relatively more anticipatory eye movements to visual objects that met certain semantic properties. For example, when participants heard the verb "eat", they quickly shifted their gaze to the only edible object on the screen. In 2007, the authors further demonstrated that participants make use of tense information to determine which referent is being referred to; when participants heard "The man has drunk...", they mostly looked at an empty glass, but when they heard "The man will drink...", they tended to fixate a full glass of beer. In sum, language-mediated eye movements reflect a dynamic process that is characterized by a continuous updating of mental representations of those events referred to by the visual environment and the heard words (Huettig, 2011).

Though eye-tracking has been indispensable in revealing how significant prediction is for language processing, problems with the method have been observed as well. In reading studies, participants tend to skip functional or short words because they process them in the fixation prior to the skip (Rayner, 2006; Rayner & Clifton, 2009). As upcoming linguistic information remains accessible in the parafoveal preview (Lai et al., 2013), spillover effects thus remain an issue. This renders extremely careful manipulation of the experimental conditions a prerequisite (Husband, 2021). A complication with visual world studies is that the pictured referents and the possible actions are much more limited in number than the actual visual world (Zhan, 2018). Apart from potentially creating task-specific strategies that lack generalizability, this closed-set problem may facilitate prediction by nudging participants, which is undesirable.

Electroencephalography

Another method that has served the field of predictive language processing to a substantial degree involves the measuring of event-related brain potentials (ERPs) through electroencephalography (EEG). Much of ERP research has focused on the N400, which is a negativity that peaks approximately 400 milliseconds after a stimulus, such as an unexpected target noun, has been presented (Rommers et al., 2013). Traditionally, the N400 has been assumed to indicate semantic processing (Kutas & Federmeier, 2000, 2011). In the case of an expectancy violation, extra processing is required to "suppress or revise the initial prediction" (Federmeier et al., 2010, p. 155). In line with this, differential N400 responses have been observed when participants encountered adjectives that mismatched highly expected head nouns in terms of grammatical gender (Otten and Van Berkum, 2008; Van Berkum et al., 2005).

Similar responses have been observed for gender-marked determiners in Spanish (e.g., Wicha et al., 2004) and the *a/an* contrast in English (DeLong et al., 2005; Husband, 2021).

In a large-scale study, however, Nieuwland et al. (2018) were unable to replicate DeLong et al.'s (2005) article-elicited N400 results. Instead, the authors found stronger N400 effects when native speakers encountered expected relative to unexpected nouns. In line with this, Luke and Christianson (2016) note that although a drop in the amplitude of the N400, typically observed when someone encounters highly predictable words, is associated with facilitated processing and lexical retrieval, the opposite need not be true. Federmeier et al. (2007) also observed no differential brain responses for unexpected words in highly versus weakly constraining contexts. Consequently, a spike in the N400 may not always be interpreted as a cost of a violated prediction.

A brain response that is generally accepted to indicate additional processing caused by a prediction error is the P600, which is a frontal late-positive component that emerges approximately 600 milliseconds after orthographic anomalies are presented (Bulkes et al., 2020). Studies have shown that the P600 modulates higher-level processes such as structural integration (e.g., Vissers et al., 2006, 2008) and that it reflects a process of language monitoring. Correspondingly, a spike in the P600 can be observed when a deviation from the expected input, for example in the form of a morphosyntactic violation (Coulson et al., 1998), is encountered; the mismatch then triggers more effortful integrative processing.

Due to its high temporal resolution, ERP research has been crucial in studying the continuous effects of incongruities that occur between adjacent words (Husband, 2021). Doing so with more traditional methods, such as self-paced reading, remains difficult. Still, EEG analysis is very expensive, laborious for the researcher, and time-consuming (Ansari, 2018). If a cheaper and faster method to study early cues in prediction were to be developed and tested, this would be of substantial scientific value (Boyce et al., 2020).

Maze task

To address some of the shortcomings of the aforementioned methods, the maze task has been put forward as an option (Forster et al., 2009). Like self-paced reading, this alternative methodology measures word-by-word reading times online. Sentences are presented as sequences of choices between two words, one being a distractor and the other the grammatical continuation of the sentence (Witzel & Foster, 2014). Participants' task is to build a grammatical sentence by choosing the correct continuations as fast as they can. This process involves a number of incremental steps, which require effort and time. Examples are (a) identifying the two candidate words, (b) deciding the extent to which they match the context,

(c) determining which of the candidates is correct, (d) "initiating and completing motor actions" (i.e., pressing the key), and (e) integrating the chosen word into sentence context (Boyce et al., 2020, p. 3). Some steps, such as (d) and (e), may be executed simultaneously.

It may be noted that the maze task process resembles that of normal and self-paced reading. Here, too, readers must identify words, integrate them into context, and determine whether they have been integrated well enough for processing to continue (Boyce et al., 2020). This, in turn, may trigger a saccade or result in a key press. However, unlike normal and self-paced reading, the necessity to choose between candidates in a maze task induces strong incremental processing: if individuals are unable to integrate either of the candidates (step b above), they cannot accurately determine which way the sentence should proceed. Consequently, low-cloze target words that are hard to integrate will result in longer reading times on those words "with minimal spillover to words" that follow (Boyce et al., 2020, p. 3). Due to this improved localization, the maze task takes away some of the complexities in interpreting eye-tracking or self-paced reading data. Different versions of the task (i.e., grammaticality maze and lexicality maze) have already been shown to provide more robust evidence of language processing difficulty than moving-window reading and eye-tracking experiments have (Witzel et al., 2012).

Additionally, the maze task allows data to be collected through online crowdsourcing services, such as Mechanical Turk (Paolacci & Chandler, 2014). This makes the task much cheaper and easier to implement than other methods. As more empirical support for its potential to reproduce findings from self-paced reading and eye-tracking experiments is being found (Forster et al., 2009; Husband, 2021; Witzel & Forstel, 2014), the maze task may prove to be a worthy alternative to investigate predictive language processing when higher-fidelity resources are unavailable. Moreover, being less tied to the lab than ERP and eye-tracking experiments, the maze task enables a more diverse (non-university student) sample to be recruited (Boyce et al., 2020). This, in turn, facilitates the study of factors that plausibly moderate anticipatory linguistic processes, such as (high) age and (poor) language skill.

Moderators

Factors that potentially moderate predictive language processes have, historically, received little scientific attention. This is strange given the fact that individual differences in experience or cognitive ability are likely to influence these processes (Huettig & Janse, 2012, 2016), either in isolation or in combination with contextual differences, such as whether spoken language relates to abstract events or some co-present visual scene (Huettig, 2015). If they do, revising the models of anticipatory language processing to take moderators into account may be a logical next step. In the coming paragraphs, experimental findings on the effects of two moderators (i.e., age and reading skill) will be discussed in detail.

Effects of age

Studies on the effects of age on anticipatory language processing have found that older adults are slower readers (e.g., Rayner & Clifton, 2009), show more regressive eye movements (Kliegl et al., 2004), and integrate final words into sentences more slowly and to a lesser extent than younger adults (Federmeier & Kutas, 2005; Huang et al., 2012). The general conclusion drawn from these results was that older individuals rely less on anticipatory processes, and so, they tend to display less prediction-related advantages during sentence processing. It is now believed that the cognitive processes underlying these individual differences are a decline in working memory capacity and a decrease in processing speed (Huettig & Janse, 2012, 2016).

To make quick predictions, sufficient cognitive resources are required (Ito et al., 2018). As older adults have less of these resources available to them (Lindenberger et al., 2008; Poon et al., 1992), this may result in anticipatory modes of comprehension being engaged to a lesser extent and slowing down. In an ERP study by Wlotko et al. (2012), this manifested itself in older adults revisiting contextual information more often when interpreting message-level meanings; even for highly predictable words, they were more likely to show the ERP response associated with the process of revisiting this information (i.e., a left-lateralized frontal negativity). For highly unexpected orthography, only the younger adults showed an enhanced positive brain response associated with the consequences of disconfirmed predictions. Generally, prediction thus seems to be "susceptible to age-related deterioration and can be associated with processing costs" (Federmeier, 2007, p. 491).

Still, normal aging has also been found to have a marginally positive effect on anticipatory processing. In an eye-tracking experiment, Huettig and Janse (2012) found higher age to be associated with poorer cognitive efficiency and slower processing speed. However, after having accounted for these variables, higher age appeared to enhance predictive processing. This is not that surprising given the fact that age and life experience go hand in hand; both increase the amount and complexity of the knowledge an individual possesses, which, according to mathematical accounts, will inevitably have processing (speed) costs (Shannon, 1948). In line with this, Salthouse (1991) found age differences in fluid aspects of cognition to be mediated by age-related reductions in working memory, which were likely to be influenced by age-induced reductions in processing speed. What follows is that it may be useful to segregate data variance caused by differences in speed or memory capacity from data variance caused by age alone.

Yet, when comprehension is the goal, the amount and complexity of the information that is to be processed by individuals may exert a greater influence on their processing speed than the amount and complexity of the information they possess. If so, being older may still do more harm than good (e.g., Eckert, 2011; Verhaeghen et al., 2003); even though it is possible that the excessive focus on cognitive decline has concealed the advantages higher age might have in everyday language contexts (cf. Ramscar et al., 2014), younger individuals generally anticipate linguistic input more habitually and to a stronger degree, whereas older individuals show reduced and delayed prediction effects (e.g., Federmeier et al., 2010, 2012; Wlotko, 2012). Still, the exact details about how linguistic anticipation varies across adulthood are yet to be explored in further detail.

Effects of reading skill

Some recent research points in the direction of prediction and literacy being strongly intertwined. Although the exact definition of literacy varies substantially across studies, it is generally referred to as the ability to read or write and often taken as a proxy for language experience (Choi et al., 2018). Being a multifaceted concept, literacy includes a wide variety of skills, such as numeracy, oral skills, and reading skills (Burnett et al., 2016), all of which can be measured in various ways. Reading ability, specifically, appears to predict academic achievement at university level (Pluck, 2018). Studies investigating how linguistic anticipation varies as a function of literacy components often derive their data from less conventional participant samples, such as adults with dyslexia or children (Huettig, 2015).

In an eye-tracking study, for example, Huettig and Brouwer (2015), compared adults with dyslexia to adults that had no reading disabilities. The authors found anticipatory spoken language processing to be influenced by differences in reading ability. Reading ability was composed of word reading skill, measured with the “Een Minuut Test” (One-Minute-Test, Brus & Voeten, 1999), and pseudoword reading skill, measured with the Klepel Test (Van den Bos et al., 1994). In both tasks, participants read out loud as many increasingly difficult (pseudo)words as they can in a limited amount of time. Participants’ scores on the two tasks positively correlated with their anticipatory eye movements; “adults with dyslexia anticipated the target objects […] much later than the controls” (p. 97).

Similarly, in a visual-world study, Mishra et al. (2012) found that Indian low literates shifted their eye gaze to target objects much later than Indian high literates, in this case postgraduate university students. Although the former group was able to read and write, they only pronounced about 6.6% of experimental words that varied in syllabic complexity correctly; high literates did so for 98.1% of the words. The finding that anticipatory eye movements are positively correlated with word reading scores is in line with experimental research on illiteracy among adults (for a review, see Huettig & Mishra, 2014).

In an eye-tracking experiment among toddlers, Mani and Huettig (2012) found predictive ability to be positively correlated with productive vocabulary size (i.e., the number of words children both understood and produced, as reported by their parents). Whereas skilled producers showed strong evidence of predicting upcoming linguistic input, low producers showed none at all. Interestingly, there was no correlation between toddlers' comprehension vocabulary size (i.e., the number of words they understood but did not produce) and their predictive ability. So, in line with the prediction-by-production theory (e.g., Martin et al., 2018), prediction appeared to depend more on production skills than on comprehension skills. In a follow-up study, specifically children's word reading skill, measured with the standardized "Salzburger Lese und Rechtschreibtest II" (Salzburger Reading and Spelling Test II), predicted anticipation of spoken language (Mani & Huettig, 2014). This aligns with the research on comprehension processes used by proactive readers, who are characterized by long saccades and many regressions, versus conservative readers, who are characterized by short saccades and few regressions (Koornneef & Mulders, 2017), or low- and high-comprehending children in general (see Kraal et al., 2018).

Language-mediated anticipatory eye movements also appear to be positively related to individuals' vocabulary size and verbal fluency. In a study by Rommers et al. (2015), vocabulary size was assessed using the nonverbal Peabody Picture Vocabulary Test (Dunn & Dunn, 1997). Here, participants indicated which of four displayed pictures corresponded to an orally presented word. Verbal fluency was measured with a category fluency task (Delis et al., 2001), whereby participants named as many members of a category (e.g., vegetables) as they could in one minute. Performance on these tasks connects to anticipatory language comprehension (see, e.g., Borovsky et al., 2012; Federmeier et al., 2002, 2010). The authors found that, before stimulus onset, even objects that resembled the target objects only in shape (e.g., a tomato when a moon is the target) were fixated more than unrelated distractor objects (e.g., a bowl). More importantly, however, individuals who performed well on the fluency and vocabulary tasks displayed this pattern to a larger degree than those performing poorly, leading the authors to conclude that the preactivation of visual representations is modulated by literacy.

Moderator interactions

Eye-tracking studies have yielded convincing evidence for larger predictability effects in younger versus older adults, as well as high versus low literates. However, only a few studies have looked into the possible interactions between age- and literacy-related effects on linguistic anticipation, with varying results. Consequently, it remains unclear whether and how age-related prediction effects interact with the literacy aspects that tend to improve with age, such as verbal ability (Cheimariou et al., 2021).

Borovsky et al. (2012) contrasted anticipatory eye movements from children aged 3 to 10 with young adults aged 18 to 28 to assess the role of vocabulary size, which has been associated with comprehension speed in visual tasks (e.g., Fernald et al., 2006), in lexical processing. Vocabulary size was measured using the Peabody Picture Vocabulary Test (Dunn & Dunn, 2007) and the Sentence Completion Subtest of the Comprehensive Assessment of Spoken Language (Carrow-Woolfolk, 1999), which requires subjects to produce a fitting final word to an orally presented sentence. Both age groups showed clear signs of predictive processing. They were also equally quick to anticipate upcoming linguistic events. But, when age was accounted for in the analyses, individuals with larger vocabularies responded faster than those with smaller vocabularies. Even though all participants were relatively young, the authors concluded that not age but vocabulary skill influences predictive language processing. Studies with older subjects, however, point in different directions and hint at more complex age-literacy interactions (cf. Mulder & Hulstijn, 2011); their findings will be discussed next.

In an eye-tracking study, Steen-Baker et al. (2017) focused on how the facilitation of sentence context, which "reflects the interaction between the build-up of message-level semantics and lexical processing" (p. 460), is affected by age and reading skill. Reading skill was measured using the Slosson oral reading task (Slosson & Nicholson, 1990) – a word-identification task that required subjects to pronounce certain words as accurately as they could in unlimited time – and a Woodcock-Johnson reading fluency task (Schrank et al., 2014), which required subjects to establish the truth value of as many sentences (e.g., "De lucht is blauw", The sky is blue) as they could within three minutes. The latter requires processes such as decoding, lexical access, and semantic integration. In both cases, the total number of correct items was mapped to a grade-level estimation of reading ability.

In general, younger participants (i.e., those aged 16 to 40) displayed advantages in psychomotor speed relative to middle-aged participants (i.e., those ages 41 to 64). The same was true when high literates (i.e., those scoring above the median reading level of grade 9.5) were contrasted with low literates (i.e., those scoring below the median). The high-literacy group also showed advantages in, among other things, phonological awareness and naming fluency. Comprehension performance was generally high and increased with both age and reading skill. It was, however, particularly low among less skilled readers in the case of low-cloze items. This may indicate that, regardless of age, low literates have more difficulty with forming mental representations when semantic features are not strongly co-activated (cf. Ng et al., 2017; Steen-Baker et al., 2017).

The authors found no evidence of slowing in reading processes at least into midlife, but an interaction effect did reveal that, as age increased, high literates showed less age-related slowing, whereas low literates showed more. This corresponds with findings from related fields; in a word-learning study, for example, Meijer et al. (2008) found that educational attainment, which correlates with literacy (Roser & Ortiz-Ospina, 2016), can compensate for age-related deteriorations in verbal learning. In other words, building up literacy experience over the years may automate low-level reading processes. This, in turn, allows efficient lexical processes to continuously develop into adulthood and reading skills to potentially operate as a buffer when predictive ability becomes compromised (Steen-Baker et al., 2017).

Similar results were obtained in an ERP experiment by Federmeier et al. (2010). Older adults failed to demonstrate anticipatory processing during comprehension even though working memory demands were minimized and the task mainly relied on "the kind of knowledge shown to be preserved or augmented with age" (p. 159), suggesting there is something quite essential about how age influences language processing. More interestingly, however, the older adults who performed well on a category fluency task that indexed verbal fluency showed a brain response typically displayed by younger adults (i.e., a sustained frontal positivity to low-typicality exemplars, indicative of the processing cost of an expectancy violation) more than their poorly-performing peers did (cf. Federmeier et al., 2002). So, age-related debilitation in predictive processing was again mitigated, at least to a certain degree, by a form of literacy.

In another experiment, Federmeier et al. (2010) showed that the reduced semantic processing effects were not caused by simple slowing of language production mechanisms; even though older adults produced language as quickly as (and more accurately than) younger adults, they generally showed less signs of using generative strategies during comprehension (i.e., they were less likely to generate expectations about potentially upcoming targets in the first experiment). This corresponds to the finding that older adults are less likely to systematically use top-down mechanisms related to language generation (e.g., Logan et al., 2002) – unless there are clear instructions to do so. Indeed, Federmeier et al.'s (2010) second experiment demanded overt production from its subjects. The authors concluded that if active involvement of the top-down circuitry is encouraged, older adults may in fact display preserved functioning of those predictive processes younger people automatically and systematically engage in, even if age in general seems to reduce the likelihood of these processes being activated.

To some extent, all these studies attempt to answer the same fundamental question of whether people *always* predict (Huettig, 2015). A popular view is that our brains are prediction machines (e.g., Clark, 2013). A growing body of evidence (e.g., Borovsky et al., 2012; Federmeier et al., 2010; Steen-Baker et al., 2017), however, insinuates that not everybody predicts, or at least not all the time; the strong link between task proficiency and anticipatory language processing certainly seems to suggest this. What follows is that it may be justified to add more nuance to the assertion that prediction is a quintessential principle of the human mind (James, 1890).

The current study

If prediction is as central to natural language processing as the empirical studies based on undergraduates suggest, then *all* proficient language users should be able to demonstrate it in some way (Mishra et al., 2012), regardless of their age or reading skill. In this study, Dutch speaking adults of various ages and reading skill levels executed an online maze task. The goals of the study were to investigate whether less skilled readers could approximate the predictive processing of highly skilled readers as they gathered more life experience, whether signs of age-related cognitive decline persisted after certain individual differences had been accounted for, and whether advanced reading skills could mitigate potential impediments in predictive language processing caused by higher age.

Participants performed an online grammaticality maze task (Boyce et al., 2020), in which they built two-sentence stories as fast as they could by repeatedly choosing the next word in the sentence out of two word options, only one of which would fit grammatically (Witzel & Foster, 2014). If participants chose the wrong (ungrammatical) word, the trial would automatically end and the first word of the next story would appear. To study prediction effects, the Dutch gender-marked articles “de” and “het” were used (cf. Fleur et al., 2020). Reading times on the articles were measured; if, based on the preceding context, participants had predicted a specific noun and its corresponding article, reading times were expected to become longer as soon as an incongruent article appeared as the only viable option for the next word. So, article reading times indexed the extent to which participants had difficulty with integrating (un)expected linguistic information into sentence context; increased predictive processing was signaled by relatively longer reading times on unpredictable versus predictable articles (i.e., a relatively larger delay)[2]. Investigating prediction effects by means of a maze task is relatively new and unconventional. Therefore, the first goal of this study was to replicate previous findings concerning the independent modulations of age and reading skill. Correspondingly, the first research question was: How do age and reading skill independently modulate predictive language processing?

Higher age has been associated with decreased processing speed, loss of memory capacity, and increased difficulty with information integration (Federmeier, 2007; Federmeier & Kutas, 2005; Huang et al, 2012; Huettig & Janse, 2016). Because the magnitude of older individuals' expectancy violation tends to be smaller (cf. Federmeier et al., 2010) and failed predictions may be more likely to escape their attention, older adults typically show less evidence of predictive language processing than do younger adults (Federmeier et al., 2002, 2010). As a decrease in predictive processing should translate to a relative speed advantage in reading time, older individuals were expected to slow down to a lesser extent than younger individuals when they encountered articles that disconfirmed predictions. Consequently, hypothesis 1 was as follows: Middle-aged adults show decreased predictive processing (as indexed by relatively shorter reading times on unpredictable than on predictable articles) compared to younger adults.

Eye-tracking studies that investigated how different components of literacy are connected to predictive ability have yielded strong evidence for a positive relationship between the two (e.g., Huettig & Brouwer, 2015; Mani & Huettig, 2012, 2014; Mishra et al., 2012; Rommers et al., 2015; Steen-Baker et al., 2017). Skilled readers look at target objects sooner than less skilled readers due to stronger activation of the anticipated objects' mental representation. This enables them to make fast predictions. As a consequence, less skilled readers, compared to skilled readers, were expected do display relatively shorter reading times on unpredictable than on predictable articles. Hypothesis 2, therefore, was as follows: Less skilled readers show decreased predictive processing (as indexed by relatively shorter reading times on unpredictable than on predictable articles) compared to highly skilled readers.

[2] Although this resembles how differential brain responses can indicate the magnitude of a prediction error (Kutas & Federmeier, 2011) or facilitated processing (Van Petten & Luka, 2012), it is worthy to note that the relation between reading times and prediction is more ambiguous and prone to factors such as word frequency and length (Hyönä & Olsen, 1995) as well as the presence of semantically similar words (Roland, et al., 2012). Consequently, these factors ought to be kept constant across experimental conditions (cf. Otten & Van Berkum 2009). Still, a large range of predictability has been found to have continuous and graded effects on reading time (Luke & Christianson, 2016) and these effects have even been observed in moderate- and low-constraining contexts (Rayner & Well, 1996). Therefore, in the current study, increased predictive processing was indexed by a relative speed disadvantage in reading time.

As there is a possibility that high literacy is able to counteract age-related deterioration in predictive processing, the second goal of this study was to investigate the interactions between the aforementioned age- and reading skill-related effects on predictive processing. The corresponding research question was: How do age and reading skill interactively modulate predictive language processing? Only a few studies have looked into moderator interactions within the predictive field (Borovsky et al., 2012; Federmeier et al., 2010; Steen-Baker et al., 2017). Their findings resemble those from related fields (e.g., Meijer et al., 2008); in general, language skill seems to be able to operate as a buffer for potential cognitive impairments caused by higher age. There are even signs of literacy experience allowing language-related processes to continue to grow and mature over the span of a lifetime. Consequently, hypothesis 3 was as follows: As age increases, less skilled readers show stronger deterioration in predictive processing (as indexed by relatively shorter reading times on unpredictable than on predictable articles) compared to highly skilled readers.

Some evidence suggests that the negative effects of higher age on predictive processing (H1) evaporate when processing speed costs and memory capacity loss, which are thought to be inevitable consequences of high age (Ramscar et al., 2014; Shannon, 1948), are accounted for in statistical analyses (Huettig, 2015). If so, having lifelong language experience may not be as costly as was previously thought. What follows is that older individuals' speed advantage could become smaller or even fully disappear (cf. Huettig & Janse, 2012). In contrast, the negative effects of low reading skill on predictive processing (H2) were not expected to disappear when the aforementioned cognitive debilitations are accounted for. Hypothesis 4, thus, was as follows: When differences in processing speed and memory capacity have been accounted for, age no longer predicts variance in predictive processing (as indexed by the difference between reading times on unpredictable and predictable articles) in addition to reading skill.[3]

Method

Design

To investigate whether potential age-related deterioration in predictive language processing can be mitigated by sophisticated reading skills, adults of varying age and reading skill levels executed an online maze task in which they built two-sentence stories word-by-

[3] It may be noted that accepting this hypothesis would essentially translate to confirming the null hypothesis. Strictly speaking, this is not allowed within the frequentist inference paradigm. We acknowledge this statistical limitation and encourage future studies to employ more advanced statistical techniques, such as Bayesian linear regression analyses.

word. All stories were highly constraining and differed in terms of the cloze value (high versus low) of their final two target words, which were an article and a noun. High-cloze endings were very predictable and low-cloze endings were very unpredictable. Reading times on the two target words were recorded. The earliest signs of prediction were to be found on the articles; predictive processing was indexed by relatively longer reading times on unpredictable versus predictable articles (the greater the difference, the greater the ability). Closing nouns were considered spillover regions; reading times on them were recorded to explore the maze task's sensitivity to spillover effects (Boyce et al., 2020; Forster et al., 2009; Witzel et al., 2012). The study had a 2 (age group: younger versus middle-aged) x 2 (reading skill group: low versus high) x 2 (predictability: low versus high) mixed design. Age group and reading skill group were the between-subjects factors and predictability was the within-subjects factor. The average reading time per participant and predictability condition (in milliseconds) was the dependent measure. Participants indicated their age prior to the experiment. Reading skill was measured using a lexical decision task and a Woodcock-Johnson reading fluency task. To assess whether age-related effects could be explained by differences in processing speed and memory capacity, participants' performance was assessed using four cognitive tasks.

Participants

102 adults concluded the online experiment. However, due to some scores on the cognitive battery failing to meet the cutoff criterion and a number of participants being unable to finish sufficient maze task trials, the data of 76 adults were analyzed (39.5% female). Age ranged between 18 and 65 years (M = 36.79, SD = 15.70). Reading skill differed substantially among the participants, with scores on the Woodcock-Johnson reading fluency task varying between 37 and 95, scores on the lexical decision task ranging between 20 and 67, and total reading skill scores ranging between 0.43 and 1.81. Age and reading skill were treated as categorical variables. Following the methodology of Steen-Baker et al. (2017), participants aged 18 to 40 were classified as younger and participants aged 41 to 65 were classified as middle-aged (40 was adopted as an arbitrary cutoff point). Participants who scored below the median reading skill score of 1.11 were classified as less skilled readers; the remaining participants were classified as highly skilled readers (cf. Steen-Baker et al., 2017). Participant characteristics are presented in Table 1 on the next page.

Table 1

Participant characteristics

	Low reading skill				High reading skill										
	Y (n = 22)		MA (n = 16)		Y (n = 26)		MA (n = 12)		Total (n = 76)		Correlations				
Measures	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*Age*	*Ed. level*	*PS*	*MC*	*RS*
Age	24.50	4.58	56.88	5.83	26.35	4.78	55.17	5.15	36.79	15.70	.	-.27	**-.39****	**-.55****	.05
Ed. level	5.95	0.95	5.06	1.00	6.12	0.99	5.50	1.17	5.75	1.07	-.27	.	**.32***	**.40****	.16
PS	25.41	3.14	23.06	2.54	29.69	4.26	24.50	3.94	26.24	4.40	**-.39****	**.32***	.	**.48****	**.41****
LC	11.27	2.14	10.38	1.26	13.50	2.40	10.17	2.41	11.67	2.51	**-.39***	.20	**.88****	**.41****	**.24**
PC	14.14	1.73	12.96	2.09	16.19	2.53	14.33	2.23	14.57	2.51	**-.30***	**.37***	**.88****	**.43****	**.47****
MC	1.53	0.31	1.18	0.27	1.70	0.15	1.25	0.34	1.47	0.33	**-.55****	**.40****	**.48****	.	.22
NWR	15.82	3.28	13.13	3.22	17.81	2.19	14.00	5.19	15.64	3.75	**-.40****	.48**	**.45****	**.82****	**.23**
BRDS	5.18	1.33	3.69	1.08	5.69	0.88	3.83	1.64	4.83	1.46	**-.52****	.20	**.37***	**.86****	.15
RS	0.92	0.21	1.00	0.08	1.38	0.19	1.38	0.24	1.17	0.29	.05	.16	**.41****	.22	.
WJ	57.55	8.49	53.69	4.50	79.35	12.02	74.17	12.69	66.82	14.87	-.18	.22	**.57****	**.41****	**.80****
LD	34.86	5.54	36.69	5.63	45.73	8.23	49.83	9.22	41.33	9.23	.12	.18	**.34***	.20	**.79****

Note. Correlation coefficients are Pearson's r (two-tailed). Bold items are significant at $p < .05$. Bold items with * are significant at $p < .01$. Bold items with ** are significant at $p < .001$. Y = younger; MA = middle-aged; PS = Processing Speed; LC = Letter Comparison; PC = Pattern Comparison; MC = Memory Capacity; NWR = Nonword Repetition; BRDS = Backward Recall Digit Span; RS = Reading Skill; WJ = Woodcock-Johnson Reading Fluency; LD = Lexical Decision.

Reading skill groups did not differ significantly from each other in terms of age, educational attainment, processing speed, or memory capacity ($t < 1$, $p > .636$ in all cases). Likewise, age groups did not differ significantly from each other in terms of reading skill ($t = .01$, $p = .994$). However, advantages were observed for the younger group in educational attainment ($t(74) = 3.30$, $p = .001$), memory capacity ($t(74) = 6.51$, $p < .001$), and processing speed ($t(74) = 4.30$, $p < .001$). No interactions between age and reading skill reached significance ($F < 1$ in all but one case). All participants were native speakers of Dutch, as reported by themselves. Still, 44 (57.9%) used languages other than Dutch on a daily basis (41 English, 7 German, 2 Afrikaans, 2 French, and 1 Spanish). Multilingualism occurred mostly among middle-aged participants with low reading skill ($F(3, 75) = 3.40$, $p = .022$). Only three (3.9%) participants had dyslexia. As they were spread across three of the four groups, they were included in the analyses.

Cognitive battery

Processing speed

Participants' psychomotor speed was assessed using online versions of the paper-and-pencil letter comparison task and pattern comparison task (Earles & Salthouse, 1995; Salthouse, 1991, 1993, 1996). In the former, participants determined of as many strings of consonants as they could within thirty seconds whether they were identical or not. In the latter, they made similar judgments on pairs of line drawings (see Figure 1). Participants did so by pressing the keys that corresponded to the labels "hetzelfde" (the same, D) or "anders" (different, J). This was a two-alternative forced choice (cf. Huettig & Janse, 2016).

Figure 1

Example trials of the letter comparison task (left) and the pattern comparison task (right)

Strings got longer and drawings increased in complexity as trials went by. Scores on the letter comparison task ($M = 11.67$, $SD = 2.51$, range 5-19) correlated significantly with scores on the pattern comparison task ($M = 14.56$, $SD = 2.51$, range 9-21), $r(74) = .54$, $p < .001$. Together, they formed a reliable measure of processing speed, α = .697. Participants' final processing speed score ($M = 26.24$, $SD = 4.40$, range 18-37) corresponded to the total number of correctly evaluated pairs on both tasks (max. 44). Higher scores reflected faster processing.

To familiarize participants with the tasks, they executed two simple practice trials (one being the same and one being different) before each task. These did not count towards the final score.

Memory capacity

A visual version of the auditory nonword repetition task (NWR, Gathercole & Baddeley, 1996; Thorn & Gathercole, 1999) was used to assess participants' short-term memory capacity. Traditionally, nonwords that obey the rules of a particular language are presented over headphones. After having heard a nonword, participants' task is to repeat it as accurately as possible. The current experiment was to be executed in the absence of a researcher, and so, to avoid technical and ethical issues with participant-generated voice recordings, a visual variant of the NWR was created. Instead of hearing them, participants saw nonwords that were orthographically legal in Dutch (De Jong & Van der Leij, 1999) on their screen. Nonwords appeared for two seconds, after which participants were tasked to type them in exactly as they appeared. For this, they had five seconds; the next nonword would then appear automatically. Nonwords ($N = 20$) differed in syllabic length from two to five syllables and increased in length as trials went by. The total number of correctly repeated nonwords ($M = 15.64$, $SD = 3.75$, range 4-20) indexed participants' short-term memory capacity. A higher score represented greater short-term memory capacity.

Working memory performance was assessed using a backwards recall digit span task (BRDS, a component of the Wechsler adult intelligence test, 2004). The BRDS requires participants to manipulate, rather than store and reproduce, the presented stimuli. Consequently, scores on the task are indicative of working memory capacity, rather than short-term memory capacity (Baddeley, 2006). In the online version of the task, participants saw a series of digits that appeared one-by-one on their screen (e.g., 2 9 7). Each digit appeared for one second. Inter-digit time was also one second. After the sequence was presented, participants typed in the digits in reverse order (e.g., 7 9 2). For this, they had unlimited time. To familiarize participants with the task, they performed one three-digit practice trial beforehand. The experimental trials that followed increased in sequence length from two to eight digits (cf. Huettig & Janse, 2012, 2016). Answer boxes were fitted to the number of digits per trial. The total number of correctly recalled digit sequences ($M = 4.83$, $SD = 1.46$, range 1-7) indicated participants' working memory performance, with higher scores reflecting better performance. Scores on the NWR and BRDS correlated moderately with each other, $r(74) = .42$, $p < .001$, but did not form a very reliable measure of memory capacity, $\alpha = .441$. As the two tasks contained different numbers of items, the sum of the relative performance on both tasks (min = 0, max = 2) indexed overall memory capacity. Higher scores indicated greater memory capacity. On average, this was 1.47 (*SD* = 0.33, range 0.49-2.00).

Reading skill

In the Woodcock-Johnson (WJ) reading fluency task, participants were presented with a list of simple sentences (e.g., "De lucht is blauw", The sky is blue) and tasked to establish the truth value of as many sentences as they could within two minutes (Schrank et al., 2014). They did so by pressing the keys corresponding to the labels "waar" (true, D) and "niet waar" (false, J). In total, there were 98 sentences. Scores on the WJ indicated participants' fluency in executing coordinated processes needed to understand simple sentences, such as "decoding, lexical access, parsing, [and] semantic integration" (Steen-Baker et al., 2017, p. 465). Participant scores corresponded to the absolute number of correctly evaluated sentences (*M* = 66.82, *SD* = 14.87, range 37-95), with higher scores reflecting more reading fluency.

In the online time-limited lexical decision task (LDT), participants decided of as many strings of letters as they could within one minute whether they formed correctly spelled words in Dutch. They did so by pressing the keys corresponding to the labels "echt woord" (real word, D) and "nep woord" (fake word, J). In total, forty nonwords, which were orthographically and phonologically legal in Dutch, were intermixed with forty real words. Performance on the LDT has been shown to correlate with scores on standardized measures of reading ability (Moreno & Van Orden, 2001; Yeatman et al., 2021). Consequently, the LDT is thought to be an accurate and reliable measure of reading ability that can be self-delivered through a web browser. Participant scores corresponded to the total number of correctly evaluated (non)words (*M* = 41.33, *SD* = 9.23, range 20-67). Scores on the WJ and the LD correlated significantly with each other, $r(74) = .65$, $p < .001$, and formed a reliable measure of reading skill, $\alpha = .736$. Again, because the tasks contained different numbers of items, the sum of the relative performance on both tasks (min = 0, max = 2) indexed overall reading skill, with higher scores reflecting greater reading skill. On average, this score was 1.17 (*SD* = 0.29, range 0.43-1.81).

Material and instrumentation

Stimuli consisted of a subset (*N* = 32) of Fleur et al.'s (2020) pretested items (accessible at https://osf.io/6drcy/). Only items were selected that had definite determiners, cloze values > 0.90 for high-cloze target words, and cloze values < 0.15 for low-cloze target words. Each item was a short story frame. Following Fleur et al.'s (2020) pretest, spillover regions were removed from the original stimuli. Each story frame consisted of two sentences, which were presented as sequences of word options. The first sentence served to set up a broader discourse context and the second sentence had either a very predictable or a very unpredictable ending, as indexed by the cloze value of the two final target words. The first word of each sentence was given. Participants made sequential forced choices between a correct sentence continuation and a contextually and grammatically inappropriate distractor (see Table 2). For each correct word, a distractor of similar length and word frequency was selected. Although Boyce et al. (2020) recently developed a version of the maze task called A(auto)-Maze that automatically generates distractors, this version is currently still limited to English. As a consequence, distractor material was matched manually using the Subtlex-NL database (Keuleers et al., 2010).

Table 2

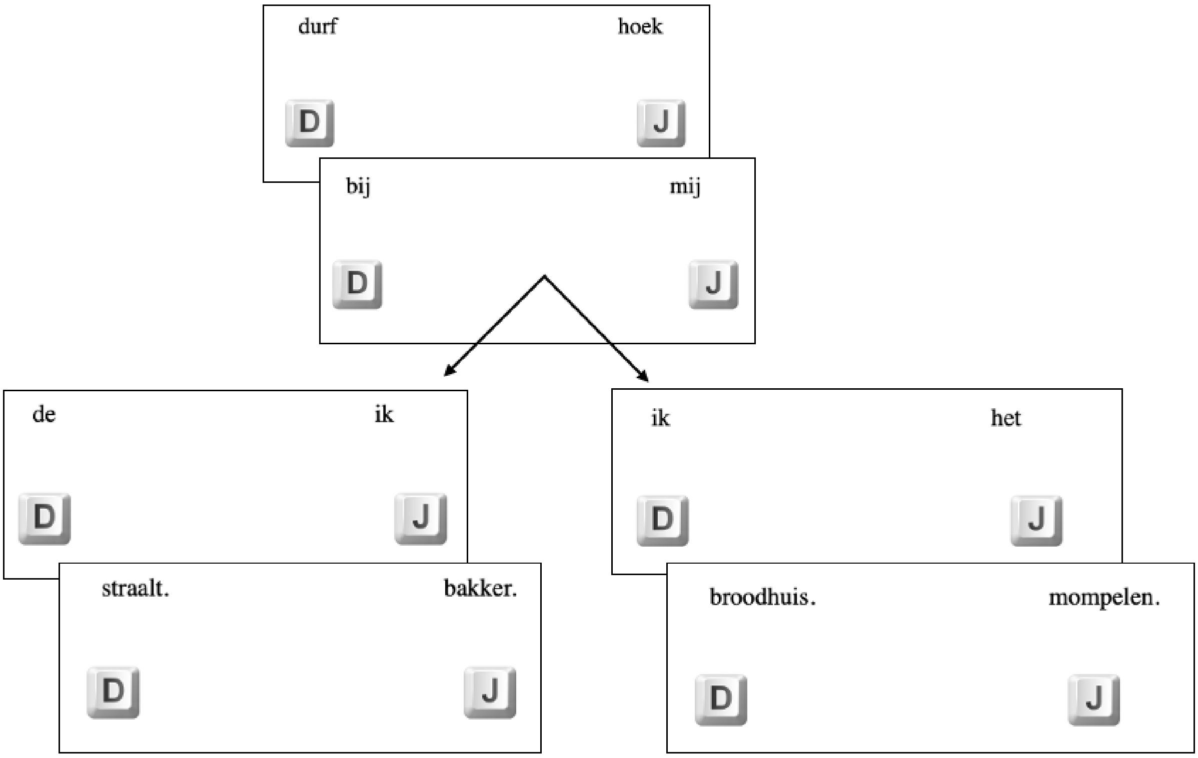
Example of Dutch two-sentence story in each predictability condition, including distractors per word choice (numbered) and loose English translations

		Predictability of critical noun phrase	
	Discourse context	Predictable	Unpredictable
Correct continuation	Het[1] is[2] zondagochtend.[3] De[4] gehele[5] gelovige[6] familie[7] gaat[8] zoals[9] altijd[10] naar[11]	de[12] kerk.[13]	het[12] gebedshuis.[13]
	It[1] is[2] Sunday morning.[3] The[4] whole[5] religious[6] family[7] goes[8], as[9] always[10], to[11]	*the[12] church.[13]*	*the[12] (house of prayer.)[13]*
Distractor	(.)[1] wat[2] geldproblemen.[3] (.)[4] beschouw[5] lammetje[6] volgens[7] jullie[8] bedankt[9] maken[10] moet[11]	op[12] belde.[13]	van[12] abstracter.[13]
	(.)[1] what[2] (financial issues.)[3] (.)[4] (consider)[5] (small lamb)[6] (according to)[7] you[8] thanks[9] (to make)[10] (have to)[11]	*on[12] phoned.[13]*	*from[12] abstracter.[13]*

In building each story, participants pressed D on their keyboards to continue with the word on the left and J to continue with the word on the right; word positions were randomized (see Figure 2 on the next page). Reading times between key presses were recorded. If a distractor was selected, the corresponding trial would automatically end. Participants were then neutrally informed, so as not to demotivate them, that the next trial would begin as soon as they clicked on a button. They would then be presented with the first word choice of the next trial. One practice trial was executed beforehand. All stimuli, including distractors, can be found in the Appendix.

Figure 2

Example of a word choice sequence in the maze task with predictable ending (left) and unpredictable ending (right)



To ensure that participants saw all 32 story frames with predictability of the final noun phrase being counterbalanced, they were randomly assigned to one of two stimulus lists. In list A, the first half of the items had a predictable ending and the second half had an unpredictable ending. In list B, this was reversed. The order in which items appeared was fully randomized per participant. The list that participants were assigned to did not influence reading times in either predictability condition, but participants who were assigned to list A (48.7%) finished significantly more whole trials (M = 26.11, SD = 4.04) than participants who were assigned to list B (M = 23.80, SD = 4.50). Article gender was purposefully not balanced across items. Of the 32 story frames, 22 (68.8%) contained target nouns of common gender (i.e., "de") and 10 (32.2%) contained target nouns of neuter gender (i.e., "het"); based on the Corpus of Gesproken Nederlands (Corpus of Spoken Dutch, Oostdijk, 2000), this disparity corresponds with the relatively high frequency of words of common gender (± 69.1%) compared to words of neuter gender (± 30.9%) in the Dutch language (Deutsch & Wijnen, 1985; Fleur et al., 2020).

Procedure

The whole experiment was programmed to be self-deliverable through a web browser using the online survey tool Qualtrics and an extension of the JavaScript-based experiment platform Ibex, called PennController for Internet Based Experiments (PCIbex, Zehr & Schwarz, 2018)[4]. The code needed to run the experiment in PCIbex Farm has been made freely available at https://osf.io/enwdy. Participants were recruited online via convenience and snowball sampling techniques. They were told that the experiment, which would take 25 minutes to complete, was to be executed individually, on a computer or laptop, and in a quiet room. After having been introduced to the study topic, participants digitally signed the informed consent form. In a Qualtrics survey, they provided information about their age, gender, educational attainment, language usage, and potential language impairments. All questions included the option “I prefer not to say”. Next, participants were redirected to PCIbex, where they were administered the cognitive battery. Tasks appeared in the same order as presented in the method. After completing the cognitive battery, which required ten minutes, participants performed the maze task, which required another fifteen minutes. Finally, participants were thanked and given the opportunity to ask questions or report technical difficulties.

Data analysis

In conformity with the Peer Reviewers' Openness Initiative (https://opennessinitiative.org, Morey et al., 2016), this study was preregistered on the Open Science Framework (Foster & Deardorff, 2017). All materials and instructions related to this study were accessible during the review process on the OSF project “Moderating predictive language processing: Age and reading skill” (https://osf.io/enwdy), where they remain accessible. To test the first three hypotheses, a 2x2x2 mixed ANOVA was performed, with age (younger versus middle-aged) and reading skill (low versus high) as between-subjects factors, predictability (low versus high) as within-subjects factor, and the average article reading time per participant and per predictability condition (in milliseconds) as the dependent measure. The first hypothesis was to be accepted if the effect of predictability on reading time was stronger among younger adults than it was among middle-aged adults. The second hypothesis was to be accepted if this effect was stronger among highly skilled readers than it was among less skilled readers. The third hypothesis was to be accepted if the interaction between predictability and age group, or more specifically, the negative effect of higher age on predictive processing (H1), was smaller among highly skilled readers than it was among less skilled readers[5]. To test the fourth hypothesis, a multiple linear regression analysis was performed, with age and reading

[4] Under the instructions of the Research Ethics and Data Management Committee of the Tilburg School of Humanities and Digital Sciences, the collection of demographic information was moved to a Qualtrics survey.

[5] The formulation of this hypothesis differs from the one in the preregistration, which, by mistake, stated that the interaction effect should be larger among highly skilled readers and smaller among less skilled readers. Conceptually, the intention was for the hypothesis to be as it is stated above.

skill (both treated as continuous variables), processing speed, and memory capacity as predictors, and delay (i.e., the difference between the by-participant average reading time on unpredictable articles and the by-participant average reading time on predictable articles)[6] as outcome measure. This hypothesis was to be accepted if age lost its significance as a (negative) predictor as soon as processing speed and memory capacity were added to the model (both negative predictors), whereas reading skill would remain a (positive) significant predictor.

Participants whose scores on the cognitive battery fell outside the cutoff criterion of three standard deviations below the grand mean were removed from the dataset, as they were believed to have misread the instructions or misunderstood the task at hand. When this cutoff criterion resulted in a negative value, which was the case for the NWR and the BRDS, participants who did not finish any trials correctly were still removed from the dataset. In total, this led to the exclusion of three (2.9%) participants. Likewise, participants who were unable to generate at least six article reading times per predictability condition in the maze task were excluded from the analyses. Nineteen (18.6%) participants were unable to generate any reading times on these articles and four (3.9%%) were unable to generate at least twelve. The average number of completely finished trials was 24.92 (SD = 4.41, range 10-32) and did not depend on participants' group, $F(3, 75) = 2.28$, $p = .087$. No trials were excluded based on reading times being too long (i.e., over ten seconds). In sum, the data of 76 adults were analyzed.

Results

Interactions between predictability, age group, and reading skill group

To test the first three hypotheses, a 2x2x2 mixed ANOVA was performed, with age group (younger versus middle-aged) and reading skill group (low versus high reading skill) as between-subjects factors, article predictability (low versus high) as the within-subjects factor, and the average article reading time per participant and per predictability condition (in milliseconds) as dependent measure. A Kolmogorov-Smirnov test indicated that reading time on predictable articles was not normally distributed among middle-aged adults with low, $D(16) = .24$, $p = .012$, and high reading skill, $D(12) = .28$, $p = .011$. However, since the ANOVA tolerates violations to its normality assumption rather well (Field, 2013) and since Levene's test of equality of error variances yielded insignificant results in both predictability conditions, this

[6] When there are only two data points per participant, a simple difference score is mathematically equivalent to a repeated measures ANOVA (Anderson, as cited in Winer, 1971). Consequently, the two will produce the same significance statistics (see Smolkowski, 2019).

was not expected to endanger the validity of the results in a serious way. The within-subjects factor had only two levels, and so, the assumption of sphericity was automatically met.

There was a highly significant main effect of article predictability on reading time, $F(1, 72) = 305.50$, $p < .001$, $\eta^2 = .809$ (large effect). In line with the literature, reading time was shorter on predictable articles and longer on unpredictable articles (see Table 3). Age group also significantly influenced reading time, $F(1, 72) = 7.24$, $p < .001$, $\eta^2 = .091$ (large effect). Middle-aged participants needed more time to read both predictable (*M* = 886, *SD* = 169) and unpredictable articles (*M* = 1404, *SD* = 307), whereas younger participants needed less time to do so (*M* = 772, *SD* = 142 and *M* = 1267, *SD* = 240, respectively). In similar fashion, reading skill group significantly influenced article reading time, $F(1, 72) = 5.88$, $p = .018$, $\eta^2 = .076$ (medium-large effect). Less skilled readers needed more time to read both predictable (*M* = 872, *SD* = 158) and unpredictable articles (*M* = 1363, *SD* = 296), whereas highly skilled readers needed less time to do so (*M* = 756, *SD* = 144 and *M* = 1271, *SD* = 242, respectively).

Table 3

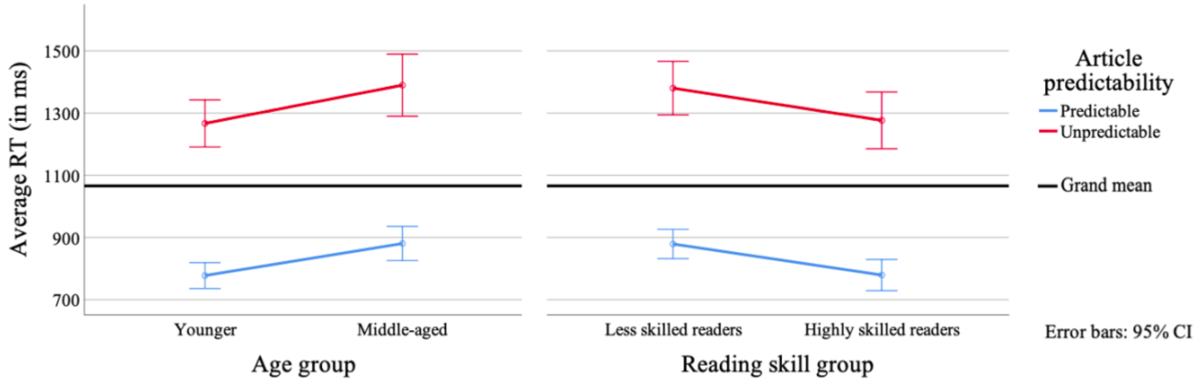
Average reading time on predictable and unpredictable articles (in milliseconds) per age and reading skill group

	Younger adults				Middle-aged adults				Total	
	Low RS		High RS		Low RS		High RS			
	M	SD	M	SD	M	SD	M	SD	M	SD
Pred.	838	148	717	113	920	163	841	172	814	161
Unpred.	1271	254	1263	232	1489	312	1291	271	1317	273

The interaction between article predictability and age group was not significant, $F(1, 72) = 0.12$, $p = .729$ (see the first graph in Figure 3). Age-related differences in reading time did not vary as a function of article predictability. As a consequence, the first hypothesis was rejected. There was also no significant interaction between article predictability and reading skill group, $F(1, 72) = 0.00$, $p = .948$ (see the second graph in Figure 3). Differences in reading time related to reading skill did not vary as a function of article predictability. For this reason, the second hypothesis was also rejected.

Figure 3

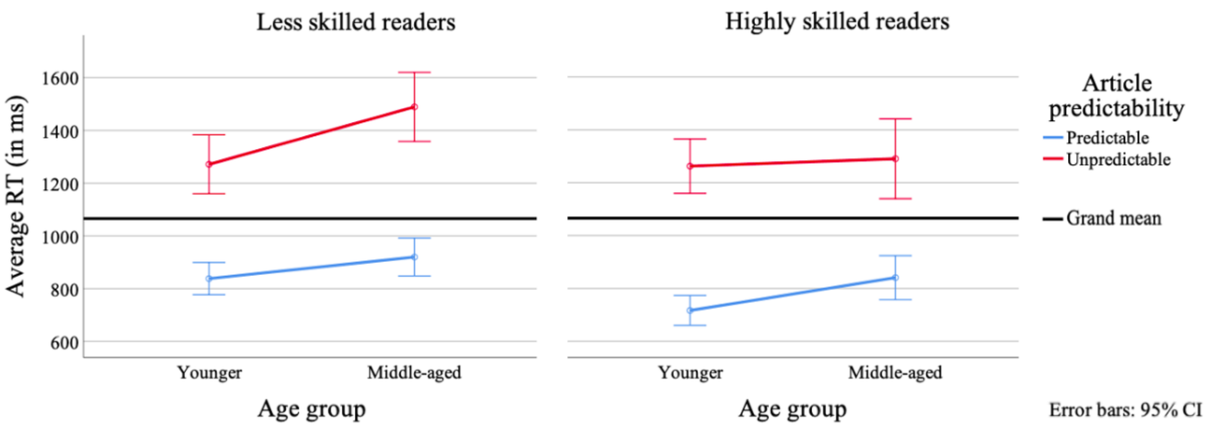
Interaction between article predictability and age group (left) and article predictability and reading skill group (right) on average article reading time (in milliseconds)



The ANOVA did reveal a significant three-way interaction between article predictability, age group, and reading skill group, $F(1, 72) = 4.12$, $p = .046$, $\eta^2 = .054$ (medium effect). A closer examination of this effect revealed that among less skilled readers, the difference between reading times on predictable and unpredictable articles tended to be larger for middle-aged participants, whereas among highly skilled readers, the difference tended to be larger for younger participants. This observation is diametrically opposed to the third hypothesis. Even though the interaction between predictability and age group was statistically significant among neither less skilled readers, $F(1, 36) = 2.73$, $p = .107$), nor highly skilled readers, $F(1, 36) = 1.49$, $p = .231$ (see Figure 4), the third hypothesis was rejected.

Figure 4

Three-way interaction effect between article predictability, age group, and reading skill group on average article reading time (in milliseconds)



The role of processing speed and memory capacity

To test the fourth hypothesis, a multiple linear regression analysis was performed. Age and reading skill were treated as continuous variables and entered the regression model as predictors, together with processing speed and memory capacity. The outcome measure, delay, was calculated as the difference between the by-participant average reading time on unpredictable articles and the by-participant average reading time on predictable articles (*M* = 503, *SD* = 241). There were no signs of multicollinearity, $r(74) < .56$ and VIF < 1.7 in all cases, average VIF = 1.37, tolerance = 0.73. None of the cases had a standardized residual above three and only four (5.3%) cases had a standardized residual above two. A Kolmogorov-Smirnov test indicated that standardized residuals were normally distributed, $D(76) = .07$, $p = .200$. They also appeared independent of each other, Durbin-Watson statistic = 2.01. Partial plots revealed no strong signs of heteroscedasticity or violations of the assumption of linearity. In sum, none of the assumptions were seriously violated.

Processing speed and memory capacity correlated positively with each other ($r(74) = .48$, $p < .001$). In line with the expectations, both factors also correlated positively with reading skill ($r(74) = .41$, $p < .001$ and $r(74) = .22$, $p = .027$, respectively). Moreover, processing speed and memory capacity declined as age increased ($r(74) = -.39$, $p < .001$ and $r(74) = -.55$, $p < .001$, respectively). However, the correlations between the delay variable and (1) age ($r(74) = .05$, $p = .345$), (2) reading skill (-.02, $p = .428$), (3) processing speed ($r(74) = -.02$, $p = .417$), and (4) memory capacity ($r(74) = -.07$, $p = .275$) all remained insignificant.

The first model, which contained only age and reading skill as predictors, was not an improvement over the null model, *F*-Change(2, 73) = 0.10, $p = .906$, $R^2a = -.025$. Neither age ($\beta = .05$, $SE = 1.80$, $p = .686$) nor reading skill ($\beta = -.02$, $SE = 98.55$, $p = .842$) predicted the delay variable. The second model, which contained the additional predictors processing speed ($M = 26.24$, $SE = 4.40$) and memory capacity ($M = 1.47$, $SE = 0.33$), was also not a significant improvement over the first model, *F*-Change(2, 71) = 0.09, $p = .911$, $R^2a = -.051$. Neither processing speed ($\beta = .02$, $SE = 8.21$, $p = .887$) nor memory capacity ($\beta = -.07$, $SE = 110.87$, $p = .669$) predicted delay. Correspondingly, the fourth hypothesis was rejected as well.

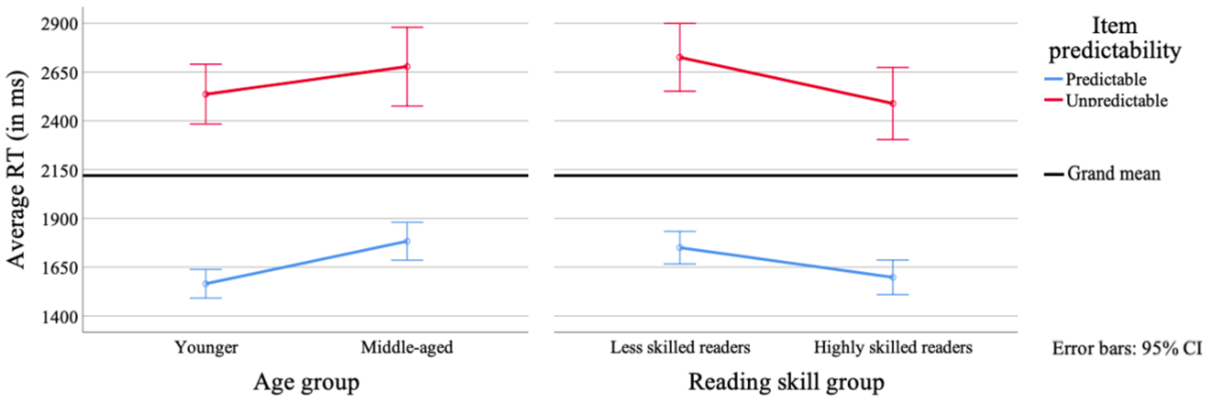
Exploratory analyses

To explore the maze task's sensitivity to spillover effects (see, e.g., Boyce et al., 2020; Forstel et al., 2009; Witzel et al., 2012), reading times on the regions that contained both the articles and the closing nouns, which were as (un)predictable as the articles that preceded them, were analyzed. A mixed ANOVA was performed, with the by-participant average reading time on the whole critical noun phrases (i.e., the sum of both reading times) per predictability condition (in milliseconds) as the dependent measure. Only completely finished trials were analyzed. Reading times were not normally distributed on predictable regions among middle-aged adults with low, $D(16) = .21$, $p = .049$, and high reading skill, $D(12) = .32$, $p = .001$. The assumption of homogeneity of variances, however, was met in all cases.

The ANOVA yielded similar results as before. Item predictability significantly impacted reading time, $F(1, 72) = 410.95$, $p < .001$, $\eta^2 = .851$ (large effect). Reading time was shorter on predictable regions ($M = 1644$, $SD = 286$) and longer on unpredictable regions ($M = 2595$, $SD = 539$). Age group again significantly influenced reading time, $F(1, 72) = 4.13$, $p = .046$, $\eta^2 = .054$ (medium effect). Middle-aged adults needed more time to read both predictable ($M = 1791$, $SD = 305$) and unpredictable regions ($M = 2703$, $SD = 575$), whereas younger adults needed less time to do so ($M = 1557$, $SD = 238$ and $M = 2532$, $SD = 513$, respectively). Reading skill group also remained to significantly affect reading time, $F(1, 72) = 4.83$, $p = .031$, $\eta^2 = .063$ (medium effect). Less skilled readers needed more time to read both predictable ($M = 1735$, $SD = 277$) and unpredictable regions ($M = 2705$, $SD = 561$), whereas highly skilled readers needed less time to do so ($M = 1552$, $SD = 268$ and $M = 2485$, $SD = 500$, respectively).

Figure 5

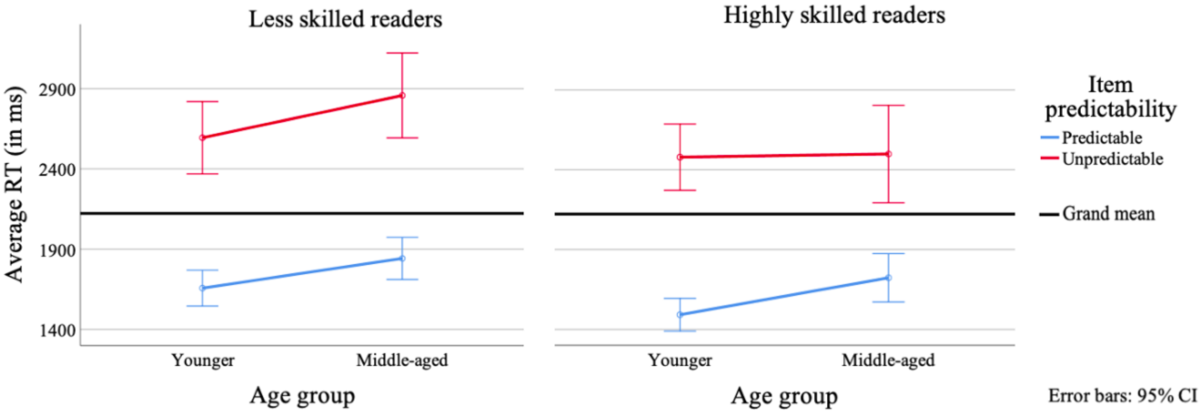
Interaction effect between item predictability and age group (left) and item predictability and reading skill group (right) on average item reading time (in milliseconds)



Again, item predictability did not interact with either age group, $F(1, 72) = 0.69$, $p = .410$, or reading skill group, $F(1, 72) = 0.85$, $p = .361$ (see Figure 5). Unlike before, however, there was no evidence of a significant three-way interaction, $F(1, 72) = 2.82$, $p = .097$ (see Figure 6).

Figure 6

Three-way interaction effect between item predictability, age group, and reading skill group on average item reading time (in milliseconds)

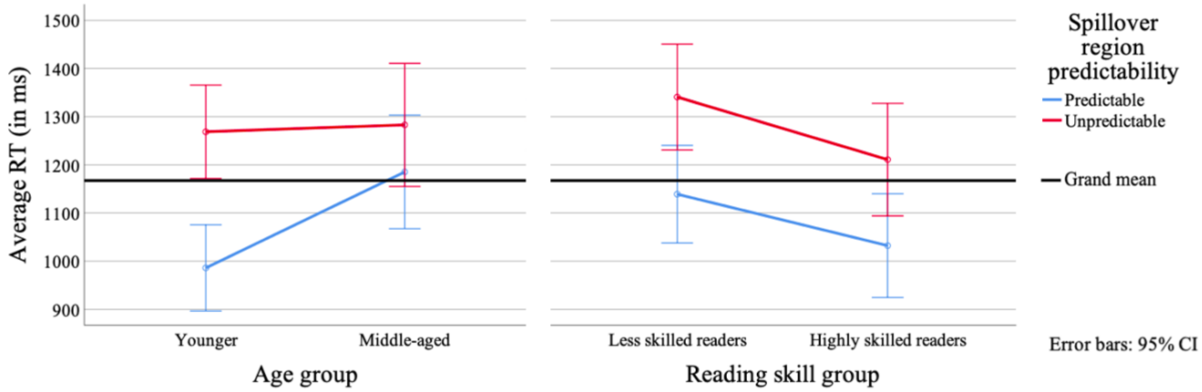


To investigate whether the reading times on closing nouns alone differed from the reading times on the whole region, a mixed ANOVA with the by-participant average reading time on the spillover region per predictability condition (in milliseconds) was performed. Reading time was not normally distributed on predictable spillover regions among younger adults with high reading skill, $D(26) = .18$, $p = .029$, and on unpredictable spillover regions among younger adults with low, $D(22) = .20$, $p = .024$, and high reading skill, $D(26) = .17$, $p = .050$. The assumption of homogeneity of variances was met in all cases.

Predictability again significantly affected reading time, $F(1, 72) = 16.04$, $p < .001$, $\eta^2 = .182$ (large effect). Reading time was shorter on predictable spillover regions ($M = 1059$, $SD = 325$) and longer on unpredictable spillover regions ($M = 1275$, $SD = 335$). Although middle-aged participants again needed more time to read both predictable ($M = 1193$, $SD = 350$) and unpredictable spillover regions ($M = 1294$, $SD = 290$), whereas younger participants needed less time to do so ($M = 981$, $SD = 285$ and $M = 1264$, $SD = 362$, respectively), this difference was no longer significant, $F(1, 72) = 3.05$, $p = .085$. In parallel, although less skilled readers again needed more time to read both predictable ($M = 1124$, $SD = 353$) and unpredictable spillover regions ($M = 1337$, $SD = 348$), whereas highly skilled readers needed less time to do so ($M = 995$, $SD = 284$ and $M = 1214$, $SD = 315$), this difference was no longer significant, $F(1, 72) = 3.74$, $p = .057$.

Figure 7

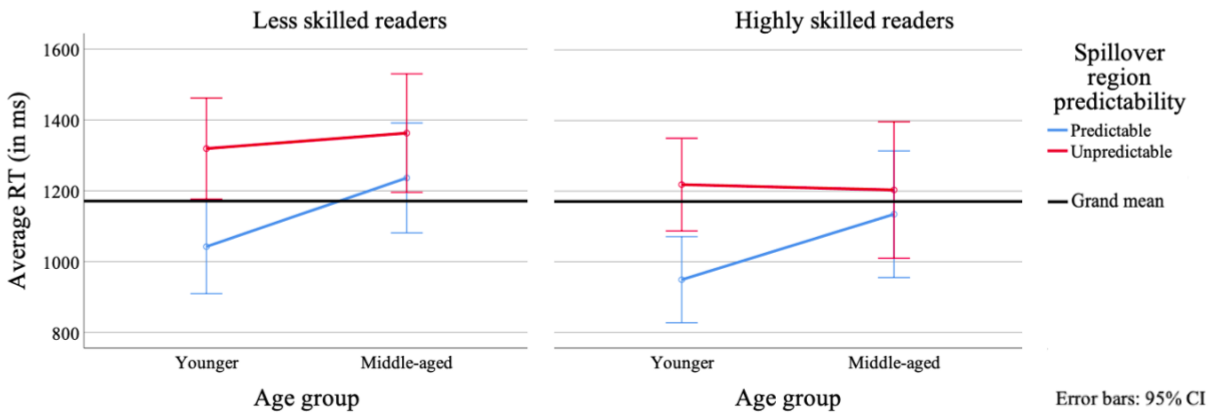
Interaction effect between spillover region predictability and age group (left) and spillover region predictability and reading skill group (right) on average spillover region reading time (in milliseconds)



Although there was a trend towards significance, predictability of the spillover region did not interact with age group, $F(1, 72) = 3.80$, $p = .055$. It also did not interact with reading skill group, $F(1, 72) = 0.06$, $p = .809$ (see Figure 7). The three-way interaction between spillover region predictability, age group, and reading skill group was also insignificant, $F(1, 72) = 0.13$, $p = .715$ (see Figure 8).

Figure 8

Three-way interaction effect between spillover region predictability, age group, and reading skill group on average spillover region reading time (in milliseconds)



To explore whether the results of the regression analysis would hold true when reading times on the spillover regions were included, a linear regression analysis was performed, with the corresponding delay variable, calculated as the difference between the by-participant average reading time on unpredictable regions and the by-participant average reading time on predictable regions ($M = 944.68$, $SD = 392.59$), as outcome measure. Models contained the same predictors as before and all assumptions were met. The correlation between delay and age remained insignificant (*r*(74) = -.08, *p* = .244). The correlation between delay and reading skill was significant, but unexpectedly negative (*r*(74) = -.26, *p* = .012). In line with the expectations, the correlation between delay and memory capacity was negative (*r*(74) = -.22, *p* = .031). The correlation between delay and processing speed was negative, but insignificant (*r*(74) = -.17, *p* = .069).

Although there was a trend towards significance in both regression analyses, the first model remained to be no significant improvement over the null-model, *F*-Change(2, 73) = 2.79, *p* = .068, R^2a = .046. Age did not predict delay (*β* = -.07, *SE* = 2.83, *p* = .548) and reading skill even negatively predicted delay (*β* = -.25, *SE* = 154.95, *p* = .027). The second model remained to be no significant improvement over the first model, *F*-Change(2, 71) = 2.63, *p* = .079, R^2a = .087. Age still did not predict delay (*β* = -.26, *SE* = 3.52, *p* = .066), reading skill lost its significance as a predictor (*β* = -.15, *SE* = 173.62, *p* = .239), and processing speed did not predict delay (*β* = -.07, *SE* = 12.47, *p* = .601). Memory capacity did negatively predict delay (*β* = -.29, *SE* = 168.40, *p* = .045). Similar effects appeared when the delay variable that corresponded to the reading time on closing nouns alone (*M* = 2010, *SD* = 402) entered a multiple linear regression analysis as outcome measure.

Discussion

Predictability effects on article reading time

The main goal of this study was to investigate how age and reading skill interactively modulate anticipatory language processing. More specifically, it was to examine whether high reading skill can compensate for potential deterioration in predictive language processing caused by higher age. When there is a mismatch between the input one encounters and one's expectations, a need arises for additional and more effortful integrative processing (Coulson et al., 1998; Vissers et al., 2006, 2008). This extra processing requires extra time (Boyce et al., 2020). In this study, articles were either very predictable or very unpredictable; the former signaled a prediction realization and the latter a prediction violation. In the maze task, predictive processing was indexed by relatively longer reading times on unpredictable articles and relatively shorter reading times on predictable articles. This pattern could indicate, either, (1) that participants had predicted not only potentially upcoming referents, "but also the word form of the noun plus the corresponding article" (Fleur et al., 2020, p. 17), or, (2) that they had predicted specific nouns, either with or without their gender, and then used article gender as soon as it became available to adjust their noun predictions.[7] The earliest signs of prediction could, thus, be observed in differential reading times on the Dutch gender-marked articles “de” or “het” that adults had to pick at the end of each self-built story in the grammaticality maze task.

Based on an extensive body of research (see, e.g., Brothers et al., 2017; Ehrlich & Rayner, 1981; Forster et al., 2009; Husband, 2021; Otten & Van Berkum, 2008, 2009; Rayner et al., 2011; Van Berkum et al., 2005; Wicha et al., 2004; Witzel & Forster, 2014), adults were expected to decelerate more when they encountered articles they had not predicted and less when they encountered articles that aligned with predictions. Across all analyses, strong and consistent evidence of this pre-nominal effect was found: reading times on predictable articles were significantly shorter than reading times on unpredictable articles. By providing a localized measure of reading time, the maze task thus allowed effects of predictability to be observed prior to the noun (Husband, 2021), thereby indicating processing time differences on exactly those words that were hypothesized to yield disparities (Witzel & Forster, 2012). Stories that ended with a low-cloze critical noun phrase indeed required more processing than stories that ended with a high-cloze critical noun phrase and the resulting difference in reading time was already observable when participants encountered the earliest cues that signaled either a prediction realization or a prediction violation. In other words, the maze task appeared sensitive to early predictive cues and may indeed be a worthy alternative method to investigate linguistic prediction when higher-fidelity resources are unavailable.

Reading skill impacting the negative effects of higher age

In line with the literature (e.g., Federmeier & Kutas, 2005; Huang et al., 2012; Kliegl et al., 2004; Rayner & Clifton, 2009; Steen-Baker et al., 2017), adults’ age and reading skill level affected how much time they needed to read articles. Even though the two generations were very similar in terms of reading skill, middle-aged adults needed substantially more time than younger adults to read articles. They also showed disadvantages in processing speed and memory capacity (cf. Federmeier et al., 2002, 2010; Huettig & Janse, 2016; Ronnlund et al., 2005). The poorer cognitive performance of middle-aged adults should, however, not necessarily be interpreted as a manifestation of cognitive decline, as it may reflect a simple consequence of lifelong learning on information processing (see Ramscar et al., 2014). More importantly, even though middle-aged adults were slower readers, their performance on the

[7] The former explanation is generally known as the article prediction mismatch hypothesis (DeLong et al., 2005; Kutas et al., 2011; Wicha et al., 2003, 2004) and the latter as the noun prediction revision hypothesis (Otten & Van Berkum, 2009; Van Berkum et al., 2005). The two are not mutually exclusive.

maze task did not appear to be compromised; the two generations delayed to a similar degree when they encountered unpredictable articles. This finding was consistent across analyses.

Analogously, less skilled readers needed significantly more time than highly skilled readers to read articles. Unexpectedly, however, the two reading skill groups displayed a delay of similar magnitude when they encountered unpredictable articles. The poor language skills of less skilled readers, thus, did not seem to negatively affect their performance on the maze task. Again, this finding was corroborated by the regression analysis. At first glance, neither age group nor reading skill group seemed to interact with article predictability.

The absence of moderation across different analyses suggests that all adults were able to demonstrate they had predicted upcoming linguistic content to a very similar degree. This would ultimately support the view that prediction indeed is a quintessential component of natural language processing (cf. Clark, 2013). Consequently, it is tempting to conclude that there may exist a potentially equal-level playing field for all proficient language users. The significant three-way interaction between age group, reading skill group, and article predictability, however, revealed this might not be the whole truth. In line with the first hypothesis, middle-aged adults showed slightly decreased prediction effects compared to younger adults (cf. Federmeier et al., 2010, 2012; Wlotko, 2012), but this pattern only appeared among highly skilled readers; among less skilled readers, predictive processing actually increased with age. This is diametrically opposed to findings from previous studies (e.g., Federmeier et al., 2010; Steen-Baker et al., 2017; Wlotko & Federmeier, 2012) and the third hypothesis, according to which high reading skill should be able to compensate for age-related degeneration in predictive language processing.

The finding that higher age only negatively affected the predictive processing of adults with high reading skill is striking and conflicts with a large body of evidence that emphasizes how positively connected prediction and literacy are (e.g., Huettig & Brouwer, 2015; Mani & Huettig, 2014, 2015; Mishra et al., 2012; Rommers et al., 2015). Most of this evidence comes from eye-tracking research, where predictive ability is mapped to anticipatory eye movements. High-literacy subjects typically outperform their less literate peers because they fixate anticipated objects sooner (e.g., Huettig & Brouwer, 2015; Mani & Huettig, 2012; Mishra et al., 2012; Rommers et al., 2015). In the maze task, increased performance takes the form of a relative speed *dis*advantage when the only grammatical sentence continuation is unpredictable (Forster et al., 2009; Husband, 2021; Witzel & Forster, 2014). This is because highly skilled readers are thought to anticipate upcoming linguistic content to a stronger degree and, therefore, to engage in more effortful integrative processing when initial predictions have to suppressed or revised (Federmeier et al., 2010). Even though these violations should give rise to a relatively larger delay in reading time, the results suggest quite the opposite: of the two reading skill groups, only the adults with high reading skill showed the expected age-related decline in predictive processing. Over the span of a lifetime, the expectancy violation may have become less prominent to these adults. This corresponds with the literature (e.g., Federmeier et al., 2010, 2012; Wlotko, 2012) and the first hypothesis. Strangely, however, this pattern was reversed, rather than stronger, among adults with low reading skill; when neither of the word options immediately matched predictions (i.e., when the only viable word option was an unpredictable article), they even tended to decelerate more as age increased. In other words, age actually appeared to have a slightly positive effect on predictive processing among less skilled readers (cf. Huettig & Janse, 2012). This suggests that, perhaps, having lifelong language experience might not be as detrimental to predictive processing as was traditionally thought (see Ramscar et al., 2014).

Even though the remarkable three-way interaction seems to suggest that, as age increases, only poor readers show more of a predictive effect, this need not be the case. Indeed, it would be strange to conclude that highly skilled readers notice the mismatch between the input and their predictions to a lesser extent than their less skilled peers. Rather, this observation may be explained by the fact that, in a grammaticality maze task, participants must complete a number of steps before they can choose which of two words correctly continues a sentence (Boyce et al., 2020). What follows is that the reading time between key presses may not necessarily be limited to variation in one underlying process. It is most definitely sensitive to variation in linguistic prediction; the pre-activation of certain word features directly links to reading time, because it facilitates processes such as word recognition (Witzel & Forster, 2012). However, reading time may also be sensitive to variation in some of the subsequent processes that a grammaticality maze task requires, such as lexical access or the integration of chosen words into the developing sentence representation.

It is not unlikely that adults' rate of incremental integration differed as a function of their reading skill. The Woodcock-Johnson reading fluency task (Schrank et al., 2014), which comprised 50% of the overall reading skill measure, even required semantic integration. If, indeed, highly skilled readers were faster integrators, they may have been better able to catch up on lost time during this process, thereby counteracting or even neutralizing a delay in reading time. The processing cost caused by an expectancy violation may then have been equally strong or, as the literature suggests, even stronger among highly skilled readers. Ultimately, differences in integrative speed could explain why skilled readers of both age groups yielded an almost identical average reading time on unpredictable articles (the difference in means was less than 30 milliseconds) and why reading skill even correlated negatively with the delay variable that corresponded to the whole noun phrase (p = .012). Less skilled readers may have struggled more with integrating words into sentence context, which resulted in a strong (sustained) processing cost and a relatively large delay. Among highly skilled readers, this processing cost may have been equally strong initially, but, due to large advantages in integration, still resulted in a moderate delay.

The results suggest that the extent to which highly skilled readers slowed down was more limited than the extent to which less skilled readers slowed down. Still, the observation that middle-aged adults with poor reading skill delayed the most of all groups (approximately 20 milliseconds more than younger adults with high reading skill) is striking and in need of future examination. If (dis)advantages in integrative speed had a large impact on the degree to which participants delayed, the way in which predictive processing was measured in this study might have concealed the preservation of predictive ability among middle-aged adults with high reading skill. It should be noted that this hypothesis is speculative; to investigate how individual differences in anticipatory language processing and the rate of lexical integration contribute to variations in reading time, future studies could employ alternative experimental designs that allow the two accounts to be better distinguished.

For example, in an ERP study, Mantegna et al. (2019) kept the ease of integration constant across conditions by employing sentences with rhyming completion. This enabled the researchers to manipulate word predictability separately and to attribute N400 effects to linguistic prediction alone. Alternatively, the current results could be further analyzed by means of an analysis of lexical co-occurrence. Words that often appear together, such as *bake* and *oven*, become co-activated once language comprehenders encounter them. As direct co-activation facilitates semantic integration, it can influence processing time (Savic et al., 2020). By analyzing the extent to which the words in the experimental sentences co-occur, one can investigate how much of the variance in data can be explained by low-level intralexical connections (Witzel & Forster, 2012).

Predictability effects on spillover region reading time

The maze task was successful in showing that predictability influences reading time (cf. Forster et al., 2009; Husband, 2021, Witzel & Foster; 2014; Witzel et al., 2012). The main effect of predictability persisted in the spillover region, but this observation should not be interpreted as sensitivity to spillover effects. This is because closing nouns, too, were either predictable or unpredictable (congruent with the article). A main effect was, thus, highly expected. Indeed, the exact definition of a spillover region may depend on the method that is used and, therefore, vary from study to study. It is possible that a different operationalization would have revealed that predictability effects did not spill over to subsequent phrases, which, in the current study, were removed to keep the task manageable.[8]

Still, if the effect of predictability on reading time was equally large on closing nouns as it was on articles (or if it was even larger), this could hint at a potential delay in prediction effects. Effect sizes were large in both cases, but predictability had a much larger effect on article reading time (η^2 = .809) than it had on spillover region reading time (η^2 = .182). This may imply that, although the violation of a prediction continued in the spillover region, readers were able to dismiss their original prediction to some degree and thereby regain time after the initial error caused by the article had been processed. This aligns with the highly incremental sentence processing that the maze task requires (Boyce et al., 2020; Forster et al., 2009; Witzel et al., 2012).

Although age group and reading skill group no longer significantly influenced reading time in the spillover region, reading times on closing nouns alone were characterized by an almost significant interaction between age group and predictability (p = .055). Younger adults delayed slightly more than middle-aged adults when choosing unpredictable closing nouns, which aligns with the first hypothesis. This finding was, however, not corroborated by the regression analysis, where age and delay did not appear to be correlated at all. Moreover, no convincing evidence was found of age group modulating the preactivation of linguistic input to an extent that would allow differences to be observed *before* participants decided upon the correct closing noun. This is somewhat problematic because differential reading times on the closing nouns are receptive to multiple interpretations; whereas a reading time difference on the articles must predominantly relate to “pure” lexical prediction (that is, as long as all the words and word options that precede the article are kept constant across conditions), the same observation on closing nouns may also be explained, in part, by lexical integration (Mantegna et al., 2019) and spreading activation accounts (Bassi, 2000). Indeed, differential reading times on nouns are more ambiguous because they can index a multitude of processes.

To reiterate, the prediction account and the integration account are not mutually exclusive. However, if one wants to establish the extent to which effects can be ascribed to

[8] In Fleur et al.’s (2020) ERP-study, from which the material originated, nouns were investigated separately. Due to the nature of ERP-research, the original stimuli contained neutral spillover regions (e.g., “in the city”) that followed the critical noun phrase. These regions had no use in the maze task and including them would increase the workload for participants with about one-hundred extra word choices. Consequently, stimuli were truncated. Closing nouns became the new spillover regions and were excluded from the primary analyses.

predictive processing alone, differentiating between them is useful. Again, to better investigate how much of the variance in spillover region reading time can be explained by prediction, future studies could perform an analysis of lexical co-occurrence (cf. Witzel & Forster, 2012), adopt alternative experimental designs (cf. DeLong et al., 2005; Mantegna et al., 2019; Van Berkum et al., 2005; Wicha et al. 2004), or compare the results of multiple experiments that employ different methods (cf. Brothers et al., 2017; Rommers et al., 2013). Following the latter strategy, future studies may also provide insight into the extent to which the pre-nominal effect can be reproduced using different methods (cf. Fleur et al., 2020) and, more importantly, whether it differs as a function of age and reading skill.

In addition, to investigate whether a delay in reading time indicates the same underlying psychological reality as more traditional indicators of linguistic prediction, we hope to see more studies examining the extent to which anticipatory eye movements or ERPs can be mapped to the specific pattern in reading behavior that indexed predictive processing in this study. Methods that combine eye-tracking with EEG, in particular, will enable the extraction of so-called fixation-related potentials during natural reading (cf. Loberg et al., 2019). They may provide a better picture of how not only age but also reading skill relates to some of the neural responses that have been linked to anticipatory language processes, such as the P600.

Limitations

As was mentioned in the method, the experiment was programmed to be self-deliverable through a web browser, partly because the pandemic complicated on-site execution of the experiment and partly because it allowed a greater number of participants to be recruited in a short window of time. Unfortunately, this approach resulted in many participants experiencing difficulties whilst executing the experiment on PCIbex (significantly more responses were received via Qualtrics than via PCIbex). A substantial number of participants were unable to finish sufficient or any maze task trials, which resulted in a considerable amount of data exclusion; the participant goal was no longer met in all groups. This limits the generalizability of results. Some participants reported having (accidentally) skipped the last instruction or having struggled with understanding the maze task despite having read the instruction. We also suspect that many participants closed their web browser prematurely, thereby not pressing the final button that would send their responses. To avoid these issues in the future, the experiment should be pretested on people who are representative of the whole final sample. It would also be useful to have the instructions visible during the execution of the tasks, so that participants can consult them multiple times. A solution to the improper logging of responses would be to send the responses to the PCIbex server immediately after participants concluded the last trial (the platform does not allow intermediate sending of responses). The debriefing will then have to be detached from the actual experiment.

Another methodological limitation concerns the validity of the six short cognitive tasks that together formed the cognitive battery. Standardized versions of these tasks exist, but most of them are still limited to English. Therefore, for some of the tasks, such as the LDT, Dutch items were created. Although, in essence, the adapted versions of the tasks were the same as the original ones, the exact items that were used were different and yet to be validated. This may explain why processing speed, memory capacity, and reading skill were not found to significantly predict the delay variable. In addition, some tasks, such as the NWR, were altered to make the experiment self-deliverable through the web. It is possible that the poorer cognitive test results of middle-aged adults related, at least in part, to the digital interface of the experiment, with which the younger adults may have been more familiar (Heimann-Steinert et al., 2021).

Still, linear regression analyses revealed that it was possible to predict adults' article reading time or reading speed based on all components of the cognitive battery (effects occurred independent of predictability condition, $p < .01$ in all but one case). To some extent, reading speed indicates comprehension (Kotzer et al., 2021), because comprehension allows for efficient reading (Perfetti & Stafura, 2014). In fact, to assess the reading ability of neurotypical adults, effective reading speed is oftentimes used (see Patching & Jordan, 2005). In line with the expectations, correlation coefficients were negative; as processing speed, memory capacity, and reading skill increased, adults' reading speed decreased (and, perhaps, their comprehension increased). Moreover, processing speed and memory capacity correlated negatively with age. This validates the adapted versions of the tasks to a certain extent. Nevertheless, we encourage future studies to compare online versions of the tasks to more traditional ones as well as develop a framework for creating versions in other languages. Both research efforts would enable a larger and more diverse sample to be recruited.

Following this line of reasoning, it may be noted that, although reading scores varied substantially among the participants, the sample was generally highly educated (half of the adults had a university degree). The younger adults had also followed higher education than the middle-aged adults. Likewise, the average age of the current sample was 36 years, which some may consider young. Future studies may provide a fuller picture of how linguistic anticipation varies across adulthood, or specifically after midlife, by including adults aged 65 and above. To better reflect the diversity of the true population, a wider and more even spread of educational attainment and age should be a priority in future research attempts.

Finally, a general methodological limitation concerns the binary nature of the predictability manipulation. In the current experiment, all stories were highly constraining and target words were either very predictable or very unpredictable. Even though most of the evidence for anticipatory language processing comes from experiments that employ similar prediction-stimulating materials (Huettig & Mani, 2016), highly predictable content words make up only a minority (i.e., about 5%) of the language people encounter in natural settings (Luke & Christianson, 2016). What follows is that results obtained thus far may not extend to more casual settings that better represent real-world situations. Consequently, the degree to which prediction is a primary feature of natural language processing remains partially unknown. To increase ecological validity in the future, experiments should make use of the full range of predictability (cf. Wlotko & Federmeier, 2013). A similar remark can be made about the categorization of adults into groups, as age and reading skill are best described as continuous factors (Bell & Perfetti, 1994). To study the interactions of multiple continuous factors, more statistically powerful approaches, such as linear mixed models, will be required.

Investigating low-cloze probability contexts will also assist in answering the question of how important prediction is for language comprehension. The observation that prediction effects are contingent on the population, the setting, and the cognitive resources/demands available implies that "successful language processing can and does take place in the absence of prediction" (Huettig & Mani, 2016, p. 27). If a smaller delay in reading time can indeed be interpreted as compromised predictive ability, future studies are encouraged to more rigorously examine the extent to which people who fail to demonstrate prediction-related benefits during sentence processing (e.g., children, adults with dyslexia, L2 learners) suffer from poor skills or other group characteristics. More research efforts should be spent on studying the factors that explain why some studies do not find the robust prediction effects that most studies do.

In 2011, Kutas et al. suggested that the scientific community move beyond the discussion of *whether* people predict upcoming language and instead examine *how* linguistic prediction operates and *what* exactly the contents of predictions may be. Indeed, the notion that language users often do predict upcoming language became widely accepted and these questions became a research topic of great interest in the decade that followed (Huettig, 2015). However, in attempting to come closer to answering the fundamental questions of just how much prediction is as necessary a feature of natural conversation or whether it is required for language comprehension, it is now time to extend the debate and focus more research efforts on finding out *who* predicts linguistic content and *when*. In this study, predictability, age group, and reading skill group were mostly found to have main effects on article reading time. Even though poor readers showed more of a predictive effect, we suspect this pattern appeared because of disadvantages in integrative speed. In sum, no convincing evidence was found for serious deterioration in predictive language processing related to compromised reading skill or higher age, at least into adults' midlife.

References

Aaronson, D., & Ferres, S. (2018). The word-by-word reading paradigm: An experimental and theoretical approach 1. In D. E. Kieras, & M. A. Just (Eds.), *New methods in reading comprehension research* (pp. 31-68). London, England: Routledge.

Aaronson, D., & Scarborough, H. S. (1977). Performance theories for sentence coding: Some quantitative models. *Journal of Verbal Learning and Verbal Behavior*, *16*(3), 277-303.

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*(3), 247-264.

Altmann, G. T., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, *57*(4), 502-518.

Ansari, A. H. (2018). *Development of neonatal EEG monitor for automated brain analysis* (Doctoral dissertation, KU Leuven, Belgium). Retrieved from: https://lirias.kuleuven.be/1658371?limo=0

Baddeley, A. D. (2006). Working memory: An overview. In S. Pickering (Ed.), *Working memory and education* (pp. 1–31). New York, NY: Academic Press.

Bassi, A. (2000). A dynamic associative semantic model for natural language processing based on a spreading activation network. In *Proceedings 20th International Conference of the Chilean Computer Science Society*, 99-108. Santiago, Chile.

Bell, L. C., & Perfetti, C. A. (1994). Reading skill: Some adult comparisons. *Journal of Educational Psychology, 86*(2), 244.

Bendixen, A., SanMiguel, I., & Schröger, E. (2012). Early electrophysiological indicators for predictive processing in audition: A review. *International Journal of Psychophysiology, 83*(2), 120-131.

Bock, K. (1990). Structure in language: Creating form in talk. *American Psychologist*, *45*(11), 1221–1236.

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology, 112*(4), 417-436.

Boyce, V., Futrell, R., & Levy, R. P. (2020). Maze made easy: Better and easier measurement of incremental processing difficulty. *Journal of Memory and Language*, *111*, 104082.

Bulkes, N. Z., Christianson, K., & Tanner, D. (2020). Semantic constraint, reading control, and the granularity of form-based expectations during semantic processing: Evidence from ERPs. *Neuropsychologia*, *137*, 107294.

Brothers, T., Swaab, T. Y., & Traxler, M. J. (2017). Goals and strategies influence lexical prediction during sentence comprehension. *Journal of Memory and Language*, *93*, 203-216.

Brus, B. T., & Voeten, M. J. M. (1973). *Eén-minuut-test: vorm A en B: verantwoording en handleiding: schoolvorderingentest voor de technische leesvaardigheid, bestemd voor groep 4 tot en met 8 van het basisonderwijs.* Nijmegen, Netherlands: Berkhout.

Carrow–Woolfolk, E. (1999). *Comprehensive Assessment of Spoken Language.* Circle Pines, MN: American Guidance Service.

Cheimariou, S., Farmer, T. A., & Gordon, J. K. (2021). The effects of age and verbal ability on word predictability in reading. *Psychology and Aging*, *36*(4), 531-542.

Choi, S. K., Rose, I. D., & Friedman, D. B. (2018). How is literacy being defined and measured in dementia research? A scoping review. *Gerontology and Geriatric Medicine*, *4*, 1-11.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181-204.

Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes*, 13(1), 21-

58.

Curcic, M., Andringa, S., & Kuiken, F. (2019). The role of awareness and cognitive aptitudes in L2 predictive language processing. *Language Learning*, *69*, 42-71.

De Jong, P. F., & Van der Leij, A. (1999). Specific contributions of phonological abilities to early reading acquisition: Results from a Dutch latent variable longitudinal study. *Journal of Educational Psychology*, *91*(3), 450.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Verbal fluency subtest of the Delis–Kaplan executive function system.* San Antonio, TX: Psychological Corp.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*(8), 1117-1121.

Deutsch, W., & Wijnen, F. (1985). The article's noun and the noun's article: Explorations into the representation and access of linguistic gender in Dutch, *Linguistics*, *23*(5), 793-810

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

Earles, J. L., & Salthouse, T. A. (1995). Interrelations of age, health, and speed. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *50*(1), 33-41.

Eckert, M. A. (2011). Slowing down: Age-related neurobiological predictors of processing speed. *Frontiers in Neuroscience*, *25*(5), 1-13.

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior*, *20*(6), 641-655.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, *44*(4), 491-505.

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.

Federmeier, K. D., & Kutas, M. (2005). Aging in context: age-related changes in context use during language comprehension. *Psychophysiology*, *42*(2), 133-141.

Federmeier, K. D., Kutas, M., & Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain and Language*, *115*(3), 149-161.

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.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics.* London, England: Sage.

Fleur, D. S., Flecken, M., Rommers, J., & Nieuwland, M. S. (2020). Definitely saw it coming? The dual nature of the pre-nominal prediction effect. *Cognition*, 204, 104335.

Forster, K. I., Guerrera, C., & Elliot, L. (2009). The maze task: Measuring forced incremental sentence processing time. *Behavior Research Methods*, *41*(1), 163-171.

Foster, E. D., & Deardorff, A. (2017). Open science framework (OSF). *Journal of the Medical Library Association: JMLA*, *105*(2), 203-206.

Frisson, S., Rayner, K., & Pickering, M. J. (2005). Effects of contextual predictability and transitional probability on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(5), 862-877.

Gathercole, S. E., & Baddeley, A. D. (1996). *The children's test of nonword repetition*. London, England: Psychological Corporation.

Heimann-Steinert, A., Latendorf, A., Prange, A., Sonntag, D., & Müller-Werdan, U. (2021).

Digital pen technology for conducting cognitive assessments: A cross-over study with older adults. *Psychological Research*, *85*(8), 3075-3083.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, *466*(7302), 29.

Hintz, F., Meyer, A.S., & Huettig, F. (2015). Doing a production task encourages prediction: Evidence from interleaved object naming and sentence reading. In E. Kaiser, T. Mintz, R. Pancheva, & J. Zevin (Eds.), *Proceedings of the 28th CUNY conference on Human Sentence Processing*, Los Angeles, CA.

Horton, W. S. (2005). Conversational common ground and memory processes in language production. *Discourse Processes*, *40*(1), 1-35.

Huang, H. W., Meyer, A. M., & Federmeier, K. D. (2012). A “concrete view” of aging: Event related potentials reveal age-related changes in basic integrative processes in language. *Neuropsychologia*, *50*(1), 26-35.

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, *1626*, 118-135.

Huettig, F., & Brouwer, S. (2015). Delayed anticipatory spoken language processing in adults with dyslexia: Evidence from eye-tracking. *Dyslexia*, *21*(2), 97-122.

Huettig, F., & Janse, E. 2012. Anticipatory eye movements are modulated by working memory capacity: Evidence from older adults. In *Proceedings of the 18th Architectures and Mechanisms for Language Processing Conference*, Riva del Garda, Italy.

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.

Huettig, F., & Mani, N. (2016). Is prediction necessary to understand language? Probably not.

Language, Cognition and Neuroscience, *31*(1), 19-31.

Huettig, F., & Mishra, R. K. (2014). How literacy acquisition affects the illiterate mind–a critical examination of theories and evidence. *Language and Linguistics Compass*, *8*(10), 401-427.

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, *137*(2), 151-171.

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

Husband, E. M. (2021). Prediction in the maze: Probabilistic pre-activation and the English a/an contrast. Paper presented at *Architectures and Mechanisms for Language Processing*, Paris, France.

Hyönä, J., & Olson, R. K. (1995). Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(6), 1430.

Ito, A., Corley, M., & Pickering, M. J. (2018). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism: Language and Cognition*, *21*(2), 251-264.

James, W. (1890). *The principles of psychology*. New York, NY: Dover.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, *111*(2), 228-238.

Kamide, Y., Altmann, G. T., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*(1), 133-156.

Kemp, N., Mitchell, P., & Bryant, P. (2017). Simple morphological spelling rules are not always used: Individual differences in children and adults. *Applied Psycholinguistics*, *38*(5), 1071-1094.

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.

Kieras, D. E. (1978). Good and bad structure in simple paragraphs: Effects on apparent theme, reading time, and recal. *Journal of Verbal Learning and Verbal Behavior, 17*(1), 13-28.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology*, *16*(1-2), 262-284.

Koornneef, A., & Mulders, I. (2017). Can we 'read' the eye-movement patterns of readers? Unraveling the relationship between reading profiles and processing strategies. *Journal of Psycholinguistic Research*, *46*(1), 39-56.

Kotzer, M., Kirby, J. R., & Heggie, L. (2021). Morphological awareness predicts reading comprehension in adults. *Reading Psychology*, *42*(3), 302-322.

Kraal, A., Koornneef, A. W., Saab, N., & van den Broek, P. W. (2018). Processing of expository and narrative texts by low-and high-comprehending children. *Reading and Writing*, *31*(9), 2017-2040.

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*(1), 32-59.

Kutas, M., DeLong, K., & Smith, N. (2011). A look around at what lies ahead: Prediction and predictability in language processing. In M. Bar (Ed.), *Predictions in the brain: Using our past to generate a future* (pp. 190-207). Oxford, England: Oxford University Press.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), 463-470.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*, 621-647.

Lai, M. L., Tsai, M. J., Yang, F. Y., Hsu, C. Y., Liu, T. C., Lee, S. W. Y., Lee, M. H., Chiou, G. L., Liang, J. C., & Tsai, C. C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review*, *10*, 90-115.

Light, L. L. (1991). Memory and aging: Four hypotheses in search of data. *Annual Review of Psychology*, *42*(1), 333-376.

Lindenberger, U., Nagel, I. E., Chicherio, C., Li, S. C., Heekeren, H. R., & Bäckman, L. (2008). Age-related decline in brain resources modulates genetic effects on cognitive functioning. *Frontiers in Neuroscience*, *2*, 234-244.

Loberg, O., Hautala, J., Hämäläinen, J. A., & Leppänen, P. H. (2019). Influence of reading skill and word length on fixation-related brain activity in school-aged children during natural reading. *Vision Research*, *165*, 109-122.

Logan, J. M., Sanders, A. L., Snyder, A. Z., Morris, J. C., & Buckner, R. L. (2002). Under-recruitment and nonselective recruitment: Dissociable neural mechanisms associated with aging. *Neuron*, *33*(5), 827-840.

Luke, S. G., & Christianson, K. (2013). SPaM: A combined self-paced reading and masked-priming paradigm. *Behavior Research Methods*, *45*(1), 143-150.

Luke, S. G., & Christianson, K. (2016). Limits on lexical prediction during reading. *Cognitive Psychology*, *88*, 22-60.

MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, *9*(2), 157-201.

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—But only for skilled producers. Jou*rnal of Experimental Psychology: Human Perception* *and Performance*, *38*(4), 843.

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.

Mantegna, F., Hintz, F., Ostarek, M., Alday, P. M., & Huettig, F. (2019). Distinguishing integration and prediction accounts of ERP N400 modulations in language processing through experimental design. *Neuropsychologia*, *134*, 107199.

Martin, C. D., Branzi, F. M., & Bar, M. (2018). Prediction is production: The missing link between language production and comprehension. *Scientific Reports*, *8*(1), 1-9.

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eye movements are modulated by reading ability: Evidence from Indian low and high literates. *Journal of Eye Movement Research*, *5*(1), 1-10.

Mitchell, D. C. 1984. An evaluation of subject-paced reading tasks and other methods for investigating immediate processes in reading. In D. Kieras, & M. A. Just (Eds.), *New methods in reading comprehension research* (pp. 69-89). Hillsdale, NJ: Erlbaum.

Meijer, W. A., de Groot, R. H., Van Boxtel, M. P., Van Gerven, P. W., & Jolles, J. (2008). Are age differences in verbal learning related to interstimulus interval and education? *Experimental Aging Research*, *34*(4), 323-339.

Moreno, M. A., & Van Orden, G. C. (2001). Word recognition, cognitive psychology of. In N. J. Smelser & P. B. Baltes (Eds.), *International Encyclopedia of Social and Behavioral Sciences* (pp. 16556-16561), Oxford, England: Pergamon.

Morey, R. D., Chambers, C. D., Etchells, P. J., Harris, C. R., Hoekstra, R., Lakens, D., Lewandowsky, S., Coker Morey, C., Newman, D. P., Schönbrodt, F. D., Vanpaemel, W., Wagenmakers, E., & Zwaan, R. A. (2016). The peer reviewers' openness initiative:

Incentivizing open research practices through peer review. *Royal Society Open Science*, *3*(1), 150547.

Mulder, K., & Hulstijn, J. H. (2011). Linguistic skills of adult native speakers, as a function of age and level of education. *Applied Linguistics*, *32*(5), 475-494.

Ng, S., Payne, B. R., Steen, A. A., Stine-Morrow, E. A., & Federmeier, K. D. (2017). Use of contextual information and prediction by struggling adult readers: Evidence from reading times and event-related potentials. *Scientific Studies of Reading*, *21*(5), 359-375.

Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., Von Grebmer Zu Wolfsthurn, S., Bartolozzi, F., Kogan, V., Ito, A., Mézière, D., Barr, D. J., A Rousselet, G., Ferguson, H. J., Busch-Moreno, S., Fu, X., Tuomainen, J., Kulakove, E., Husband, E. M., Donaldson, D. I., Kohút, Z., Rueschemeyer, S., & Huettig, F. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *ELife*, 7, e33468.

Oostdijk, N. H. J. (2000). Het corpus gesproken Nederlands. *Nederlandse Taalkunde, 5*, 280-284.

Otten, M., & Van Berkum, J. J. (2008). Discourse-based word anticipation during language processing: Prediction or priming? *Discourse Processes*, *45*(6), 464-496.

Otten, M., & Van Berkum, J. J. (2009). Does working memory capacity affect the ability to predict upcoming words in discourse? *Brain Research*, *1291*, 92-101.

Paolacci, G., & Chandler, J. (2014). Inside the turk: Understanding mechanical turk as a participant pool. *Current Directions in Psychological Science*, *23*(3), 184-188.

Patching, G. R., & Jordan, T. R. (2005). Spatial frequency sensitivity differences between adults of good and poor reading ability. *Investigative Ophthalmology & Visual Science*, *46*(6), 2219-2224.

Perfetti, C. A., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*(1), 22–37.

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, *144*(10), 1002-1044.

Pluck, G. (2018). Lexical reading ability predicts academic achievement at university level. *Cognition, Brain, Behavior*, *22*(3), 175-196.

Poon, L. W., Martin, P., Clayton, G. M., Messner, S., Noble, C. A., & Johnson, M. A. (1992). The influences of cognitive resources on adaptation and old age. *The International Journal of Aging and Human Development*, *34*(1), 31-46.

Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., & Baayen, H. (2014). The myth of cognitive decline: Non-linear dynamics of lifelong learning. *Topics in Cognitive Science*, *6*(1), 5-42.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372-422.

Rayner, K., & Clifton Jr, C. (2009). Language processing in reading and speech perception is fast and incremental: Implications for event-related potential research. *Biological Psychology*, *80*(1), 4-9.

Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and Aging*, *21*(3), 448-465.

Rayner, K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(2), 514-528.

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, *3*(4), 504-509.

Roland, D., Yun, H., Koenig, J. P., & Mauner, G. (2012). Semantic similarity, predictability, and models of sentence processing. *Cognition*, *122*(3), 267-279.

Rommers, J., Meyer, A. S., & Huettig, F. (2015). Verbal and nonverbal predictors of language-mediated anticipatory eye movements. *Attention, Perception, & Psychophysics*, *77*(3), 720-730.

Rommers, J., Meyer, A. S., Praamstra, P., & Huettig, F. (2013). The contents of predictions in sentence comprehension: Activation of the shape of objects before they are referred to. *Neuropsychologia*, *51*(3), 437-447.

Ronnlund, M., Nyberg, L., Backman, L., & Nilsson, L. G. (2005). Stability, growth, and decline in adult life span development of declarative memory: Cross-sectional and longitudinal data from a population-based study. *Psychology & Aging*, *20*(1), 3–18.

Roser, M., & Ortiz-Ospina, E. (2016). *Literacy*. Retrieved from https://ourworldindata.org/literacy

Salthouse, T. A. (1991). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science*, *2*, 179-183

Salthouse, T. A. (1993). Speed mediation of adult age differences in cognition. *Developmental Psychology*, *29*(4), 722-738.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403-428.

Savic, O., Unger, L., & Sloutsky, V. (2020). Exposure to co-occurrence regularities in language drives semantic integration of new words. doi.org/10.31234/osf.io/p23z6

Schomers, M. R., Kirilina, E., Weigand, A., Bajbouj, M., & Pulvermüller, F. (2015). Causal influence of articulatory motor cortex on comprehending single spoken words: TMS evidence. *Cerebral Cortex*, *25*(10), 3894-3902.

Schrank, F. A., Mather, N., & McGrew, K. S. (2014). *Woodcock-Johnson IV tests of* *cognitive abilities*. Rolling Meadows, IL: Riverside.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379-423.

Silbert, L. J., Honey, C. J., Simony, E., Poeppel, D., & Hasson, U. (2014). Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. *Proceedings of the National Academy of Sciences*, *111*(43), 4687-4696.

Slosson, R. L., & Nicholson, C. L., (1990). *Slosson oral reading test*. East Aurora, NU: Slosson Educational.

Smolkowski, K. (2019) *Gain score analysis*. Retrieved from: http://homes.ori.org/keiths/Tips/Stats_GainScores.html

Staub, A. (2015). The effect of lexical predictability on eye movements in reading: Critical review and theoretical interpretation. *Language and Linguistics Compass*, *9*(8), 311-327.

Staub, A., & Clifton Jr, C. (2006). Syntactic prediction in language comprehension: Evidence from either... or. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(2), 425.

Steen-Baker, A. A., Ng, S., Payne, B. R., Anderson, C. J., Federmeier, K. D., & Stine-Morrow, E. A. (2017). The effects of context on processing words during sentence reading among adults varying in age and literacy skill. *Psychology and Aging*, *32*(5), 443-467.

Thorn, A. S., & Gathercole, S. E. (1999). Language-specific knowledge and short-term memory in bilingual and non-bilingual children. *The Quarterly Journal of Experimental Psychology: Section A*, *52*(2), 303-324.

Tombaugh, T. N., Kozak, J., & Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Archives of Clinical Neuropsychology*, *14*(2), 167-177.

Burnett, N., Packer, S., & Bella, N. (2016). *Education for all: Literacy for life*. Paris, France: United Nations Educational, Scientific and Cultural Organization. Retrieved from: http://www.unesco.org/new/en/education/themes/leading-the-international-agenda/efareport/reports/2006-literacy/

Van Berkum, J. J., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(3), 443-476.

Van den Bos, K. P., Spelberg, H. C. L., Scheepstra, A. J. M., & De Vries, J. R. (1994). *De KLEPEL test voor de leesvaardigheid van pseudo-woorden*. Nijmegen, Netherlands: Berkhout.

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, *83*(2), 176-190.

Verhaeghen, P., Steitz, D. W., Sliwinski, M. J., & Cerella, J. (2003). Aging and dual-task performance: A meta-analysis. *Psychology and Aging*, *18*(3), 443–460.

Vissers, C. T. W., Chwilla, D. J., & Kolk, H. H. (2006). Monitoring in language perception: The effect of misspellings of words in highly constrained sentences. *Brain Research*, *1106*(1), 150-163.

Vissers, C. T. W., Kolk, H. H., Van de Meerendonk, N., & Chwilla, D. J. (2008). Monitoring in language perception: Evidence from ERPs in a picture-sentence matching task. *Neuropsychologia*, *46*(4), 967-982.

Wechsler, D (2004). *WAIS-III Nederlandstalige bewerking Wechsler adult intelligence scale:*

Afname en scoringshandleiding. Amsterdam, Netherlands: Harcourt Test.

Wicha, N. Y., Bates, E. A., Moreno, E. M., & Kutas, M. (2003). Potato not pope: Human brain potentials to gender expectation and agreement in Spanish spoken sentences. *Neuroscience Letters*, *346*(3), 165-168.

Wicha, N. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*, *16*(7), 1272-1288.

Winer, B. J. (1971). *Statistical principles in experimental design* (2nd ed.). New York, NY: McGraw-Hill.

Witzel, J., & Forster, K. (2014). Lexical co-occurrence and ambiguity resolution. *Language, Cognition and Neuroscience*, *29*(2), 158-185.

Witzel, N., Witzel, J., & Forster, K. (2012). Comparisons of online reading paradigms: Eye tracking, moving-window, and maze. *Journal of Psycholinguistic Research*, *41*(2), 105-128.

Wlotko, E. W., & Federmeier, K. D. (2012). Age-related changes in the impact of contextual strength on multiple aspects of sentence comprehension. *Psychophysiology*, *49*(6), 770-785.

Wlotko, E. W., & Federmeier, K. D. (2013). Two sides of meaning: The scalp-recorded N400 reflects distinct contributions from the cerebral hemispheres. *Frontiers in Psychology*, *4*(181). doi:10.3389/fpsyg.2013.00181

Wlotko, E. W., Federmeier, K. D., & Kutas, M. (2012). To predict or not to predict: age-related differences in the use of sentential context. *Psychology and Aging*, *27*(4), 975.

Whitford, V., & Titone, D. (2016). Eye movements and the perceptual span during first-and second-language sentence reading in bilingual older adults. *Psychology and Aging*, *31*(1), 58.

Yeatman, J. D., Tang, K. A., Donnelly, P. M., Yablonski, M., Ramamurthy, M., Karipidis, I. I., Caffarra, S., Takada, M. E., Kanopka, K., Ben-Shachar, M., & Domingue, B. W. (2021). Rapid online assessment of reading ability. *Scientific Reports*, *11*, 3696.

Zehr, J., & Schwarz, F. (2018). *Penncontroller for internet based experiments (IBEX)*. doi.org/10.17605/OSF.IO/MD832

Zhan, L. (2018). Using eye movements recorded in the visual world paradigm to explore the online processing of spoken language. *Journal of Visualized Experiments: JoVE, 140*, 58086.

Appendix

All story frames with high- and low-cloze noun phrases (left) and distractors (right).

1	Ik	.
	lees	extra
	thuis	gezegd
	graag	helpen
	het	ik
	nieuws	schuld
	op	zijn
	papier.	hoelang.
	Daarom	.
	lees	laatst
	ik	dat
	dus	weg
	iedere	missen
	ochtend	zwemmen
	de	*is*
	krant.	*adres.*
	het	*is*
	dagblad.	*beverig.*
2	Roelof	.
	hielp	dader
	met	ben
	de	ik
	voorbereidingen	geheimzinnig
	voor	weet
	het	is
	kerstdiner.	lastigval.
	Hij	.
	begon	omhoog
	eerst	denkt
	met	was
	het	is
	dekken	warmte
	van	hij
	de	*ik*
	tafel.	*kapot.*
	het	*ik*
	krukje.	*graast.*
3	Theo	.
	leerde	twijfel
	zijn	heb
	zoontje	gehoopt
	thuis	vertel
	Mozart	opzoek
	spelen.	moment
	Gezellig	.
	zaten	rechts
	ze	op
	weer	komt
	samen	vertel
	achter	moest
	de	*is*
	piano.	*apart.*
	het	*is*
	keyboard.	*retoriek.*
4	Het	.
	is	wat
	zondagochtend.	geldproblemen.
	De	.
	gehele	beschouw
	gelovige	lammetje
	familie	volgens
	gaat	jullie
	zoals	bedankt
	altijd	maken
	naar	moet
	de	*op*
	kerk.	*belde.*
	het	*op*
	gebedshuis.	*abstracter.*
5	Vanuit	.
	de	is
	auto	eten
	hadden	meisje
	we	van
	telefonisch	sterveling
	alvast	menigte
	een	is
	kamer	vergat
	gereserveerd.	democratie.
	We	.
	kwamen	begrijpt
	die	heb
	avond	hoorde
	namelijk	reputatie
	pas	oh
	laat	doe
	aan	hem
	bij	zou
	het	*ik*
	hotel.	*slechts.*
	de	*ik*
	camping.	*opzoekt.*
6	Neil	.
	Armstrong	oostelijke

is
een
bekende
astronaut.
Hij
was
namelijk
de
eerste
mens
die
voet
zette
op
de
maan.
het
maanlandschap.
7 Femke
wil
de
kerstverlichting
aanzetten.
Ze
stopt
de
stekker
in
het
stopcontact.
de
stekkerdoos.
8 Mijn
moeder
was
aan
het
bakken.
Precies
op
tijd
haalde
ze
de
cake
uit
de
oven.
het
fornuis.

niet
we
starten
daaronder.
.
dit
gedeelte
is
geven
erin
heb
grap
bijbel
ze
is
vecht.
is
identificeerden.
.
nog
hij
welgemanierd
vriendelijke.
.
stilte
is
begeeft
we
ik
terechtkwam.
ik
ruziemaakte.
.
zullen
ben
moet
van
finale.
.
er
geef
excuus
in
is
exact
dat
is
bind.
is
nodigt.

9 Julius
had
de
soepgroenten
gesneden.
Ondertussen
kookte
het
water
al
in
de
pan.
het
potje.
10 We
zouden
vanaf
Frankfurt
naar
New York
vliegen.
Met
de
taxi
kwamen
we
die
ochtend
aan
bij
het
vliegveld.
de
luchthaven.
11 Tijdens
koningsdag
waren
de
koning
en
Maxima
in
Groningen.
Ze
werden
allereerst
officieel
ontvangen
door

.
m'n
is
consulteert
diensten.
.
hengel
te
boven
zou
er
ik
geel.
ik
teams.
.
politie
hield
toepassen
weet
met rust
gedacht.
.
is
heus
succes
in
heb
genoemd
hem
zou
van
tegenover.
van
weggegooid.
.
hooghoudt
andere
is
gebeld
is
omsluit
ze
migreerde.
.
perfect
borgtocht
kasteel
verdachte
eens

de
burgemeester.
het
bestuurshoofd.
12 De
knuffel
was
uit
de
wandelwagen
in
het
stof
gevallen.
Ik
reinigde
hem
thuis
op
zestig
graden
in
de
wasmachine.
het
wastoestel.
13 Mijn
opa
had
een
opvouwbaar
mes
uit
Zwitserland.
Van
dichtbij
keek
ik
nieuwsgierig
naar
het
zakmes.
de
stiletto.
14 Er
kwam
een
boodschap
vanuit
de

en
interesseert.
en
lichtsnelheden.
.
vreemds
dit
bent
van
slachtafval
er
ik
hang
hopelijk.
.
verderop
aan
nieuwe
er
herfst
haalde
hij
we
zuidwesten.
we
onvervaard.
.
belt
meer
kan
oliepompen
lui
ook
afgestudeerd.
.
waarover
ouwe
en
gerechtigheid
moet
van
klapte.
van
naderden.
.
bang
hij
vanmorgen
rennen
is

cockpit.
Ik
luisterde
naar
de
rustige
stem
van
de
piloot.
het
crewlid.
15 Sjaak
wil
afvallen
en
meer
bewegen.
In
plaats
van
de
lift
neemt
hij
nu
altijd
de
trap.
het
trappenhuis.
16 We
kwamen
aan
bij
een
kruispunt.
Ik
remde
vanwege
de
rode
kleur
van
het
stoplicht.
de
waarschuwings-
lamp.
17 Richard

leverde.
.
chocola
weet
we
verveel
viel
een
ik
Mexico.
ik
gebruld.
.
nog
handelde
je
doe
koningin.
.
ervan
wat
is
eraf
kwaad
in
ja
gedaan
er
enig.
er
overeenkomt.
.
vraagt
kan
zal
van
overnieuw.
.
snakt
gelooft
is
belt
flink
hij
ik
vertraagt.
ik
wegwerp-
cameraatje.
.

voelde
een
pukkel
op
zijn
wang.
Hij
keek
meteen
in
de
spiegel.
het
raam.
18 Mariska
houdt
niet
van
uitgaan.
Ze
leest
liever
een
goed
boek
thuis
op
de
bank.
het
bankstel.
19 De
politieagenten
hadden
de
verdachte
opgepakt.
Hij
moest
direct
in
de
auto
mee
naar
het
bureau.
de
cel.
20 De

degene
ben
venijn
er
maar
ergs.
.
ouwe
handen
hij
is
prettig.
is
liep.
.
eigen
is
wat
betrapt.
.
amper
redden
is
hoe
zeer
nieuwe
zijn
is
welk.
is
afleidde.
.
geïnteresseerd
wereld
is
bespreken
geheugen.
.
onder
gebouw
hij
ik
bang
man
mijn
niet
gelukt.
niet
der.
.

auto
voor
mij
ging
opeens
erg
langzaam.
Om
niet
te
botsen
trapte
ik
hard
op
de
rem.
het
pedaal.
21 Jitse
komt
thuis
en
gooit
haar
jas
op
de
bank.
Haar
moeder
hangt
de
jas
aan
de
kapstok.
het
haakje.
22 Milan
verwacht
vandaag
een
brief.
Door
het
raam
ziet
hij
eindelijk

want
als
zou
idee
draait
zie
bedoelde.
.
ik
we
perzik
vetzak
is
baas
er
te
fit.
te
leunde.
.
eens
denkt
is
tekst
bent
rood
me
ik
viel.
.
binnen
eikel
ik
vlug
hem
en
omleiden.
en
maaide.
.
zichzelf
probleem
hij
mijne.
.
ik
gooi
leuk
van
begrijpen

de
postbode.
het
bestelbusje.
23 Tijdens
het
sjieke
kerstdiner
morste
Evert
rode
wijn
over
de
gehele
tafel.
Daarom
zit
er
nog
steeds
een
vlek
op
het
tafelkleed.
de
tafelloper.
24 Men
verwachtte
een
toename
van
het
aantal
leerlingen.
Gelukkig
werken
er
nu
wel
genoeg
leerkrachten
op
de
school.
het
internaat.
25 Ik
had
er
verraadt.
er
verwonderen.
.
niet
gruwel
weggebracht
pollen
oogde
belt
lees
doen
is
inzien
kapot.
.
z'n
is
nee
geven
en
zulk
er
we
spendeerde.
we
vaatwassen.
.
middernacht
is
heetten
hij
ik
honden
beschikbaar.
.
rustig
is
ja
kan
vinden
vergrendelen
er
is
tussen.
is
kwijtraak.
.
meer
in
de
stad
verkeerd
geparkeerd.
Mijn
oma
bood
financiële
hulp
bij
het
betalen
van
de
boete.
het
dwangbevel.
26 Iris
deed
mee
aan
een
schaaktoernooi.
Ze
zette
de
stukken
klaar
op
het
schaakbord.
de
rand.
27 Na
de
storm
was
er
schade
aan
de
schoorsteen
van
het
huis.
Ook
waaiden
er
veel
hij
ik
ligt
president
dramatisch.
.
ach
zon
investeert
juni
zal
hij
vandaan
een
is
grijs.
is
vervangers.
.
eten
man
wil
van
paardenbloemen.
.
nicht
is
belang
bedoel
er
is
grenzeloos.
is
koos.
.
ik
leert
als
we
erheen
wil
is
vermoordden
een
is
doet.
.
vastzet
hij
kijk

	dakpannen van *het dak.* *de dakbedekking.*	boerengat een *is lol.* *is improviseert.*
28	Het broeden van de kip had weken geduurd. Gister kwam het kuiken uit *het ei.* *de kippenstal.*	. opbelt een ik erom meer hield opkomen. . onder is hernia kom *ik ha.* *ik signeerden.*
29	Peter houdt van goed brood. Hij kocht een brood om de hoek bij *de bakker.* *het broodhuis.*	. eigen hij weet sterke. . valse ben zonde ja is durf mij *ik straalt.* *ik mompelen.*
30	Anne verbrandt snel. Als ze buiten	. binnenste auto. . te vertel

	is, blijft ze het liefst in *de schaduw.* *het park.*	we, kleine in hij afgaat er *is hielden.* *is lieg.*
31	Ze zijn boven op de Eiffeltoren geklommen. Na een lange klim genoten ze van *het uitzicht.* *de uitkijkpost.*	. maar praat we is concurreren irritante. . was zodra fans congres in dat *op verander.* *op transparant.*
32	Alexander is aan het voetballen met vrienden. Hij rent en schopt hard tegen *de bal.* *het net.*	. ik wil van goochelaar die vermoord. . muis we toerist baby weten *is zul.* *is wij.*