Central Lancashire Online Knowledge (CLoK)

Title	Adult Age Differences in Parafoveal Preview Effects during Reading: Evidence from Chinese
Type	Article
URL	https://clok.uclan.ac.uk/38868/
DOI	https://doi.org/10.1037/pag0000639
Date	2021
Citation	He, Liyuan, Ma, Weidong, Shen, Fengdan, Wang, Yongsheng, Wu, Jie, Warrington, Kayleigh. L, Liversedge, Simon Paul and Paterson, Kevin. B (2021) Adult Age Differences in Parafoveal Preview Effects during Reading: Evidence from Chinese. Psychology and Aging, 36 (7). pp. 822-833. ISSN 0882-7974
Creators	He, Liyuan, Ma, Weidong, Shen, Fengdan, Wang, Yongsheng, Wu, Jie, Warrington, Kayleigh. L, Liversedge, Simon Paul and Paterson, Kevin. B

Adult Age Differences in Parafoveal Preview Effects during Reading:

Evidence from Chinese

Liyuan He[1,2,3]

Weidong Ma[2,4]

Fengdan Shen[2]

Yongsheng Wang[1,2,3]

Jie Wu[1,2,3]

Kayleigh L. Warrington[5]

Simon P. Liversedge[6]

Kevin B. Paterson[7]

1. Key Research Base of Humanities and Social Sciences of the Ministry of Education, Academy of Psychology and Behavior, Tianjin Normal University, Tianjin, China

2. Faculty of Psychology, Tianjin Normal University, Tianjin, China

3. Tianjin Social Science Laboratory of Students' Mental Development and Learning, Tianjin, China

4. Tianjin University of Traditional Chinese Medicine, Tianjin, China.

5. School of Psychology, Nottingham Trent University, Nottingham, UK.

6. School of Psychology, University of Central Lancashire, Preston, UK.

7. Department of Neuroscience, Psychology and Behaviour, University of Leicester,

Leicester, UK.

Abstract

We investigated parafoveal processing by 44 young (18-30 years) and 44 older (65+ years) Chinese readers using eye movement measures. Participants read sentences which included an invisible boundary after a two-character word (N) and before two one-character words (N+1, N+2). Before a reader's gaze crossed the boundary, N+1 and N+2 were shown normally or masked (i.e., as valid/invalid previews), after which they reverted to normal. Young adults obtained preview benefits (a processing advantage for valid over invalid previews) for both words. However, older adults obtained N+2 preview benefits only when N+1 was valid, suggesting their parafoveal processing is more limited.

Keywords: Cogntive aging, eye movements in reading, Chinese, parafoveal processing; N+2 preview effects

Despite their greater experience of reading, older adults (aged 65+ years) tend to read less fluently than young adults (aged 18-30 years). This is reflected in slower reading by older adults, due to their dwelling for longer on words (i.e., making more and longer eye fixations) and making more backwards eye movements (i.e., regressions) to re-read text (for reviews, see Gordon et al., 2016; Paterson et al., 2020). Various factors are hypothesized to contribute to this less fluent reading in older age, including that older adults process information from upcoming words (i.e., in parafoveal vision) less effectively compared to young adults (e.g., Rayner et al., 2009, 2010).

It is well-established that parafoveal processing makes an important contribution to skilled reading. It allows readers to use low-resolution information from outside central (i.e., foveal) vision to begin processing upcoming words and to help guide the targeting of the next eye movement (for reviews, see Cutter et al., 2015; Schotter et al., 2012). Evidence for this comes primarily from studies that use gaze-contingent text-change techniques, such as the moving window and boundary paradigms, to limit the availability of parafoveal information on each fixational pause (for a review, see Rayner, 2014). This is achieved in the moving window paradigm by making changes to text dependent on where the reader currently is looking. On each eye fixation, text is shown as normal within a region around the reader's gaze (the 'moving window") while text outside this region is masked. This restricts the availability of parafoveal information, and the experimenter can assess its effects on eye movement behavior. By comparison, the boundary paradigm is used to investigate effects of limiting parafoveal information on the processing of specific words in text. In this paradigm, an invisible boundary is placed immediately in front of a target word in a sentence, which is shown either as normal or masked so that it unavaible for parafoveal processing. Once the reader's gaze crosses the boundary, the word reverts to normal, allowing the experimenter to assess effects of denying parafoveal preview of this word on its subsequent processing.

Studies using these paradigms show that the eye movements of skilled, young adult readers are disrupted when rightward parafoveal information (i.e., about upcoming words) is denied when reading left-to-right scripts like English (e.g., McConkie & Rayner, 1975, 1976; Rayner, 1975). Other studies show that slower or less skilled readers are disrupted less when rightward parafoveal information is denied, implying that such readers process this information less effectively (Chace et al., 2005; Häikiö et al., 2009; Rayner et al., 2010; Underwood & Zola, 1986; Veldre & Andrews, 2015). As older adults show reduced peripheral visual processing in non-reading tasks (e.g., Ball et al., 1988; Sekuler et al., 2000; for a review, see Owsley, 2011), this has led some researchers to argue that older readers might also process parafoveal information less effectively (e.g., Rayner et al., 2009). Consistent with this view, several studies using the moving window paradigm suggest that older readers can obtain less parafoveal information compared to young adult readers (Rayner et al., 2009, 2014). Similarly, findings from studies using the boundary paradigm suggest that the processing of specific target words in sentences is disrupted more for young compared to older adults when parafoveal preview of these words is denied, consistent with older adults processing this information less effectively (Rayner et al., 2010). However, other studies using these paradigms report conflicting findings that show no such adult age differences (Choi et al., 2017; Risse & Kliegl, 2011; Whitford & Titone, 2016).

These inconsistencies in the literature, and the fact that research to date has been conducted almost exclusively with alphabetic scripts, highlight the need for further work to better understand aging effects on parafoveal processing. The present research therefore investigated this issue using a non-alphabetic script (Chinese) which has characterisics that promote parafoveal processing and so may be well-suited to revealing age differences in the use of this information. Like English, Chinese is read from left to right. However, by contrast with alphabetic languages like English, Chinese uses a non-alphabetic script in which words

are created from box-like logograms called characters rather than letters (Campbell, 1997). Words in this script generally are short, with most comprising only one or two characters, while spaces are not used to separate words in text (see Li, X. et al., 2015; Zang et al., 2011). The Chinese script therefore tends to be physically dense compared to alphabetic scripts (Bai et al., 2008), with the result that upcoming words are typically closer to fixation, and so more available for parafoveal processing.

Consistent with this, studies using the boundary paradigm report preview benefits (i.e., a processing advantage when parafoveal words are available rather than denied) for up to two words to the right of a fixated word (e.g., Yan et al., 2010; Yang et al., 2009, 2012). In these studies, the fixated word is usually referred to as word N, while the two words to its right are words N+1 and N+2 respectively. By comparison with these findings, preview benefits for N+2 words are less common in alphabetic scripts. This is most likely because these are further from fixation and so in lower resolution vision, and such benefits have been reported only when N+1 is a short function word (e.g., Angele & Rayner, 2001; Risse & Kleigl, 2011), or N+1 and N+2 form a spaced compound (e.g., *teddy bear*, Cutter et al., 2014; see also Cui et al., 2013, Yu et al., 2016, Zang et al., 2021; and for a meta-analysis of N+1 and N+2 parafoveal preview effects in alphabetic scripts and Chinese, see Vasilev & Angele, 2017). This and other evidence suggests that parafoveal processing makes a unique contribution to reading fluency in Chinese, by allowing multiple parafoveal words to be processed to some degree, or even fully identified, on each fixation. Such findings have influenced the development of computational models of eye movement control for Chinese reading. In particular, current models emphasize the importance of parafoveal processing for reading fluency, by assuming that readers use specific processing strategies to maximize the number of characters they can identify parafoveally on each fixation, and to direct the next eye movement beyond these characters (Li & Pollatsek, 2020; see also Yu et al., 2020).

Other research shows that Chinese is read much less fluently by older than younger adults. In particular, older adults appear to read Chinese almost twice as slowly as young adults (and so appear to experience a larger reading speed slowdown as compared to older readers of alphabetic scripts). They also dwell for longer on words, by making more and longer fixations and also more regressions, and make generally shorter forward eye movements (e.g., Li, L. et al., 2019; Li, S. et al., 2018; Wang et al., 2018a,b; Zang et al., 2016; Zhao et al., 2019, 2020; for a review, see Paterson et al., 2020). These shorter forward eye movements, in particular, suggest that eye guidance may benefit less from parafoveal processing in Chinese reading for older relative to younger adults. Specifically, the finding suggests that older adults may identify fewer characters parafoveally on each fixation and so progress less far in text with each eye movement. However, few studies to date have investigated adult age differences in parafoveal processing in Chinese reading (but see Xie et al., 2020), although, as we have noted, this script's capacity to support both N+1 and N+2 parafoveal preview effects suggests it is well-suited for assessing such age differences. Consequently, such studies may be important for understanding aging effects on parafoveal processing across different writing systems, and will help inform the future development of models of eye movement control for Chinese reading (Li & Pollatsek, 2020; Yu et al., 2020).

This was therefore the focus of the present experiment. However, the potential for observing N+2 preview effects in Chinese also afforded the opportunity to assess whether these effects are contingent on the availability of preceding words (e.g., word N+1), which may shed further light on mechanisms underlying parafoveal processing. This issue has been investigated in alphabetic scripts (see, e.g., Angele et al., 2011). Especially clear evidence comes from research using spaced compounds as parafoveal words (e.g., *teddy bear;* Cutter et al., 2014), showing that N+2 preview benefits (i.e., for *bear*) are observed only when the N+1 constituent (i.e., *teddy*) is also available for parafoveal processing (i.e., is not masked).

This is explained in terms of contingencies between words, such that the first constituent licenses processing of the second as part of a larger lexical unit. Consequently, when that first constituent in masked, readers benefit less from parafoveal preview of the N+2 word. Evidence of a similar effect in Chinese reading has become available recently (Zang et al., 2021). In this study, Zang et al. presented two word idiomatic phrases with an adjective noun structure (e.g., 垫脚石, ,meaning "stepping stones"). As with the Cutter et al. study, they found that preview benefits for the second constituent of the idiom only occurred when the first constituent was parafoveally available. Crucially, however, susch effects have not been explored in older adults to date. We therefore also investigated this issue.

Accordingly, we used eye movement measures and the boundary paradigm to assess age differences in N+1 and N+2 preview benefits for young (18-30 years) and older (65+ years) Chinese readers. Participants read sentences containing a two-character word (N) followed by two one-character target words (N+1 and N+2). An invisible boundary was placed after word N, and therefore directly in front of N+1 and N+2. Prior to the reader making an eye movement that crossed this boundary, the availability of N+1 and N+2 was manipulated by showing either both words as normal or masking either N+1 or N+2, or both, using a visually similar pseudo-character. As soon as the reader's gaze crossed the boundary, these displays quickly reverted to normal.

Consistent with previous aging research, we expected that older adults' would exhibit less fluent reading, by making more and longer fixations and more regressions, and so reading more slowly than young adults. We also expected that young adults would show both N+1 and N+2 preview effects, such that using a mask to deny parafoveal preview of these words would cause readers to make longer subsequent fixations on them. Beyond this, and quite crucially, the experiment enabled an assessment of age differences in the use of parafoveal previews. One possibility is that young and older readers make similar use of this

information, so that denying N+1 and N+2 previews produces comparable increases in subsequent fixation times on words, ruling out the possibility that generally longer reading times by older adults are due to less effective parafoveal processing. Alternatively, if older adults have less effective parafoveal processing, they may exhibit reduced preview effects relative to younger adults for N+2 and possibly also N+1 words, so that they exhibit smaller or no increase in subsequent fixation times on these words when masked compared to displayed normally. Moreover, by independently manipulating the two parafoveal words, the present experiment might be informative about whether the availability of preceding words affects N+2 processing, by revealing if masking N+1 can disrupt N+2 processing, and whether any such effects differ as a function of age group.

Method

Ethics Statement. The study was approved by the research ethics committee in the Academy of Psychology and Behavior at Tianjin Normal University and conducted in accordance with the principles of the Declaration of Helsinki.

Participants. Forty-four young adults aged 18-29 years ($M = 23$ years) from Tianjin Normal University and 44 older adults aged 65-74 years ($M = 67$ years) from the local Tianjin community participated in the experiment. All were native Chinese readers who had normal or corrected vision (at least 20/40 high-contrast acuity, in Snellen values), assessed using an eye chart, and reported reading for several hours (at least) each week. Older adults were screened for non-impaired cognitive abilities using the Beijing version of the Montreal Cognitive Assessment (applying a conventional exclusion criterion of scores < 26/30; Nasreddine et al., 2005). Vocabulary and working memory were assessed for both age groups using the Vocabulary Knowledge Test from the Chinese version of WAIS-III (Wechsler et al., 2002) and the WAIS-III digit-span subtest, assessing forward and backward spans (Wechsler,1997). Vocabulary scores were similar across age groups, but digit spans were

smaller for the older adults, as is typical (Ryan et al., 2000). Table 1 summarizes participant

characteristics.

Figure 1 & Table 1

No previous study has examined adult age differences in N+1 and N+2 orthographic

preview effects in Chinese. Accordingly, we used power analysis to estimate the smallest

effect size that our design could detect with at least 80% power. We focused on estimating the

power to detect an N+2 preview effect for either age group in measures of early fixational

processing (i.e., first-fixation duration, single-fixation duration, gaze duration), using

software created by Westfall (https://jakewestfall.shinyapps.io/two_factor_power/). Default

values were used for estimates of variance components (following Judd et al., 2016). The

analysis provided estimates in the region of Cohen's $d = .45$ to .51, dependent on the eye

movement measure, which might indicate a medium or large effect size (e.g., Cohen, 1988).

We compared these estimates with calculations of observed effect sizes in published studies.

They compared well with observed effect sizes in experiments by Yu et al. (2016), which we

computed as N+2 effects of $d = .55$ for gaze durations in their Experiment 1, and $d = .48$ for

first-fixation durations, and $d = .30$ for gaze durations, in their Experiment 2. Our estimates

compared less favorably with observed effect sizes in Yan et al. (2010), which we computed

as N+2 effects of around $d = .30$ in first-fixation durations and gaze durations, and Yang et al.

(2009), which we computed as N+2 effects of between $d = .07$ and .09 in the same measures.

Finally, Yang et al. (2012) reported null effects. The large differences in observed effect sizes

across these studies may reflect variation in the stimulus characteristics, including the length

and lexical frequency of words (see, e.g., Yan et al., 2010, and Yang et al., 2012, for

discussion) as well as differences in the number of participants and stimuli per condition

across these experiments. Whereas these previous studies assessed N+2 preview effects on

N+2 target words, we examined the consequences for both N+1 and N+2 processing, which

may afford a better opportunity to detect such effects.

Materials and Design. Stimuli consisted of 60 sentences. These were presented using the boundary paradigm, with an invisible boundary placed after word N (see Figure 1), followed by two one-character target words (N+1 and N+2). Word N was a two-character word in 55 stimuli, and in the other five a multiword unit expressing a single meaning (e.g., "每个", meaning "everyone"). N+1 and N+2 words had a variety of constructions (e.g., noun+verb, noun+noun, verb+adverb, verb+verb), and were selected so they could not combine with each other or an adjacent character to form a word. Prior to a reader making a saccade that crossed the boundary, N+1 and N+2 words were shown normally (i.e., valid previews) or as invalid previews created by replacing either N+1 or N+2, or both, with a visually similar pseudo-character. The sentences were therefore shown in one of four preview conditions, which reverted to normal as soon as a saccade crossed the boundary.

The sentences were counterbalanced within each age group, so that each participant viewed each sentence in only one preview condition and an equal number of sentences in each preview condition. The sentence were viewed equally often in each preview condition across an age group. These stimuli were intermixed with 48 filler sentences, with 10 practice sentences presented at the start of the experiment. Sentence-level analyses were conducted with the between-participants factor age group (young adult, older adult). Word-level analyses for N, N+1 and N+2 were conducted with the between-participants factor age group (young adult, older adult) and within-participants factors N+1 preview validity (invalid, valid) and N+2 preview validity (invalid, valid). In each case, the invalid preview condition was used as the reference (i.e., 1 = invalid in statistical models).

Apparatus and Procedure. An EyeLink 1000 eye-tracker (SR Research), using a tower-mounted camera and chin and forehead rests, recorded each participant's right eye movements during binocular reading at 1000 Hz. Stimuli were displayed as black-on-white

text in Song font on a high-resolution (1024×768 pixels) monitor with a fast (144 Hz) refresh-rate. At 65cm viewing distance, each character subtended approximately 1° horizontally and so was of typical size for reading (Xu & Jordan, 2009).

Participants took part individually and were instructed to read normally and for comprehension. At the start of the experiment, a 3-point horizontal calibration procedure was conducted across the same line as a sentence was presented (ensuring spatial error of .3° or less for all participants). Calibration accuracy was checked before each trial and the eye-tracker recalibrated as necessary to maintain this high spatial accuracy. At the start of each trial, a fixation square equal in size to a character space was presented on the left side of the screen. Once this was fixated, a sentence was presented with the first character replacing the square. Participants pressed a key once they finished reading each sentence. This was replaced by a yes/no comprehension question on 30% of trials, which participants answered by pressing one of two keys. At the end of the experiment, participants were asked if they noticed the display changes and, if they had, how many they had noticed.

Thirty-six percent of young adults and 18% of older adults reported noticing a display change. The mean estimates of display changes by those who noticed any was 8.7%, with no participant reporting more than 10% of display changes. Therefore, following customary practice, no participant's data were excluded from statistical analyses. We note that while parafoveal preview effects appear highly reliable (see Vasilev & Angele, 2017), they can represent a mix of preview benefits and costs, whereby different masking procedures for creating invalid previews may incur a processing cost (for discussion, see, e.g., Hutzler et al., 2013; Kleigl et al., 2013; Vasilev & Angele, 2017; Vasilev et al., 2021). In the present experiment, we used visually similar characters as masks in invalid preview conditions to minimize preview costs, although such effects may also be mediated by cross-group differences in change-detection rates, as observed in the present experiment.

Results

Accuracy answering comprehension questions (analyzed using linear mixed-effects models, see subsequently) was high for all participants (> 80%), although higher for young adults (*M* = 90%, range = 80% to 100%) than older adults (M = 85%, range = 80% to 100%, β = .51, SE = .15, *z* = 3.35). Both age groups therefore comprehended sentences well, while the slightly poorer comprehension by the older adults is consistent with other evidence that older adults perform more poorly on tasks that require them to comprehend or remember information, most likely as a consequence of sensory and cognitive declines (see, e.g., Radvansky, 1999; Stine-Morrow et al., 2006).

Before analyzing the data, we combined short fixations (< 40 ms) with adjacent fixations, then deleted fixations shorter than 80 ms or longer than 800 ms. Data were analyzed only from trials in which word N was fixated during first-pass reading (i.e., and so not skipped). Trials also were excluded if: (a) there were fewer than five fixations on a sentence (affecting 43 trials, 17 for young adults and 26 for older adults); (b) a regression was made from word N during first-pass reading (affecting 679 trials, 240 for young adults and 439 trials for older adults); (c) a blink occurred during a display-change or fixation on N+1 or N+2 (affecting 154 trials, 117 for young adults and 37 for older adults); and (d) the display-change occurred prematurely (because a "j-hook" saccade crossed the boundary but terminated to its left) or the display-change was delayed (implemented > 10 ms into the next fixation; affecting 377 trials, 253 for young adults and 124 for older adults). There were 4404 trials (2266 for young adults, 2138 for older adults) following these procedures. Note that word-level data were analyzed only for trials in which word N was fixated during first-pass reading (prior to a fixation to its right and without a first-pass regression from this word), resulting in analyses based on 3556 trials (1781 for young adults, 1775 for older adults). Finally, we removed observations for each measure more than 3 SD from each participant's

mean. Table 4 reports the number of trials contributing to each word-level analysis.

Data were analyzed by Linear Mixed-Effects Models (LMEMs, Baayen et al., 2008) using R (R Development Core Team, 2019) and the lme4 package (Bates et al., 2015). For binomial variables, generalized LMEMs were conducted with the Laplace approximation. Reading times were log-transformed. Contrasts were computed using the contr.sdif function in the MASS package (Venables & Ripley, 2002). Participants and stimuli were specified as crossed random effects (Barr et al., 2013). If a full model failed to converge, the random effects structure was reduced, first by removing correlations between factors, then interactions, then random factors, first by stimuli then participants (these models are specified in R scripts available online: 10.25392/leicester.data.12652298). Following convention, t/z values > 1.96 were considered statistically significant. Effect sizes (d) are reported for critical comparisons by dividing beta values (b) by the square root of variances for intercepts, slopes and residuals, following a procedure outlined by Brysbaert and Stevens (2018).

Recently, von der Malsburg and Angele (2017) proposed that false positives can be elevated when computing multiple dependent measures in eye movement studies. We took account of this issue in the following ways. First, as we stipulate a pattern of predicted age group effects for sentence-level measures (i.e., longer reading times, more and longer fixations, more regressions, and potentially shorter forward eye movements, for older compared to younger adults), we applied no correction to these analyses. For these analyses, we report sentence reading time (time from the onset of a sentence display until the participant pressed a key to indicate they had finished reading), number of fixations, average fixation duration (mean length of all fixations), forward saccade length (mean length, in characters, of progressive eye movements) and number of regressions (backwards eye movements) as standard sentence-level measures.

At the word-level, we tested hypotheses concerning the influence of parafoveal preview

on the initial processing of N+1 and N+2 words. Accordingly, we selected eye movement measures that are sensitive to first-pass processing of these words (the initial processing of a word prior to a saccade to its right or a regression from it) for hypothesis-testing. These comprised first-fixation duration (FFD, length of the first fixation on a word); single-fixation duration (SFD, the length of a fixation that is the only first-pass fixation on a word); and gaze duration (GD, sum of all first-pass fixations). Von der Malsburg and Angele suggest two approaches to control for an inflated risk of a Type I error. One is to apply a Bonferroni correction, and the second is to consider if an effect is statistically significant in two or more measures. We adopted the second approach in the present experiment. We also report word-level measures of skipping rates (SP, the probability of not fixating a word) as exploratory analyses for words N+1 and N+2. In addition, we report first-pass processing measures (i.e., first-fixation duration, single-fixation duration, gaze duration) as exploratory analyses of effects of parafoveal preview on the foveal processing of this word (i.e., parafoveal-on-foveal effects; for discussion, see, e.g., Drieghe, 2011).

Sentence-Level Analyses. Table 2 shows sentence-level means and summarizes statistical effects. Compared to the young adults, the older adults had longer reading times, made more fixations and regressions and marginally shorter forward saccades, with no age difference in average fixation duration, resonant with previous findings (Li, L. et al., 2019; Li, S. et al., 2018; Wang et al., 2018a,b; Zang et al., 2016; Zhao et al., 2019, 2020).

Tables 2-4

Word-Level Analyses. Table 3 shows word-level means and Table 4 summarizes statistical effects for N, N+1 and N+2 analyses.

Word N. Reading times for word N were longer for older than younger adults. Total reading times were longer when N+1 had an invalid than valid preview, indicating that invalid N+1 previews disrupted reading. Age group differences at this word are in line with

sentence-level analyses, while the lack of first-pass preview effects suggests the absence of parafoveal-on-foveal effects showing sensitivity to the lexical status of parafoveal characters prior to their fixation.

Word N+1. Compared to the young adults, the older adults had lower skipping rates for N+1. We also obtained the predicted N+1 preview validity effects, with valid N+1 previews producing shorter reading times than invalid N+1 previews (in FFD, SFD, and GD, and so meeting the hypothesis-testing criterion we applied). N+2 preview validity effects were qualified by a two-way interaction between N+1 and N+2 preview validity in GD only, possibly reflecting an N+2 preview effect only when the N+1 preview was valid (valid N+1 preview, $b = 0.10$, $SE = 0.03$, $t = 3.88$; invalid N+1 preview, $b = 0.02$, $SE = 0.03$, $t = 0.92$). Beyond this, we observed a three-way interaction between age group, N+1 preview validity, and N+2 preview validity in FFD and SFD (with GD showing a similar numerical effect, and again meeting the hypothesis-testing criterion we set). Focusing on FFD and SFD, an N+2 preview effect for young adults (bs > 0.05, ts > 2.20) was obtained regardless of N+1 preview validity (i.e., with no interaction, bs < 0.05, ts < 1). However, for older adults, an N+2 preview effect was observed only when N+1 previews were valid (i.e., valid N+1 preview, N+2 preview effect, bs > 0.08, ts > 2.54; invalid N+1 preview, N+2 preview effect, bs < 0.02, ts < 0.92). This suggests that while young adults processed N+2 regardless of the lexical status of N+1, older adults processed N+2 to a lesser degree when N+1 previews were invalid than valid.

Figure 2 illustrates FFD and SFD effects for word N+1. These depict the effects of N+2 preview validity as a function of N+1 preview validity. For young adults, fixation times are longer for invalid compared to valid N+2 previews regardless of N+1 preview validity, showing that young adults are sensitive to the orthographic form of N+2 words. For older adults, fixation times are longer for invalid than valid N+2 previews only when N+1 is valid,

suggesting that their parafoveal processing of N+2 orthography is less robust.

Figure 2

Word N+2. Older adults read N+2 more slowly (in GD and TT). An N+2 preview validity effect in word-skipping reflected increased skipping for invalid N+2 previews.

Effects of individual difference variables. Despite efforts to match the young and older adult participants in terms of key individual differences, there was some between-group variance in educational background, typical between-group differences in visual acuity, vocabulary knowledge and memory capabilities (as measured using the digit span task), and between-group differences in accuracy responding to comprehension questions in the eye movement experiment. To explore the influence of these individual difference variables, we conducted additional LMM analyses of word-level measures which included chronological age and these variables as covariates. The analyses produced a pattern of effects that was very similar to that in analyses without these variables included, suggesting that effects observed in the main analyses were a consequence of age group differences rather than the influence of these individual differences. A summary of the results of these covariate analysis is included as a Supplemental Material.

The analyses also showed limited influence of the individual difference variables on eye movement measures, other than an effect of near acuity on fixation times for word N. Acuity was not predictive of preview effects, however, quite likely because conventional acuity tests provide an assessment of foveal rather than parafoveal visual abilities. Other techniques, such as measures of the visual span (e.g., Legge et al., 1997) have been shown to be sensitive to adult age group differences in parafoveal acuity for letters and characters (e.g., Liu et al., 2017; Warrington et al., 2019; Xie et al., 2019). However, whether the visual span is predictive of parafoveal preview is unclear (Risse, 2014).

Discussion

The present study investigated age differences in N+1 and N+2 preview effects in Chinese reading. The findings replicate previous evidence for less fluent reading by older adults (Li et al., 2018; Wang et al., 2018; Zang et al., 2016; Zhao et al., 2019, 2020), as well as previous evidence for N+1 and N+2 preview benefits (Yan et al., 2010; Yang et al., 2009, 2012). Specifically, we found that masking N+1 or N+2 (or both) attenuated preview benefits in subsequent fixations on words. The robustness of N+2 effects of this nature has previously been taken as evidence for more efficient parafoveal processing in Chinese reading (Vasilev & Angele, 2017), and the present findings add to this evidence.

Crucially, we found that N+1 and N+2 previews were attenuated for both age groups when these words were replaced by visually similar pseudo-characters. This provides novel evidence that older adults, as well as young adults, can obtain detailed orthographic information from at least two (one-character) words to the right of a fixated word. Ostensibly, this seems contrary to the view that older readers obtain less rightward parafoveal information (Rayner et al., 2009, 2014). However, the present findings also revealed an age difference in N+2 preview effects which might imply subtle limitations in the parafoveal processing capabilities of older relative to younger adults.

Specifically, we found that older adults produced N+2 preview effects only when the N+1 preview also was valid. By comparison, younger adults produced N+2 preview effects irrespective of N+1 preview validity. This suggests the availability of the preceding word influenced N+2 processing, although only for older adults in the present experiment. One possibility is that older, but not younger, adults did not process beyond the point from which meaningful information could be obtained. That is, their parafoveal processing may have been more limited when the lexical status of N+1 was difficult to establish (because it was a pseudo-character). This may be a consequence of age decline in cognitive resources, such that when costs associated with N+1 processing are high, fewer resources are available for

N+2 processing. This is consistent with findings by Payne and Stine-Morrow (2012) showing that increased text processing demands (in this case due to sentence wrap-up costs) can reduce preview benefits for older relative to younger adults. Why no evidence of attenuation was observed for young adults is less clear; however, this may be a consequence of using high frequency, single character words as N+2 stimuli, which may have been relatively easily identified by the younger adults even when N+1 was a pseudoword.

Clearly encountering a pseudo-character in text is unusual. It will therefore be important to establish if similar effects manifest under naturalistic conditions. This might include factors, such as the lexical frequency of N+1 words, shown previously to modulate N+2 preview benefits (Yan et al., 2010; Yang et al., 2012). These findings show that the cognitive load associated with identifying a low (relative to high) frequency word can reduce N+2 preview effects for young adults. Consequently, if older adults' parafoveal processing suffers more from cognitive load, as the present findings and those reported by Payne and Stine-Morrow (2012) suggest, increased costs associated with N+1 processing may attenuate N+2 preview effects more for older than younger adults. Such findings add to our understanding of aging effects on parafoveal processing in reading, especially in the context of current models of eye movement control in Chinese reading, which emphasize the central importance of parafoveal processing for reading efficiency (Li & Pollatsek, 2020; Yu et al., 2020).

In sum, our findings demonstrates that older adults retain the capacity to obtain detailed parafoveal information from up to two words to the right of fixation in Chinese reading, while revealing that age differences in N+2 preview benefits may emerge when processing demands are high, perhaps because of more limited cognitive resources in older age.

References

Angele, B., & Rayner, K. (2011). Parafoveal processing of word n + 2 during reading: Do the

preceding words matter? *Journal of Experimental Psychology: Human Perception and*

Performance, 37, 1210-1220. https://doi.org/10.1037/a0023096

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed

random effects for subjects and items. *Journal of Memory and Language, 59,* 390-412.

https://doi.org/10.1016/j.jml.2007.12.005

Bai, X., Yan, G., Liversedge, S. P., Zang, C., & Rayner, K. (2008). Reading spaced and

unspaced Chinese text: evidence from eye movements. *Journal of Experimental*

Psychology: Human Perception and Performance, 34, 1277-1287.

https://doi.org/10.1037/0096-1523.34.5.1277

Ball, K. K., Beard, B. L., Roenker, D. L., Miller, R. L., & Griggs, D. S. (1988). Age and

visual search: expanding the useful field of view. *Journal of the Optical Society of*

America A, 5, 2210. https://doi.org/10.1364/JOSAA.5.002210

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for

confirmatory hypothesis testing: keep it maximal. *Journal of Memory & Language, 68*,

255-278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., & Bolker, B. (2015). *lme4: Linear mixed-effects models using S4*

classes (R Package Version 0.999999–0). R-project. http://CRAN.R-

project.org/package=lme4

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models:

A tutorial. *Journal of Cognition*, 1(1), 9, 1–20, DOI: https://doi.org/10.5334/joc.10

Campbell, G. L. (1997). *Handbook of scripts and alphabets*. Psychology Press.

Chace, K. H., Rayner, K., & Well, A. D. (2005). Eye movements and phonological

parafoveal preview: effects of reading skill. *Canadian Journal of Experimental

Psychology - Revue Canadienne de Psychologie Expérimentale, 59*, 209-217.

https://doi.org/10.1037/h0087476.

Choi, W., Lowder, M.W., Ferreira, F., Swaab, T.Y., & Henderson, J.M. (2017). Effects of

word predictability and preview lexicality on eye movements during reading: A

comparison between young and older adults. *Psychology and Aging*, *32*, 232-242.

https://doi.org/10.1037/pag0000160

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 2nd ed. Hillsdale,

NJ: Erlbaum.

Cui, L., Drieghe, D., Yan, G., Bai, X., Chi, H., & Liversedge, S. P. (2013). Parafoveal

processing across different lexical constituents in Chinese reading. *Quarterly Journal of

Experimental Psychology*, *66*, 403-416. https://doi.org/10.1080/17470218.2012.720265

Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2014). Preview benefit in English spaced

compounds. *Journal of Experimental Psychology Learning Memory & Cognition, 40*,

1778-1786. https://doi.org/10.1037/xlm0000013

Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2015). How is information integrated across

fixations in reading? In A. Pollatsek and R. Treiman (Eds.), *The Oxford Handbook of

Reading* (pp. 245-260). Oxford University Press.

doi:10.1093/oxfordhb/9780199324576.001.0001

Drieghe, D. (2011). Parafoveal-on-foveal effects on eye movements during reading. In S. P.

Liversedge, I. D. Gilchrist, & S. Everling (Eds.), The Oxford Handbook of Eye

Movements (pp. 839–855). Oxford University Press.

Gordon, P. C., Lowder, M. W., & Hoedemaker, R. S. (2016). Reading in normally aging

adults. In Wright H.H. (Ed.), *Cognitive-linguistic processes and aging* (pp. 165-191).

Amsterdam, Netherlands: John Benjamins Publishing.

https://doi.org/10.1075/z.200.07gor.

Hilz, R., & Cavonius, C. R. (1974). Functional organization of the peripheral retina:

Sensitivity to periodic stimuli. *Vision Research, 14*, 1333-1337.

https://doi.org/10.1016/0042-6989(74)90006-6

Häikiö, T., Bertram, R., Hyönä, J., & Niemi, P. (2009). Development of the letter identity

span in reading: Evidence from the eye movement moving window paradigm. *Journal of

Experimental Child Psychology, 102,* 167-181.

https://doi.org/10.1016/j.jecp.2008.04.002

Hutzler, F., Fuchs, I., Gagl, B., Schuster, S., Richlan, F., Braun, M., & Hawelka, S. (2013).

Parafoveal X-masks interfere with foveal word recognition: Evidence from fixation-

related brain potentials. *Frontiers in Systems Neuroscience*, *7*, e33.

https://doi.org/10.3389/fnsys.2013.00033

Judd, C. M., Westfall, J., & Kenny, D. A. (2017). Experiments with more than one random

factor: Designs, analytic models, and statistical power. *Annual Review of Psychology*,

68, 601-625. DOI: https://doi.org/10.1146/annurev-psych-122414-033702

Kliegl, R., Hohenstein, S., Yan, M., & McDonald, S. A. (2013). How preview space/time

    translates into preview cost/benefit for fixation durations during reading. *Quarterly*

    Journal of Experimental Psychology, *66*(3), 581-600.

    https://doi.org/10.1080/17470218.2012.658073.

Li, L., Li, S., Xie, F., Chang, M., Mcgowan, V. A., Wang, J., & Paterson, K. B. (2019).

    Establishing a role for the visual complexity of linguistic stimuli in age-related reading

    difficulty: Evidence from eye movements during Chinese reading. *Attention Perception*

    & Psychophysics, 81, 2626-2634. https://doi.org/10.3758/s13414-019-01836-y

Li, S., Li, L., Wang, J., McGowan, V. A., & Paterson, K. B. (2018). Effects of word length

    on eye guidance differ for young and older Chinese Readers. *Psychology and Aging, 33*,

    685-692. https://doi.org/10.1037/pag0000258

Li, X., Zang, C., Liversedge, S.P., & Pollatsek, A. (2015). The role of words in Chinese

    reading. In A. Pollatsek, & T. Rebecca (Eds.), *The Oxford Handbook of Reading*.

    Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199324576.013.14

Li, X., & Pollatsek, A. (2020). An integrated model of word processing and eye-movement

    control during Chinese reading. *Psychological Review*, *127*(6), 1139-1162.

    https://doi.org/10.1037/rev0000248

Liu, R., Patel, B. N., & Kwon, M. (2017). Age-related changes in crowding and reading

    speed. *Scientific Reports*, *7*(1), 8271.

Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in*

    Cognitive Sciences, *4*, 6-14. https://doi.org/10.1016/S1364-6613(99)01418-7

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation

in reading. *Perception & Psychophysics, 17*, 578-586.

https://doi.org/10.3758/BF03203972

McConkie, G. W., & Rayner, K. (1976). Asymmetry of the perceptual span in

reading. *Bulletin of the Psychonomic Society, 8*, 365-368.

https://doi.org/10.3758/BF03335168

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I.,

Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA:

a brief screening tool for mild cognitive impairment. *Journal of the American Geriatric*

Society, 53, 695-699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Owsley, C. (2011). Aging and vision. *Vision Research, 51*,1610-1622.

https://doi.org/10.1016/j.visres.2010.10.020

Paterson, K.B., McGowan, V.A., Warrington, K.L., Li, L., Li, S., Xie, F., Chang, M., Zhao,

S., Pagán, A., White, S.J., & Wang, J. (2020). Effects of normative aging on eye

movements during reading. *Vision*, 4, 7. https://doi.org/10.3390/vision4010007

Payne, B. R., & Stine-Morrow, E. A. (2012). Aging, parafoveal preview, and semantic

integration in sentence processing: testing the cognitive workload of wrap-up.

Psychology and Aging, *27*, 638-649. https://doi.org/10.1037/a0026540

R Core Team. (2019). R: A Language and Environment for Statistical Computing. Vienna: R

Foundation for Statistical Computing.

Radvansky, G. A. (1999). Aging, memory, and comprehension. *Current Directions in*

Psychological Science, 8, 49-53.

Rayner, K. (1975). The perceptual span and peripheral cues during reading. *Cognitive*

   Psychology, 7, 65-81. https://doi.org/10.1016/0010-0285(75)90005-5

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual

   search. *Quarterly Journal of Experimental Psychology, 62,* 1457-1506.

   https://doi.org/10.1080/17470210902816461

Rayner, K. (2014). The gaze-contingent moving window in reading: Development and

   review. *Visual Cognition*, *22*, 242-258, DOI: 10.1080/13506285.2013.879084

Rayner, K., Castelhano, M. S., & Yang, J. (2010). Preview benefit during eye fixations in

   reading for older and younger readers. *Psychology and Aging, 25*,

   https://doi.org/10.1037/a0019199

Rayner, K., Yang, J., Schuett, S., & Slattery, T. J. (2014). The effect of foveal and parafoveal

   masks on the eye movements of older and younger readers. *Psychology & Aging, 29*,

   205-212. https://doi.org/10.1037/a0036015

Risse, S. (2014). Effects of visual span on reading speed and parafoveal processing in eye

   movements during sentence reading. *Journal of Vision*, 14(8), 11. doi:

   https://doi.org/10.1167/14.8.11.

Risse, S., & Kliegl, R. (2011). Adult age differences in the perceptual span during reading.

   Psychology and Aging, *26*, 451-460. https://doi.org/10.1037/a0021616

Ryan, J.J., Sattler, J.M., & Lopez S.J. (2000). Age effects on Wechsler Adult Intelligence

   Scale-III subtests. *Archives of Clinical Neuropsychology, 15,* 311-317.

   https://doi.org/10.1016/S0887-6177(99)00019-0

Sekuler, A. B., Bennett, P. J., & Mamelak, M. (2000). Effects of aging on the useful field of

view. *Experimental Aging Research, 26*, 103-120.

https://doi.org/10.1080/036107300243588

Schotter, E. R., Angele, B. & Rayner, K. (2012). Parafoveal processing in reading. *Attention,

Perception & Psychophysics*, *74*, 5-35. https://doi.org/10.3758/s13414-011-0219-2

Stine-Morrow, E. A., Miller, L. M., & Hertzog, C. (2006). Aging and self-regulated language

processing. *Psychological Bulletin*, *132*(4), 582–606. https://doi.org/10.1037/0033-

2909.132.4.582

Underwood, N.R., & Zola, D. (1986). The span of letter recognition of good and poor

readers. *Reading Research Quarterly,*21, 6-19. https://doi.org/10.2307/747956.

Vasilev, M. R., & Angele, B. (2017). Parafoveal preview effects from word N+1 and word

N+2 during reading: A critical review and Bayesian meta-analysis. *Psychonomic

Bulletin & Review, 24*, 666-689. https://doi.org/10.3758/s13423-016-1147-x

Vasilev, M. R., Yates, M., Pruitt, E., & Slattery, T. J. (2021). Parafoveal degradation during

reading reduced preview costs only when it is not perceptually distinct. *Quarterly

Journal of Experimental Psychology*, *74*(2), 254-276.

Veldre, A., & Andrews, S. (2015). Semantic preview benefit in English: Individual

differences in the extraction and use of parafoveal semantic information. *Journal of

Experimental Psychology: Learning, Memory, and Cognition, 41, 586-595.*

http://dx.doi.org/10.1037/ xlm0000212

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S* (4th ed.).

Springer. https://doi.org/10.1007/978-0-387-21706-2

von der Malsburg, T., & Angele, B. (2017). False Positives and Other Statistical Errors in

    Standard Analyses of Eye Movements in Reading. Journal of Memory and Language,

    94, 119-133. https://doi.org/10.1016/j.jml.2016.10.003

Wang, J., Li, L., Li, S., Xie, F., Chang, M., Paterson, K. B., White, S. J., & McGowan, V. A.

    (2018a). Adult age differences in eye movements during reading: the evidence from

    Chinese. *Journals of Gerontology: Series B, Gerontological Society of America, 73*,

    584-593. https://doi.org/10.1093/geronb/gbw036

Wang, J., Li, L., Li, S., Xie, F., Liversedge, S. P., & Paterson, K. B. (2018b). Effects of aging

    and text-stimulus quality on the word-frequency effect during Chinese reading.

    Psychology and Aging, *33*, 693-712. https://doi.org/10.1037/pag0000259

Warrington, K.L., Xie, F., Wang, J., & Paterson, K.B. (2019). Aging effects on the visual

    span for alphabetic stimuli. *Experimental Aging Research*, *45*, 387-399.

Wechsler, D. (1997). WAIS-III administration and scoring manual. The Psychological

    Corporation.

Wechsler, D., Chen, Y.H., & Chen, X.Y. (2002). WAIS-III Chinese Version Technical

    Manual. The Psychological Corporation.

Westfall, J. (2015). *PANGEA: Power analysis for general ANOVA designs.* Unpublished

    manuscript. http://jakewestfall.org/publications/pangea.pdf

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, 58-70. https://doi.org/10.1037/a0039971

Xie, F., McGowan, V. A., Chang, M., Li, L., White, S. J., Paterson, K. B., Wang, J., &

    Warrington, K. L. (2020). Revealing similarities in the perceptual span of young and

    older Chinese readers. *The Quarterly Journal of Experimental Psychology*, *73*, 1189-

    1205. https://doi.org/10.1177/1747021819899826

Xie, F., Li, L., Zhou, S., Wang, J., Paterson, K.B., White, S., & Warrington, K. (2019). Aging

    and pattern complexity effects on the visual span: Evidence from Chinese character

    recognition. *Vision*, *3*, 11.

Xu, M., & Jordan, T. R. (2009). Assessing effects of viewing distance on normal Chinese

    reading: Some methodological and theoretical implications. *Behavior Research*

    Methods, 41, 971-976. https://doi.org/10.3758/BRM.41.4.971

Yan, M., Kliegl, R., Shu, H., Pan, J., & Zhou, X. (2010). Parafoveal load of word N+ 1

    modulates preprocessing effectiveness of word N+2 in Chinese reading. *Journal of*

    Experimental Psychology: Human Perception and Performance, 36, 1669-1677.

    https://doi.org/10.1037/a0019329

Yang, J., Wang, S., Xu, Y., & Rayner, K. (2009). Do Chinese readers obtain preview benefit

    from character n + 2? Evidence from eye movements. *Journal of Experimental*

    Psychology: Human Perception and Performance, 35, 1192-1204.

    http://dx.doi.org/10.1037/a0013554

Yang, J., Wang, S., Tong, X., & Rayner, K. (2012). Semantic and plausibility effects on

    preview benefit during eye fixations in Chinese reading. *Reading and Writing, 25*, 1031-

    1052. https://doi.org/10.1007/s11145-010-9281-8

Yu, L., Cutter, M.G., Yan, G., Bai, X., Fu, Y., Drieghe, D., & Liversedge, S.P. (2016). Word

    n+2 preview effects in three-character Chinese idioms and phrases. *Language,*

    Cognition and Neuroscience, 31,1130-1149.

    https://doi.org/10.1080/23273798.2016.1197954

Yu, L., Liu, Y., & Reichle, E. D. (2020). A corpus-based versus experimental examination of

    word- and character-frequency effects in Chinese reading: Theoretical implications for

    models of reading. *Journal of Experimental Psychology: General*. Advance online

    publication. https://doi.org/10.1037/xge0001014

Zang, C., Fu, Y., Bai, X., Yan, G., & Liversedge, S.P. (2021). Foveal and parafoveal

    processing of Chinese three-character idioms in reading. Journal of Memory and

    Language, 119 . ISSN 0749-596X

Zang, C. L., Liversedge, S. P., Bai, X., & Yan, G. L. (2011). Eye movements during Chinese

    reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook*

    of eye movements (pp. 961–978). Oxford University Press.

Zang, C., Zhang, M., Bai, X., Yan, G., Paterson, K. B., & Liversedge, S. P. (2016). Effects of

    word frequency and visual complexity on eye movements of young and older Chinese

    readers. *Quarterly Journal of Experimental Psychology, 69*, 1409-1425.

    https://doi.org/10.1080/17470218.2015.1083594

Zhao, S., Li, L., Chang, M., Xu, Q., Zhang, K., Wang, J., & Paterson, K. B. (2019). Older

    Adults Make Greater Use of Word Predictability in Chinese Reading. *Psychology and*

    Aging, *34*, 780-790. https://doi.org/10.1037/pag0000382

Zhao, S., Li, L., Chang, M., Wang, J., & Paterson, K. B. (2020). A Further look at ageing and

    word predictability effects in Chinese reading: Evidence from one-character words. The

Quarterly Journal of Experimental Psychology. Advance online publication.

https://doi.org/10.1177/1747021820951131

Figure Legends

Figure 1. An example sentence stimulus.

Figure 2. Effects of N+2 preview validity as a function of N+1 preview validity in (a) first-fixation durations and (b) single-fixation durations on word N+1.

Figure 1

N+1 Preview	N+2 Preview	Sentence
Valid	Valid	黄建安**居然**|**肯讲**自己并不认可的理论观点。
	Invalid	黄建安**居然**|**肯佤**自己并不认可的理论观点。
Invalid	Valid	黄建安**居然**|**莆讲**自己并不认可的理论观点。
	Invalid	黄建安**居然**|**莆佤**自己并不认可的理论观点。
Translation		Huang Jianan is actually **willing to talk about** theoretical views that he does not agree with.

Note: Vertical lines indicate the location of the invisible boundary and words, N, N+1 and N+2 are shown in bold. The two-character word before the boundary (居然) is word N, the first character after the boundary is word N+1 (肯 in the valid preview condition), and the second character after the boundary is word N+2 (讲 in the valid preview condition). Invalid previews were visually similar pseudo-characters. The sentences were shown normally in the experiment.

Figure 2

(a)
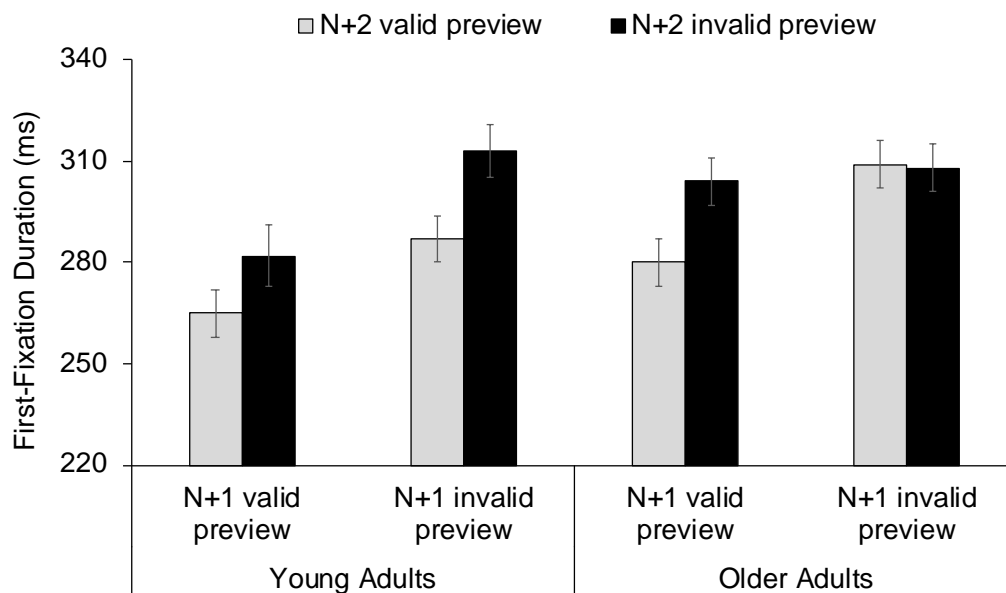
(b)
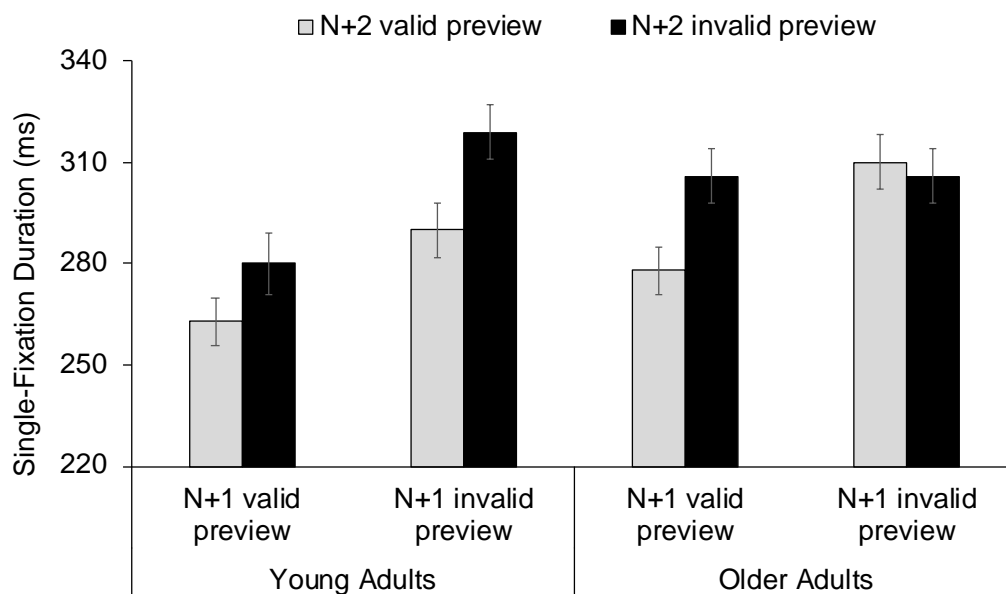
Table 1a. Summary of participant characteristics

	Young Adults	Older Adults	t
Formal education (years)	M = 15.5, SD = 2.1, Range = 12-19	M = 11.7, SD = 2.9, range = 9-16	7.14*
Visual acuity (3 m)	M = 20/17, range = 20/6 – 20/40	M = 20/24, range = 20/8 – 20/40	3.32*
Visual acuity (40 cm)	M = 16/22, range = 16/16 – 16/32	16/30, range = 16/20 – 16/50	5.85*
Vocabulary knowledge	13.16 (1.80)	14.02 (1.10)	-2.71*
Digit span	13.43 (1.58)	12.16 (2.18)	8.17*

Note. Acuity is reported in Snellen values, and vocabulary knowledge and digit span are reported as raw assessment scores (and not measures of either vocabulary or digit span size). *t*-values > 1.96 were considered significant.

Table 1b. Correlations between participant characteristics, comprehension accuracy and sentence-level measures.

	Sentence reading time	Average fixation duration	Number of fixations	Forward saccade length	Number of regressions
Age	.66*	.19	.63*	-.18	.58*
Comprehension Accuracy	-.13	.03	-.13	-.04	-.13
Formal education	-.41*	-.14	-.41*	.15	-.33*
Visual acuity (3 m)	.27*	.17	.26*	-.14	.25*
Visual acuity (40 cm)	.13	-.01	.12	.03	.04
Vocabulary knowledge	.17	-.01	.16	-.16	.14
Digit span	-.27*	-.07	-.22*	-.01	-.32*

Note. Comprehension accuracy relates to comprehension questions used in the eye movement experiment. Pearson correlations were considered significant where $p < .05$.

Table 2. Means and statistics for age-group effects in sentence-level measures

Measures	Age Group		Statistics			
	Young Adults	Older Adults	*b*	*SE*	*t/z*	*d*
Sentence reading time (ms)	3728 (28)	5813 (48)	0.42	0.05	7.67*	1.06
Average fixation duration (ms)	244 (1)	252 (1)	0.03	0.02	1.44	.12
Number of fixations	13.1 (.1)	19.1(.1)	0.35	0.05	7.22*	.97
Forward saccade length (characters)	2.68 (.02)	2.44 (.11)	-0.09	0.05	-1.95	.27
Number of regressions	3.25 (.04)	5.12 (.06)	0.47	0.07	6.46*	.76

Note. The Standard Error of the Mean is shown in parentheses. Asterisks indicate statistically significant effects at $t/z > 1.96$. Effect sizes (*d*) were calculated by dividing beta values (*b*) by the square root of variances for intercepts, slopes and residuals, as described by Brysbaert and Stevens (2018).

Table 3. Means for word-level measures

	Number of observations in analysis	Young Adults				Older Adults			
		N1 Valid		N1 Invalid		N1 Valid		N1 Invalid	
		N2 Valid	N2 Invalid	N2 Valid	N2 Invalid	N2 Valid	N2 Invalid	N2 Valid	N2 Invalid
Word N									
FFD	3552	253 (4)	250 (4)	252 (5)	253 (4)	274 (5)	269 (4)	271 (5)	276 (5)
SFD	2674	250 (5)	247 (4)	250 (5)	255 (5)	268 (6)	266 (5)	261 (5)	272 (5)
GD	3548	294 (7)	289 (7)	303 (8)	311 (9)	372 (10)	378 (10)	398 (13)	385 (10)
Word N+1									
FFD	2055	265 (7)	282 (9)	287 (7)	313 (8)	280 (7)	304 (7)	309 (7)	308 (7)
SFD	1834	263 (7)	280 (9)	290 (8)	319 (8)	278 (7)	306 (8)	310 (8)	306 (8)
GD	2054	271 (7)	310 (12)	324 (9)	338 (9)	306 (9)	344 (11)	338 (9)	347 (10)
SP	3556	.53 (.02)	.53 (.02)	.48 (.02)	.45 (.02)	.39 (.02)	.33 (.02)	.33 (.02)	.32 (.02)
Word N+2									
FFD	2306	289 (8)	285 (6)	271 (7)	279 (6)	294 (6)	289 (7)	294 (7)	295 (7)
SFD	2078	288 (7)	288 (7)	272 (7)	281 (7)	298 (7)	295 (8)	293 (8)	301 (7)
GD	2305	302 (9)	304 (8)	290 (8)	298 (8)	324 (8)	332 (9)	327 (10)	325 (8)
SP	3556	.43 (.02)	.37 (.02)	.37 (.02)	.35 (.02)	.32 (.02)	.30 (.02)	.34 (.02)	.32 (.02)

Note: SP = word skipping; FFD = first fixation duration; SFD = single fixation duration; GD = gaze duration. The Standard Error of the Mean is shown in parentheses.

Table 4. Summary of Statistical Effects for Word-Level Analyses

Measure	Effect	b	CI	SE	t/z	d
Word N						
FFD	Intercept	5.51	[5.47, 5.54]	.02	331.54	
	Word N+1	.01	[-.01, .03]	.01	.56	
	Word N+2	.00	[-.01, .03]	.01	.39	
	Age-Group	.07	[.01, .12]	.03	2.30*	.19
	Word N+1×Word N+2	.01	[-.03, .05]	.02	.49	
	Word N+1× Age-Group	.01	[-.03, .05]	.02	.55	
	Word N+2× Age-Group	.01	[-.03, .05]	.02	.50	
	Word N+1×Word N+2× Age-Group	.03	[-.05, .11]	.04	.70	
SFD	Intercept	5.50	[5.47, 5.54]	.02	323.76	
	Word N+1	.01	[-.01, .03]	.01	.79	
	Word N+2	.01	[-.01, .04]	.01	1.20	
	Age-Group	.07	[.02, .13]	.03	2.52*	.19
	Word N+1×Word N+2	.03	[-.02, .07]	.02	1.19	
	Word N+1× Age-Group	-.01	[-.05, .04]	.02	-.34	
	Word N+2× Age-Group	.02	[-.03, .06]	.02	.70	
	Word N+1×Word N+2× Age-Group	.04	[-.05, .13]	.05	.90	
GD	Intercept	5.69	[5.64,5.74]	.03	216.28	
	Word N+1	.02	[-.00, .05]	.01	1.67	
	Word N+2	.01	[-.02, .04]	.01	.66	
	Age-Group	.21	[.12, .29]	.04	4.95*	.40
	Word N+1×Word N+2	.00	[-.06, .06]	.03	-.07	
	Word N+1× Age-Group	-.01	[-.06, .05]	.03	-.14	
	Word N+2× Age-Group	.02	[-.04, .08]	.03	.76	
	Word N+1×Word N+2× Age-Group	-.01	[-.12, .11]	.06	-.13	
Word N+1						
FFD	Intercept	5.59	[5.55,5.62]	.02	298.33	
	Word N+1	.07	[.03, .10]	.02	3.79*	.17
	Word N+2	.04	[.01, .07]	.02	2.62*	.10
	Age-Group	.02	[-.05, .09]	.04	.53	
	Word N+1×Word N+2	-.03	[-.09, .04]	.04	-.79	
	Word N+1× Age-Group	-.04	[-.11, .03]	.04	-1.04	
	Word N+2× Age-Group	-.02	[-.08, .04]	.03	-.62	
	Word N+1×Word N+2× Age-Group	-.14	[-.27, -.02]	.07	-2.20*	.35

SFD	Intercept	5.59	[5.56,5.63]	.02	294.90	
	Word N+1	.08	[.04, .12]	.02	3.91*	.21
	Word N+2	.05	[.01, .08]	.02	2.60*	.15
	Age-Group	.02	[-.05, .09]	.04	.65	.
	Word N+1×Word N+2	-.03	[-.10, .05]	.03	-.71	
	Word N+1× Age-Group	-.06	[-.14, .02]	.04	-1.60	
	Word N+2× Age-Group	-.02	[-.09, .05]	.04	-.50	
	Word N+1×Word N+2× Age-Group	-.16	[-.31, -.02]	.07	-2.23*	.41
GD	Intercept	5.66	[5.61,5.70]	.02	270.15	
	Word N+1	.09	[.05, .13]	.02	4.36*	.21
	Word N+2	.06	[.02, .09]	.02	3.27*	.13
	Age-Group	.03	[-.04, .11]	.04	.87	
	Word N+1×Word N+2	-.07	[-.14, -.00]	.04	-1.98*	.15
	Word N+1× Age-Group	-.07	[-.15, .01]	.04	-1.77	
	Word N+2× Age-Group	.06	[-.07, .08]	.04	.15	
	Word N+1×Word N+2× Age-Group	-.07	[-.21, .08]	.07	-.91	
SP	Intercept	-.28	[-.46, -.11]	.09	-3.25	
	Word N+1	-.20	[-.35, -.06]	.07	-2.76*	.28
	Word N+2	-.11	[-.25, .04]	.08	-1.41	
	Age-Group	-.65	[-.97, -.32]	.17	-3.91*	.83
	Word N+1×Word N+2	.09	[-.20, .37]	.15	.60	
	Word N+1× Age-Group	.10	[-.19, .38]	.15	.67	
	Word N+2× Age-Group	-.12	[-.42, .17]	.15	-.81	
	Word N+1×Word N+2× Age-Group	.28	[-.29, .86]	.29	.97	

Word N+2

FFD	Intercept	5.58	[5.55, 5.61]	.02	358.73	
	Word N+1	-.02	[-.05, .02]	.02	-1.02	
	Word N+2	.00	[-.03, .04]	.02	.28	
	Age-Group	.03	[-.03, .09]	.03	.97	
	Word N+1×Word N+2	.05	[-.01, .11]	.03	1.59	
	Word N+1× Age-Group	.04	[-.02, .10]	.03	1.26	
	Word N+2× Age-Group	-.03	[-.10, .03]	.03	-1.06	
	Word N+1×Word N+2× Age-Group	.01	[-.11, .14]	.06	.23	
SFD	Intercept	5.59	[5.56, 5.62]	.02	342.31	
	Word N+1	-.02	[-.06, .01]	.02	-1.19	
	Word N+2	.01	[-.02, .05]	.02	.86	
	Age-Group	.04	[-.02, .11]	.03	1.35	
	Word N+1×Word N+2	.06	[-.01, .12]	.03	1.74	

	Word N+1× Age-Group	.03	[-.04, .10]	.04	.81	
	Word N+2× Age-Group	-.03	[-.10, .03]	.03	-.93	
	Word N+1×Word N+2× Age-Group	.03	[-.10, .16]	.07	.51	
GD	Intercept	5.64	[5.61, 5.68]	.02	313.96	
	Word N+1	-.01	[-.05, .02]	.02	-.78	
	Word N+2	.02	[-.03, .06]	.02	.85	.
	Age-Group	.07	[.01, .14]	.03	2.11*	0.40
	Word N+1×Word N+2	.03	[-.04, .11]	.04	.83	
	Word N+1× Age-Group	.01	[-.06, .09]	.04	.39	
	Word N+2× Age-Group	-.02	[-.10, .05]	.04	-.58	
	Word N+1×Word N+2× Age-Group	.01	[-.14, .16]	.08	.11	
SP	Intercept	-.63	[-.76, -.51]	.06	-9.96	
	Word N+1	-.05	[-.21, .11]	.08	-.70	
	Word N+2	-.14	[-.30, .02]	.08	-1.99*	0.26
	Age-Group	-.22	[-.45, -.01]	.11	-2.01*	0.50
	Word N+1×Word N+2	.09	[-.20, .38]	.15	.65	
	Word N+1× Age-Group	.24	[-.05, .56]	.16	1.56	
	Word N+2× Age-Group	.11	[-.18, .42]	.15	.73	
	Word N+1×Word N+2× Age-Group	-.19	[-.79, .37]	.29	-.66	

Note. Asterisks indicate statistically significant effects at $t/z > 1.96$. Effect sizes (*d*) for critical comparisons were calculated by dividing beta values (*b*) by the square root of variances for intercepts, slopes and residuals, as described by Brysbaert and Stevens (2018).