Proficient L2 readers do not have a risky reading strategy

Fernandez, Leigh; Xheladini, Agnesa; Allen, Shanley

Published in:
Linguistic Approaches to Bilingualism

DOI:
10.1075/lab.22064.fer

Publication date:
2023

Citation for published version (APA):
Fernandez, L., Xheladini, A., & Allen, S. (2023). Proficient L2 readers do not have a risky reading strategy. *Linguistic Approaches to Bilingualism*, *13*(6), 854-872. https://doi.org/10.1075/lab.22064.fer

Proficient L2 readers do not have a risky reading strategy

Leigh B. Fernandez[1], Agnesa Xheladini[1], & Shanley E. M. Allen[1]

[1]RPTU Kaiserslautern-Landau

Abstract

Proficient first-language (L1) readers of alphabetic languages that are read left-to-right typically have a perceptual span of 3-4 characters to the left and 14-15 characters to the right of the foveal fixation. Given that second-language (L2) processing requires more cognitive resources, we hypothesize that L2ers will have a smaller perceptual span than L1ers, and may rely on a compensatory risky reading strategy with a more symmetrical perceptual span similar to that seen in older L1 adults. Here, we test the size and symmetry of the perceptual span in German L1/English L2ers reading in English. We manipulate the amount of information available (3,6,9 characters-left/3,9,15 characters-right) during reading, and also account for the influence of English skills. Results show that L2ers benefit from an increase of window size from 3 to 6 characters to the left, and from 3 to 9 characters to the right, with higher-skilled L2ers further benefiting from an increase to 15 characters to the right. Contrary to our hypothesis, proficient L2ers exhibit an asymmetric perceptual span similar to college-aged L1ers and do not employ a compensatory risky reading strategy. This suggests that L1 and L2 language processing are not qualitatively different, but are rather modulated by individual differences.

Keywords: perceptual span, risky reading strategy, L2 reading, parafoveal processing

Proficient L2 readers do not have a risky reading strategy

1. Introduction

Considerable debate has focused on whether second language (L2) speakers process language in the same way as first language (L1) speakers, and whether the mechanisms employed during L1 and L2 processing are qualitatively similar (e.g., Kaan, 2014; Kaan & Grüter, 2021). More recently, however, research agendas have shifted to focus more on the factors that modulate L2 differences, rather than focusing on qualitative L1 and L2 differences. This is particularly salient in the L2 prediction literature, where it has been argued that L1 and L2 processing mechanisms do not differ qualitatively, but are rather modulated by sources of variation such as the quality of lexical representations, language exposure, processing strategies, etc. (Kaan, 2014; Kaan & Grüter, 2021). In the current study, we follow this line of reasoning and apply it to L2 reading strategies. To the knowledge of the authors, very little research has systematically investigated classical oculomotor behavior such as parafoveal processing in L2 readers, and particularly whether L2 readers engage in a compensatory reading strategy through their parafoveal processing patterns. Therefore, we investigate foveal and parafoveal processing during reading by L2 speakers of English with L1 German using the Gaze Contingent Moving Window paradigm (GCMWP; McConkie & Rayner, 1975) while controlling for both L2 quality of lexical representation and L2 proficiency.

When making a fixation during reading, we receive information primarily from the central two degrees of our visual field, known as the foveal area, which is the point of highest visual acuity. However, we are typically also able to extract information a further 2-5 degrees, from what is known as the parafoveal area. While visual acuity may not be as high in the parafoveal area, the processing of parafoveal information is important for fluent reading as it allows us to plan upcoming eye-movements as well as pre-process upcoming information. The GCMWP is used to establish the perceptual span, which is the amount of foveal and parafoveal information that is necessary to read fluently. Using this paradigm, it has been well established that L1 speakers of alphabetic languages, such as English, have an asymmetrical perceptual span consisting of approximately 4 characters to the left and 15 characters to the right during reading (Schotter et al., 2012). The perceptual span is not fixed and may decrease in the face of cognitive difficulty, both in terms of the text itself and in terms of individual differences of the readers. For example, the perceptual span decreases when the text being read is more complex (e.g., Henderson & Ferreira, 1990; Veldre & Andrews, 2014), individuals with higher reading skills have a larger perceptual span (e.g., Häikiö et al., 2009; Rayner et al., 2010; Veldre & Andrews, 2014), and readers with a higher quality of lexical representations are more efficient at extracting parafoveal information and have larger perceptual spans (see Andrews et al., 2020)[1]. It has also been found that L2 speakers have a smaller perceptual span relative to L1 speakers of the same language, at least in part due to the increased cognitive costs associated with reading in a second language (e.g., Leung et al., 2014).

Interestingly, it has been found that older adult L1 speakers use a "riskier" reading strategy relative to younger adults (e.g., Rayner et al., 2006); in particular, they exhibit smaller more symmetrical perceptual spans (in the form of a larger left span and smaller right span), longer forward saccades, slower reading times, increased skipping rates, and less proficient use of the parafoveal area. This strategy has been argued to reflect a reliance on contextual information (to guess upcoming information) as an attempt to compensate for age-related declines and an overall decline in lexical processing efficiency. However, not all research has found that older adults use a risky reading strategy or are less efficient at extracting parafoveal information. Whitford and Titone (2015) found that while older English/French bilingual adults showed slower reading times and more regressions, they had a rightward perceptual span of 14 characters (they did not investigate leftward span), similar to that of younger readers. Risse and Kliegl (2011) found that older German readers were also able to extract parafoveal information to the same extent as younger German readers[2].

Veldre et al., (2021) speculated that the inconsistency in finding evidence of the risky reading strategy in older adults may be the result of individual differences, which were not controlled consistently across previous research. Therefore, they designed a study to test the symmetry of the perceptual span of older L1 readers, while also controlling for written proficiency (in terms of vocabulary, reading rate, comprehension, spelling skill, and exposure to print) and manipulating text difficulty (with low frequency/high difficulty and high frequency/low difficulty text – for ease we refer to the former only as low frequency and the latter as high frequency). In terms of the right span, they found that older adults with higher proficiency showed a right perceptual span up to 15 characters (the same as college-aged adults, Veldre et al., 2014), while older adults with lower proficiency showed a right perceptual span up to 9 characters. In terms of the left span, they found that regardless of proficiency, older adults had a larger left span than college-aged adults (Veldre et al., 2014). In addition, they found that older adults were less likely to regress back into the text as the left span increased, suggesting that increasing the left span allowed for greater reading fluency (and not increased reanalysis). Veldre et al. (2021) argued that if older adults were indeed using a risky reading strategy based on guessing from context, this should be reflected in less efficient parafoveal processing to the right of the fixation (smaller right span) in conjunction with a larger left span and more regressions to the left. In addition, they found that frequency only impacted reading behavior in terms of the right span. Given that they did not find increased regressions with a larger left span and that they did not find parallel effects on both the left and right span in terms of proficiency and frequency, they argued that older adults do not engage in a risky reading strategy. However, the increased left span relative to college-aged L1 speakers suggests that older adults do use the left span to engage in late confirmatory processing (to reconfirm what they previously read) rather than to reanalyze risky guesses.

Like older L1 speakers, L2 speakers exhibit slower lexical access (e.g., Shook et al., 2016), slower reading times (e.g., Cop et al., 2015), smaller perceptual span size (e.g., Leung et al., 2014), and seem to be less efficient at extracting parafoveal information (e.g., Leung et al., 2014) relative to L1 speakers. Fernandez et al. (2020) even found slightly higher rates of skipping by L2 speakers relative to college-aged L1 speakers (though it did not reach significance), which led them to tentatively argue that L2 readers may be employing a risky reading strategy.

Given the parallels between older reading behavior and L2 reading behavior, we conduct a conceptual replication of Veldre et al. (2021) in an aim to directly test whether L2 speakers employ a risky reading strategy. If L2 speakers and L1 older speakers pattern similarly, the cognitive reasons for these similarities may or may not be the same. For example, older adults are more likely to have age-related cognitive and sensory changes that can impact language processing (e.g., reduced inhibition, declines in vision, hearing loss, etc.; e.g., Salthouse, 2010), which is unlikely to be the case with college-aged L2 speakers. However, it has also been argued that decreased performance in psychometric tasks as we age may not reflect cognitive decline, but rather reflect memory demands associated with increased knowledge and larger vocabularies (Ramscar et al., 2014). Within this framework it can be argued that L2 speakers and older L1 may have similar reasons for their performance in that both groups will have a larger vocabulary which may increase processing demands. For example, it has been found that L2 speakers pattern more closely to older monolinguals (relative to younger monolinguals) in picture-naming latencies (Gollan et al., 2008) which may reflect the increased processing costs associated with having a bilingual lexicon. While the exact cognitive and perceptual reasons may be different, we believe that it is possible L2 speakers may use a similar compensatory strategy to older L1 speakers. Alternatively, L2 speakers may behave qualitatively similar to college-aged L1 speakers (Kaan, 2014) in terms of an asymmetric perceptual span (rather than a more symmetrical perceptual span driven by a larger left span), with reading behaviors being modulated by individual differences such as L2 proficiency and quality of lexical representation.

Therefore, in the current study, we aim to test both the left and right perceptual span of late L2 speakers of English (with an L1 of German), controlling for both L2 proficiency and L2 spelling skill (as a measure of quality of lexical representation), in a conceptual replication of Veldre et al. (2021) (using the same paradigm, items, and statistical approach). To the knowledge of the authors, this is the first study directly investigating the symmetry of the L2 perceptual span, and the extent of the leftward span in L2 speakers. To directly test whether L2 speakers behave similarly to older L1 adults, we compare the L1 older adult reading data from Veldre et al. to the current L2 data. If L2 speakers adopt a risky reading strategy, we hypothesize that they will have a smaller and more symmetrical perceptual span relative to college-aged L1 speakers, but similar to older L1 adults. If they adopt a late confirmatory processing strategy similar to Veldre et al. (2021), we hypothesize that they will show a larger left span (up to window size 9) but a similar right span size (up to window size 15) relative to college-aged L1 speakers. If they behave more similarly to college-aged L1 adults we hypothesize that they will have an asymmetrical perceptual span approximately 4 characters to the left and 15 to the right of the fixation, with the span being modulated by individual differences.

2. Methods

2.1 Participants

A total of 60 college-aged L1 German speakers were recruited from the University of *removed for review*. All participants were L1 speakers of German raised in a monolingual German household, and did not learn a second language before the age of 6. All had normal or corrected-to-normal vision. See Table 1 and the supplementary material (S1) for additional information. We consider these participants to be high-intermediate to high proficiency speakers based on their proficiency score[3].

Table 1: Participant information (standard deviation in parentheses)

N (male/female)	Mean age	Mean proficiency score (%)	Mean spelling score (%)	Mean Age of English Acquisition
60 (41/19)	24.60 (3.38)	75.66 (12.15)	75.90 (7.34)	9.52 (2.19)

2.2 Apparatus

Viewing was binocular, but only the right eye was tracked using an EyeLink 1000 sampling at 1000 Hz. Stimuli were presented on a Dell P1130 19” flat screen cathode ray tube (1024 X 768 resolution; 120 Hz refresh rate) with approximately 3.5 characters subtended 1° of visual angle. Participants sat approximately 70 cm away from the monitor with their head stabilized using a chin rest.

2.3 Materials

The critical items were taken from Veldre and Andrews (2014) and consisted of 63 sentences pairs of approximately 10 words in length. Each pair consisted of a high frequency

(HF) and a low frequency (LF) version. Frequency was manipulated such that the sentence contained either three high frequency words (*finished the healthy food*) or three low frequency words (*devoured the nourishing feast*), as shown in the examples in Table 2; see Veldre & Andrews (2014) for additional item information.

Table 2

Frequency	**Sentence**
High	The man finished the healthy food before relaxing on the sofa.
Low	The man devoured the nourishing feast before relaxing on the sofa.

Sentences were presented using the GCMW paradigm. The window sizes were 3, 6, or 9 letter spaces to the left and 3, 9, or 15 letter spaces to the right of the fixation, and characters outside of the window were replaced with upper case Xs (see Appendix 1). Nine versions were created and were rotated in a fully factorial design. Participants saw both versions of each sentence pair, but the HF and LF versions were separated such that the participant would see one of each pair in the first half of the study and the other of the pair in the second half of the study. All sentences were presented across one line in Courier New font (monospaced) size of 14 point.

Participants completed the English Oxford Placement Test (OPT) proficiency test involving morphosyntactic knowledge where they selected a sentence continuation from three options. The test consisted of 50 items and the score was converted into a percentage. Participants also completed a misspelling identification task to measure their quality of lexical representation. Participants identified misspelled words from a list of 215 words (of which 51 words were misspelled). The score was calculated as a percentage taking into account hits, misses, false alarms, and correct rejections (see Table 1). The proficiency score and spelling score were combined (spelling score + proficiency score/2) to give us a measure of English skill (see supplementary information (S2) for visualization of these measures). The proficiency and spelling scores were moderately correlated (r=.56).

2.4 Procedure

Participants first filled out a language background questionnaire to provide information about their language experience. This was followed by the OPT, and the spelling task. They then took part in the eye-tracking task which consisted of 4 practice trials, and two blocks of 63 sentences. A 3-point calibration was performed prior to the practice trials, between the blocks, and if the participant took a break and removed their head from the chinrest. Each trial began with a fixation dot that participants had to fixate on for the sentence to appear; if this was not successful the participant was recalibrated. After 1/3 of the trials a true/false comprehension question was displayed probing the interpretation of the sentence that was just read. The study took approximately 75 minutes and participants were compensated 10 Euro or participant credit hours.

2.5 Analysis

Comprehension accuracy was good (ca. 87%), indicating that participants were reading for comprehension. Accuracy is not considered further. The main dependent variable in the current study was reading rate operationalized as words read per minute (WPM). Additionally, we looked at forward saccade length (FSL), which is the length of saccades made in a forward direction (in number of character spaces), forward fixation duration (FFD) which is the duration of fixations made in a forward direction (in milliseconds), and the regression count which is the number of regressions that were made backward through the text (left). Given the likelihood of increased false positives in the eye-tracking measures (due to multiple comparisons between dependent measures), we a-priori decided to focus on four dependent variables. Thus, we applied a Bonferroni correction yielding an alpha threshold of .0125 (.05/4 = .0125; see von der Malsburg and Angele, 2017).

Eye movement data were analyzed using linear mixed effects models in R (R Core Team, 2018) and lme4 (Bates et al., 2018), while p-values were estimated from the lmerTest package (Kuznetsova et al., 2018). The models included fixed effects of window size with sliding contrasts (right: 3 vs. 9, 9 vs. 15; left: 3 vs. 6, 6 vs. 9), frequency with sum contrasts (–.5/.5), and the continuous centered combined English skill predictor (spelling score + proficiency score/2). The model included two-way interactions between frequency and window size, and between English skill and window size. Omnibus tests for main effects and interactions were run using log likelihood ratio tests, and only contrasts from comparisons reaching significance are reported and discussed; see the supplementary materials (S3) for omnibus test output. The random effects structure was maximally specified (Barr et al., 2013). In the case of convergence errors or singularity issues, the model was reduced in complexity until it successfully converged without singularity issues. Below t and p values are reported; for full model information, see Appendix 2. See Appendix 3 for mean values (see OSF (https://osf.io/7jrcb/) for data and statistical code).

3. Results

3.1 Words per minute (WPM)

WPM increased for high frequency items (t = 8.76, p < .0001), from left 3 to 6 (t = 4.30, p < .0001) and from right 3 to 9 (t = 11.13, p < .0001). There were also two interactions: right 3 to 9 with English skill (t = 2.81, p < .01) and right 9 to 15 with English skill (t = 2.69, p < .01). To break down these interactions, we separated the L2 speakers into two groups by English skill (low scorers: 55-77% (n=30) / high scorers: 76-92% (n=30)) and visualized them for easier interpretation; see Figure 1. The high scorers showed a greater increase in WPM from 3 to 9 relative to the low scorers, and the high scorers showed an additional increase from 9 to 15 while low scorers showed no such increase.

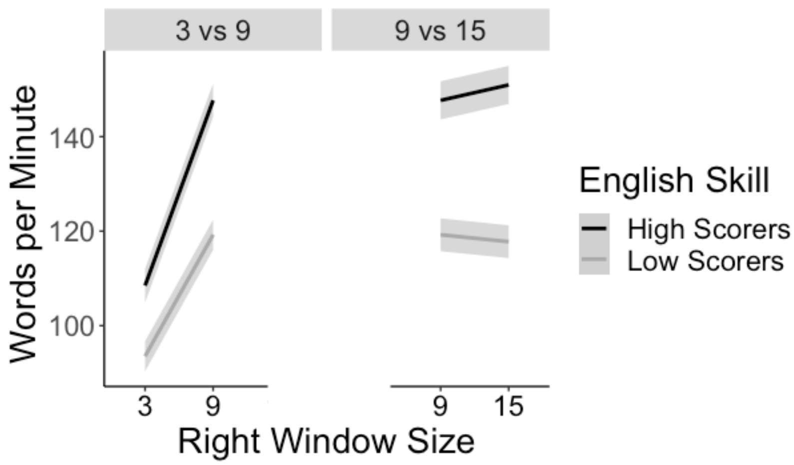

Figure 1: Right window WPM for high and low English skill scorers.

3.2 Forward saccade length (FSL)

FSL increased for high frequency items (t = 4.72, p < .0001), from left 3 to 6 (t = 3.44, p < .01), and from right 3 to 9 (t = 17.31, p < .0001) and 9 to 15 (t = 6.54, p < .0001). There was also an interaction between right 3 to 9 and English skill (t = 2.71, p < .01), with high scorers showing a greater increase in FSL from 3 to 9 relative to the low scorers. Further, there was an interaction between right 3 to 9 and frequency (t = 2.91, p < .01), with FSL increasing more for high frequency items with the increase from window size 3 to 9 relative to the low frequency items.

3.3 First fixation duration (FFD)

Given the inherent skew in duration measures, FFD was log transformed. FFD decreased for high frequency items (t = -7.17, p < .0001), from left 3 to 6 (t = -5.99, p < .0001), and from right 3 to 9 (t = -14.20, p < .0001) and 9 to 15 (t = -3.88, p < .001). Additionally, there was an effect of English skill (t = -3.42, p < .01) with FFD decreasing with English skill.

3.4 Regression count

Backwards regressions decreased for high frequency items (t = -33.07, p < .0001) relative to low frequency items.

3.5 Combined analysis

To determine whether L2 speakers behaved similarly to older adults, we directly compared the current WPM data to the older adult data from Veldre et al. (2021) which used the same items and paradigm. See the supplementary materials (S4) for the outcome of the omnibus tests and models. The main patterns revealed that L1 speakers showed greater WPM benefits than L2 speakers when increasing the left window 3 to 6, and when increasing the right window 3 to 9, and right window 9 to 15. Due to differences in the individual differences tests across the two studies, they could not be included in the analysis.

4. Discussion

In the current study we used the GCMWP to test the left and right perceptual span of college-aged L2 speakers of English (with German L1) in order to investigate L2 reading strategy. Older adult L1 speakers tend to have a smaller and more symmetrical perceptual span, which may reflect a “riskier” reading strategy relative to younger adults (Rayner et al., 2006). When individual differences are taken into account, however, older L1 readers seem to show a similar right perceptual span to college-aged L1 readers. Nonetheless, they still show a greater left perceptual span which may reflect late confirmatory processing as opposed to reanalysis from risky reading (Veldre et al., 2021). L2 speakers exhibit slower lexical access, slower reading times, smaller perceptual span size, and seem to be less efficient at extracting parafoveal information relative to college-aged L1 speakers (e.g., Cop et al., 2015; Leung et al., 2014). Given the similarities between older L1 reading and college-aged L2 reading, it has even been tentatively argued that college-aged L2 readers may be employing a risky reading strategy

(Fernandez et al., 2020). Alternatively, college-aged L2 speakers may pattern similarly to college-aged L1 readers with their perceptual span being asymmetric and modulated by individual differences (Kaan, 2014). However, to our knowledge no research has investigated the possibility of a risky reading strategy or the extent of the left span in L2 speakers.

We found evidence that L2 speakers show an asymmetrical perceptual span that is modulated by English skill, similar to what is seen with college-aged L1 readers. In terms of the left perceptual span, L2 speakers showed a benefit from the increase of window size 3 to 6, but not 6 to 9, across WPM, FSL, and FDD. In terms of the right span, WPM increased from 3 to 9 and from 9 to 15, with the latter increasing only for those who scored high on English skill. In addition, we found that FSL increased from 3 to 9 (and this increase was greater for high frequency items and for those who scored high on English skill) and from 9 to 15, and that FFD decreased from 3 to 9 and from 9 to 15. This pattern is very similar to what was found with college-aged L1 speakers using the same materials (Veldre et al., 2014), but not with older L1 speakers using the same materials (who further benefit from an increase from the left window 6 to 9; Veldre et al., 2021) or with older adults using different materials (Rayner et al., 2006). Therefore, we argue that skilled L2 speakers do not engage in a risky reading strategy or a late confirmatory strategy. Rather they behave strikingly similarly to college-aged L1 readers. This suggests that L1 and L2 reading strategies do not differ qualitatively, but are modulated by individual differences (Kaan, 2014).

Additionally, we found that higher frequency items led to faster reading times (WPM), longer FSLs, shorter FFDs, and fewer backward regressions relative to low frequency items. This is in line with both L1 literature (Veldre et al., 2014, 2021) and L2 literature (Fernandez et al., 2021) using the same materials. We also found that as English skill increased FFD decreased. Additionally, English skill interacted with the right span in several ways (as discussed above), but never the left span. This further supports the idea that L2 speakers do not engage in a risky reading strategy, as that would assume parallel effects of text frequency and individual differences on both the left and right span (Veldre et al., 2021). These findings again highlight the similar reading behavior between college-aged L1 and L2 speakers, and show the importance of controlling for individual differences in reading studies for both L1 and L2 speakers (e.g., Cop et al., 2015; Veldre et al., 2014, 2021).

When we directly compare the WPM data from the L2 speakers and older L1 data, we found that older L1 speakers have a faster WPM, and both groups showed an increase in WPM from left 3 to 6, and from right 3 to 9 and 9 to 15. However, this benefit was greater for L1 speakers relative to L2 speakers. We urge some caution when interpreting these results given that we could not include the individual difference measures in the comparison (due to differences in the tests across the two studies). It is clear from the L2 data here and the L1 data in Veldre et al. (2021) that individual differences play an important role during reading, so we encourage future research comparing these groups using the same standardized individual difference measures (e.g., Andrews et al., 2020).

Overall, this data suggests that when the individual differences of proficiency and quality of lexical representation (in terms of spelling skill) are taken into account, L2 speakers show a similar perceptual span to college-aged L1 readers - approximately 6 characters to the left and 15 to the right (e.g., Schotter et al., 2012; Veldre et al., 2014). We argue that previous research finding that L2 speakers show a smaller perceptual span has not consistently accounted for proficiency in their models, has used different individual difference measures (e.g., Leung et al., 2014; Whitford & Titone, 2015), or may have tested a subset of lower proficiency speakers,

speakers with lower quality of lexical representation, or some combination (given that we found in the current study that only low proficiency readers have a restricted right span). Again, the present study confirms the importance of individual differences for assessing L2 perceptual span and parafoveal processing.

One alternative possibility not yet discussed is that college-aged L2 speakers may perform similarly to developing L1 readers. Indeed, Cop et al. (2015) compared reading behavior of college-aged L2 speakers, college-aged L1 speakers, and developing L1 readers. They found several similarities between the college-aged L2 speakers and developing L1 readers which they argued both stem from slower lexical access (see Reichle et al, 2013), though the differences between developing L1 and college-aged L1 readers were greater than the differences between college-aged L1 and L2 speakers. Developing readers also show small increases in right perceptual span size with age, which cannot fully account for the subsequent increase in reading speed. Rather, reading skill improves with age, reflecting more global improvements in things like general knowledge and statistical learning (Meixner et al., 2022). As a result, we hesitate to draw too many parallels between developmental reading and L2 processing, given that adult L2 readers have already learned reading skills in their L1 and have greater knowledge and experience. Additionally, to the knowledge of the authors, there is very limited research investigating the left perceptual span in developing readers (e.g., Rayner, 1986), so it is unclear how our data will compare to developing readers, but we believe this could be a compelling line of future research.

Additionally, a methodological issue is worth briefly discussing. In the current study we used X's (or X-masks) to determine perceptual span size (conceptually replicating previous work). However, research investigating the role of X-masks on reading behavior has shown that they can cause interference that leads to inhibition and inflated reading times (for an L1 meta-analysis see Vasilev & Angele, 2017; for evidence in L2 reading see Fernandez et al., 2020). However, there is no consensus on what is the best type of mask to use when testing the perceptual span. Future research investigating whether these effects remain with less obtrusive masks would be welcome.

In conclusion, we found that L2 speakers do not engage in either a risky reading strategy or a late confirmatory strategy to compensate for declines when reading in their L2. Rather, proficient L2 speakers behave similarly to college-aged L1 readers, with both showing an asymmetrical perceptual span approximately 6 characters to the left and (when proficiency is taken into account) 15 characters to the right (Veldre et al., 2014). This may seem surprising given that previous research has found that L2 speakers seem to use parafoveal information less efficiently (e.g., Cop et al., 2015; Fernandez et al., 2021; Leung et al., 2014), have slower lexical access (Shook et al., 2016), and have smaller perceptual spans (e.g., Leung et al., 2014). It may be that college-aged proficient L2 speakers have enough cognitive resources available to invest in both foveal and parafoveal processing and/or L2 speakers with high enough proficiency and quality of lexical representation do not need to invest in a strategy (because they have nothing to compensate for). This study lends support to research showing that the mechanisms employed during L1 and L2 language processing (in this case reading/reading strategies) are not qualitatively different, but are rather modulated by individual differences (e.g., Kaan, 2014; Kaan & Grüter, 2021).

5. References

Andrews, S., Veldre, A., & Clarke, I. E. (2020). Measuring lexical quality: The role of spelling ability. *Behavioral Research Methods, 52,* 2257–2282. https://doi.org/10.3758/s13428-020-01387-3

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255-278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D. M., Maechler, M., & Bolker, B. (2018). *lme4: Linear mixed-effects models using S4 classes*, R package version 1.1-27.1.

Cop, U., Drieghe, D., & Duyck, W. (2015). Eye movement patterns in natural reading: A comparison of monolingual and bilingual reading of a novel. *PLoS ONE, 10*, 1-38. https://doi.org/10.1371/journal.pone.0134008

Fernandez, L. B., Bothe, R. & Allen. S. E. M. (2021). The role of L1 reading direction on L2 perceptual span: An eye tracking study investigating Hindi and Urdu. Second Language Research. https://doi.org/10.1177/02676583211049742

Fernandez, L. B., Scheepers, C., & Allen, S. E. M. (2020). The impact of uninformative parafoveal masks on L1 and late L2 speakers. *Journal of Eye Movement Research, 13*(6), 3. https://doi.org/10.16910/jemr.13.6.3

Gollan, T. H., Montoya, R. I., Cera, C. M., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language, 58*, 787–814.

Häikiö, T., Bertram, R., & Hyönä, J. (2010). Development of parafoveal processing within and across qords in reading: Evidence from the boundary paradigm. *Quarterly Journal of*

Experimental Psychology, 63(10), 1982-1998. https://doi.org/10.1080/17470211003592613

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16,* 417–429. https://doi.org/10.1037//0278-7393.16.3.417

Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism, 4*(2), 257-282. https://doi.org/10.1075/lab.4.2.05kaa

Kaan, E., & Grüter, T. (2021). Prediction in second language processing and learning: Advances and directions. In: Kaan, E. and Grüter, T (Eds). *Prediction in second-language processing and learning*, pp 1-24, John Benjamins.

Kuznetsova, A., Brockhoff, P. B., & Bojesen, C. (2018). *lmerTest: Tests in linear effects models,* R package version 3.1-3.

Leung, C. Y., Sugiura, M., Daisuke, A., & Yoshikawa, L. (2014). The perceptual span in second language reading: An eye-tracking study using a gaze-contingent moving window paradigm. *Open Journal of Modern Linguistics, 4*, 585-594. https://doi.org/10.4236/ojml.2014.45051

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

Meixner, J. M., Nixon, J. S., & Laubrock, J. (2022). The perceptual span is dynamically adjusted in response to foveal load by beginning readers. *Journal of Experimental Psychology, 151*, 1219–1232. https://doi.org/10.1037/xge0001140

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–213). Amsterdam, The Netherlands: John Benjamins. https://doi.org/10.1075/swll.11.14per

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*, 5-42.

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology*, *41*(2), 211-236. https://doi.org/10.1016/0022-0965(86)90037-8

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*, 448– 465. https://doi.org/10.1037/0882-7974.21.3.448

Rayner, K., Slattery, T. J., & Bélanger, N. N. (2010). Eye movements, the perceptual span, and reading speed. *Psychomic Bulletin & Review*, *17*, 834–839. https://doi.org/10.3758/PBR.17.6.834

Reichle, E. D., Liversedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S., White, S. J., & Rayner, K. (2013). Using E-Z Reader to examine the concurrent development of eye-movement control and reading skill. *Developmental review, 33*(2), 110–149. https://doi.org/10.1016/j.dr.2013.03.001

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

Salthouse, T. A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychology Society, 16*, 754–760. https://doi.org/10.1017/S1355617710000706

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.

Shook, A., Goldrick, M., Engstler, C., & Marian, V. (2015). Bilinguals show weaker lexical access during spoken sentence comprehension. *Journal of Psycholinguistic Research*, *44*(6), 789–802. https://doi.org/10.1007/s10936-014-9322-6

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

Veldre, A., & Andrews, S. (2014). Lexical quality and eye movements: Individual differences in the perceptual span of skilled adult readers. *The Quarterly Journal of Experimental Psychology*, *67*, 703–727. https://doi.org/10.1080/17470218.2013.826258

Veldre, A., Wong, R., & Andrews, S. (2021). Reading proficiency predicts the extent of the right, but not left, perceptual span in older readers. *Attention Perception & Psychophysics*, *83*, 18–26. https://doi.org/10.3758/s13414-020-02185-x

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/0.1016/j.jml.2016.10.003

Whitford, V., & Titone, D. (2015). Second-language experience modulates eye movements during first- and second-language sentence reading: Evidence from a gaze- contingent moving window paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*, 1118-1129. https://doi.org/10.1037/xlm0000093

Competing interests

Competing interests: The author(s) declare none.

Data availability

The data and code that support the findings of this study are openly available on the Open Science Framework (https://osf.io/7jrcb/).

Address for all authors
RPTU Kaiserslautern-Landau
Psycholinguistik und Sprachentwicklung
Postfach 3049
67653 Kaiserslautern Germany

Email addresses
Leigh B. Fernandez: leigh.fernandez@rptu.de
Agnesa Xheladini: xheladin@rptu.de
Shanely E. M. Allen: allen@rptu.de

Appendix 1. Window size manipulations, the * represents the fixation location (L= Left window size/ R=Right window size)

Window size	Sentence
3L 3R:	XXX XXX XXXXshed thx XXXXXXX XXXX XXXXXX XXXXXXXX XX XXX XXXX. *
3L 9R:	XXX XXX XXXXshed the healXXX XXXX XXXXXX XXXXXXXX XX XXX XXXX. *
3L 15R:	XXX XXX XXXXshed the healthy foXX XXXXXX XXXXXXXX XX XXX XXXX. *
6L 3R:	XXX XXX Xinished thx XXXXXXX XXXX XXXXXX XXXXXXXX XX XXX XXXX. *
6L 9R:	XXX XXX Xinished the healXXX XXXX XXXXXX XXXXXXXX XX XXX XXXX. *
6L 15R:	XXX XXX Xinished the healthy foXX XXXXXX XXXXXXXX XX XXX XXXX. *
9L 3R:	XXX XXn finished thx XXXXXXX XXXX XXXXXX XXXXXXXX XX XXX XXXX. *
9L 9R:	XXX XXn finished the healXXX XXXX XXXXXX XXXXXXXX XX XXX XXXX. *
9L 15R:	XXX XXn finished the healthy foXX XXXXXX XXXXXXXX XX XXX XXXX. *

Appendix 2. Model output

Words per Minute

Model parameter	Estimate	SE	t value	Pr(>|t|)
Intercept	**123.23**	**5.35**	**23.04**	**< .0001**
High frequency	**23.81**	**2.72**	**8.76**	**< .0001**
Right window 9-3	**32.45**	**2.92**	**11.13**	**< .0001**
Right window 15-9	1.09	1.40	.78	.44
Left window 6-3	**6.53**	**1.52**	**4.30**	**< .0001**
Left window 9-6	1.53	1.22	1.25	.21
English skill	1.92	.60	3.18	< .01[4]
High frequency: Right window 9-3	5.54	2.61	2.12	.04
High frequency: Right window 15-9	2.52	2.70	.93	.35
High frequency: Left window 6-3	3.71	2.36	1.57	.12
High frequency: Left window 9-6	-1.83	2.39	-.77	.44
Right window 9-3: English skill	**.93**	**.33**	**2.81**	**< .01**
Right window 15-9: English skill	**.38**	**.14**	**2.69**	**< .01**
Left window 6-3: English skill	.10	.18	.57	.57
Left window 9-6: English skill	.11	.14	.79	.43

Forward Saccade Length

Model parameter	Estimate	SE	t value	Pr(>|t|)
Intercept	**6.16**	**.13**	**45.80**	**< .0001**
High frequency	**.25**	**.05**	**4.72**	**< .0001**
Right window 9-3	**1.05**	**.06**	**17.31**	**< .0001**
Right window 15-9	**.21**	**.03**	**6.54**	**< .0001**
Left window 6-3	**.11**	**.03**	**3.44**	**< .01**
Left window 9-6	-.01	.03	-.40	.69
English skill	.04	.02	2.70	< .01[5]
High frequency: Right window 9-3	**.14**	**.05**	**2.71**	**< .01**
High frequency: Right window 15-9	.01	.05	.27	.79
High frequency: Left window 6-3	-.11	.05	-2.17	.03
High frequency: Left window 9-6	.08	.05	1.65	.10
Right window 9-3: English skill	**.02**	**.01**	**2.91**	**< .01**
Right window 15-9: English skill	.01	.00	2.20	.03
Left window 6-3: English skill	.00	.00	.30	.77
Left window 9-6: English skill	.00	.00	-.55	.59

Forward Fixation Duration

Model parameter	Estimate	SE	t value	Pr(>|t|)
Intercept	**5.46**	**.01**	**434.22**	**< .0001**
High frequency	**-.03**	**.00**	**-7.17**	**< .0001**
Right window 9-3	**-.13**	**.01**	**-14.20**	**< .0001**
Right window 15-9	**-.01**	**.00**	**-3.88**	**< .001**
Left window 6-3	**-.03**	**.00**	**-5.99**	**< .0001**
Left window 9-6	.00	.00	.90	.37
English skill	**.00**	**.00**	**-3.42**	**< .01**
High frequency: Right window 9-3	-.01	.01	-1.08	.28
High frequency: Right window 15-9	.00	.01	-.63	.53
High frequency: Left window 6-3	.01	.01	2.07	.04
High frequency: Left window 9-6	.00	.01	-.59	.56
Right window 9-3: English skill	.00	.00	-.98	.33
Right window 15-9: English skill	.00	.00	-.52	.60
Left window 6-3: English skill	.00	.00	.37	.71
Left window 9-6: English skill	.00	.00	-.86	.40

Regression count

Model parameter	Estimate	SE	t value	Pr(>|t|)
Intercept	**4.05**	**.20**	**19.91**	**< .0001**
High frequency	**-3.58**	**.11**	**-33.07**	**< .0001**
Right window 9-3	-.07	.10	-.72	.47
Right window 15-9	.11	.08	1.24	.22
Left window 6-3	-.14	.09	-1.55	.13
Left window 9-6	-.02	.08	-.30	.76
English skill	-.02	.02	-.98	.33
High frequency: Right window 9-3	.37	.18	2.10	.04
High frequency: Right window 15-9	.04	.17	.26	.80
High frequency: Left window 6-3	-.03	.16	-.15	.88
High frequency: Left window 9-6	.05	.15	.32	.75
Right window 9-3: English skill	.00	.01	.33	.75
Right window 15-9: English skill	.01	.01	.63	.53
Left window 6-3: English skill	.00	.01	-.21	.83
Left window 9-6: English skill	-.01	.01	-.82	.42

Appendix 3. Mean values (standard deviation) of dependent variables

	High frequency								
	3-Left			6-Left			9-left		
	3R	9R	15R	3R	9R	15R	3R	9R	15R
Word per minute	103.61 (45.75)	14.45 (61.59)	143.7 (65.13)	112.91 (47.47)	148.25 (66.9)	152.19 (7.2)	116.03 (51.78)	15.12 (7.34)	149.86 (68.85)
Forward saccade length (in characters)	5.37 (3.09)	6.43 (3.21)	6.61 (3.30)	5.40 (3.05)	6.51 (3.48)	6.55 (3.17)	5.42 (3.16)	6.43 (3.20)	6.79 (3.64)
Forward fixation duration (ms)	283.14 (142.7)	247.9 (118.14)	242.9 (12.86)	276.07 (135.14)	239.75 (112.61)	239.41 (112.87)	275.66 (127.92)	243.36 (116.1)	241.06 (116.36)
Regression count	2.61 (2.00)	2.84 (2.14)	2.87 (2.02)	2.52 (1.86)	2.78 (2.07)	2.74 (1.94)	2.51 (1.73)	2.61 (1.83)	2.84 (2.19)

	Low frequency								
	3-Left			6-Left			9-left		
	3R	9R	15R	3R	9R	15R	3R	9R	15R
Word per minute	86.05 (4.84)	117.98 (79.13)	116.5 (58.75)	92.69 (46.73)	121.66 (6.48)	119.94 (61.13)	94.70 (48.35)	122.5 (85.96)	124.18 (61.58)
Forward saccade length (in characters)	5.14 (3.01)	6.05 (3.16)	6.22 (3.36)	5.35 (3.15)	6.14 (3.20)	6.35 (3.45)	5.20 (2.98)	6.26 (3.31)	6.36 (3.45)
Forward fixation duration (ms)	29.67 (136.01)	257.83 (127.41)	256.42 (131.92)	281.06 (129.95)	248.78 (116.23)	248.12 (122.19)	283.50 (135.72)	249.48 (113.86)	248.69 (117.56)
Regression count	6.10 (4.94)	5.96 (4.82)	6.03 (5.20)	6.06 (4.95)	5.64 (4.34)	6.02 (4.46)	5.97 (4.71)	5.94 (4.86)	5.65 (4.15)

[1]High quality of lexical representation is knowing a word's orthographic, phonological, semantic, and syntactic qualities, which in turn leads to easier word identification during reading (Perfetti & Hart, 2001).
[2]Additionally, riskier reading strategies have not been found with older adults of non-alphabetic languages like Chinese (see for example, Wang et al., 2018).
[3]Previous studies run in our lab using the same proficiency test show that monolingually raised L1 speakers of English (n=97; 46 - L1 Canadian English/ 51- British English) scored 93.61% (sd: 5.54).
[4,5] These main effects are not interpreted given that they did not research significance in the omnibus testing.