Bilingualism: Language and Cognition

cambridge.org/bil

Skilled reading in a second language: Core reading processes, L2 proficiency and first-language background

Lin Chen[1], Xiaoping Fang[2] and Charles Perfetti[3]

[1]Department of Educational Psychology, Beckman Institute for Advanced Science and Technology, Department of Linguistics, University of Illinois Urbana-Champaign, USA; [2]School of Psychology and Cognitive Sciences, Beijing Language and Culture University, China and [3]Learning Research & Development Center, University of Pittsburgh, USA

Research Article

Cite this article: Chen, L., Fang, X. and Perfetti, C. (2026). Skilled reading in a second language: Core reading processes, L2 proficiency and first-language background. *Bilingualism: Language and Cognition* 1–13. https://doi.org/10.1017/S1366728926101898

Received: 1 October 2025 UTC
Revised: 10 July 2026 UTC
Accepted: 25 August 2026 UTC

Keywords:
L2 reading; proficiency; self-paced reading; authentic texts; surprisal

Corresponding author:
Lin Chen;
Email: linchen8@illinois.edu

This research article was awarded Open Data and Open Materials badges for transparent practices. See the Data Availability Statement for details.

Abstract

This study examined incremental processes in skilled second-language (L2) reading by identifying those shared with first-language (L1) reading and those modulated by L1 background and L2 proficiency. Skilled college-level L2 English readers (L1 Spanish, Korean or Chinese) were compared with native English speakers in self-paced reading of *New York Times* texts. Each word was annotated for word-, sentence- and passage-level predictors to test how these variables shape reading. Both L1 and L2 reading were affected by the same factors, indicating shared incremental processes. In L2, the magnitude of these effects was shaped by proficiency and readers' L1 background, with the most pronounced differences observed among Chinese–English readers. These findings support a skilled reading acquisition principle: Fluent reading relies on largely shared processes grounded in target-language knowledge and procedural skills for using this knowledge efficiently, even as proficiency and L1 background introduce variation in skilled L2 reading.

Highlights

- Skilled first-language (L1) and second-language (L2) reading is shaped by the same text factors.
- L2 reading variation is closely tied to proficiency.
- L1 background shows group-specific modulation.
- Skilled reading requires acquiring target-language knowledge and developing procedural skills.

1. Introduction

Reading is essential for acquiring knowledge and fully participating in society, making fluent reading the primary goal of school instruction. After years of literacy practice, a skilled native English reader can read about 238 words per minute (Brysbaert, 2019). However, this ability is not exclusive to native speakers; in today's multilingual and information-rich environments, many individuals must achieve advanced reading proficiency in a second language (L2), particularly when that language is critical for education, work, or daily communication. To promote fluent reading in L2, it is critical to understand the cognitive processes that support skilled L2 reading.

One long-standing debate in this area concerns whether skilled L2 readers rely on the same core processes as skilled first-language (L1) readers (Bley-Vroman, 2009; Clahsen & Felser, 2006; Gu et al., 2025; van Hell & Tokowicz, 2010). Some theoretical accounts propose that L2 reading relies on processing mechanisms that differ from those underlying L1 reading, although these accounts differ in the aspects of language processing to which they attribute these differences (see Hopp, 2022, for a review). One example is the shallow structure hypothesis, which argues that compared with L1, L2 syntactic processing is guided more strongly by semantic, pragmatic, probabilistic or surface-level information, resulting in shallower analyses of syntactically complex sentences (Clahsen & Felser, 2006, 2018). Other accounts propose that, as their language proficiency improves, L2 readers increasingly rely on processes similar to those of L1 readers (Brothers et al., 2021; Brysbaert et al., 2017). According to this view, many observed L1−L2 differences primarily reflect differences in proficiency or processing efficiency rather than fundamental differences in underlying processing mechanisms (Brysbaert et al., 2017). Under this view, once L2 readers achieve sufficient proficiency, their reading processes increasingly resemble those of L1 readers.

Beyond the question of L1–L2 differences, it is also important to understand the individual differences among skilled L2 readers. Examining such differences can help distinguish reading

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

processes that are shared across skilled L2 readers from those that remain sensitive to learner characteristics, such as language proficiency and L1 background (see Hopp, 2022, for a review). To address both questions, a useful framework is to consider how skilled reading develops more generally.

1.1. Skilled reading acquisition principle

Successful learning requires acquiring the knowledge components that define expertise within a domain (Koedinger et al., 2012). These components can take different forms, including knowledge of the domain and procedural skills for applying that knowledge. Building on this view, we propose the *skilled reading acquisition (SRA)* principle: Becoming a skilled reader, in any language, requires acquiring knowledge of the target language and developing procedural skills needed to use this knowledge efficiently during reading.

Language knowledge includes, for instance, knowledge of word forms, including their phonological and orthographic forms, knowledge of mappings between written and spoken language (e.g., orthography–phonology mappings), word meanings and the language's syntactic structure. Importantly, the specific language knowledge that must be acquired to support skilled reading differs across languages. For example, in English, learning to read requires acquiring phoneme–grapheme mappings (Castles et al., 2018; National Reading Panel, 2000; Rayner et al., 2001). In Chinese, readers must learn mappings between written characters and their associated syllables (morphemes; Verhoeven & Perfetti, 2017). Beyond these examples, other types of language knowledge, such as morphological, syntactic, discourse and pragmatic knowledge, also play important roles in reading and shape what learners need to prioritize across languages.

In addition to acquiring language knowledge, readers develop procedural skills in using this knowledge rapidly and efficiently to support fluent reading. For example, beginning readers may possess knowledge of orthography–phonology mappings but apply this knowledge slowly and effortfully when decoding written words. With practice, decoding skill develops as these mappings are applied increasingly rapidly and efficiently (Rayner et al., 2001). More generally, repeated practice allows acquired knowledge to be applied with increasing efficiency and automaticity.

According to the SRA principle, the same fundamental requirements for skilled reading apply to both L1 and L2 reading development. This is supported by prior findings that L2 learners must acquire orthography–phonology mappings specific to the target language, just as L1 learners do (Chen et al., 2025; Chen & Perfetti, 2024), while also developing efficiency in using this knowledge during word recognition (Droop & Verhoeven, 2003).

The SRA principle further has implications for the online processes of skilled reading. During online reading, acquired language knowledge and procedural skills support multiple moment-by-moment cognitive processes, including decoding, word identification, lexical access, syntactic parsing, semantic integration, inferencing, prediction and mental model updating (Chen et al., 2025; Perfetti & Stafura, 2014). Because the same fundamental requirements for skilled reading apply to both L1 and L2 reading development, online processes of skilled L1 and L2 reading are expected to be largely shared. As L2 language knowledge becomes better established and procedural skills become more efficient and automatic through practice, L2 reading processes should increasingly resemble those observed in skilled L1 reading.

Although these core processes are expected to be largely shared, the SRA principle also provides a basis for understanding variation in how they operate among skilled L2 readers. Such variation may arise from differences in readers' L2 proficiency and L1 background. L2 proficiency reflects the extent to which readers have acquired knowledge of the target language and developed the procedural skills needed to use that knowledge efficiently. Readers' L1 background, in turn, influences the extent to which language knowledge and procedural skills acquired in the L1 support or constrain L2 reading through cross-linguistic transfer, thereby shaping how shared processes operate during L2 reading. The following sections review evidence supporting these predictions.

1.2. L2 proficiency as a source of variation in skilled L2 reading processes

Consistent with the SRA principle, accumulating evidence identifies L2 proficiency as an important source of variation in skilled L2 reading. At the lexical level, higher-proficiency L2 readers show smaller word-frequency effects, suggesting more efficient word identification and lexical access (Mor & Prior, 2022). Beyond lexical processing, higher L2 proficiency is also associated with more efficient construction and updating of discourse representations. For example, when a text requires readers to revise their interpretation, higher-proficiency L2 readers show more inferential revision (Pérez et al., 2019). Proficiency likewise influences contextual processing, with higher-proficiency L2 readers showing stronger effects of word surprisal during meaning processing (Oralova et al., 2026).

Converging evidence from L1−L2 comparisons further supports the important role of L2 proficiency by demonstrating that higher L2 proficiency is associated with increasingly L1-like processing. Although L2 readers typically show larger word frequency effects than L1 readers (Cop et al., 2015), L1 and L2 readers show comparable word frequency effects when language proficiency is controlled (Brysbaert et al., 2017; Diependaele et al., 2013). Similarly, higher-proficiency L2 readers increasingly approximate L1-like syntactic processing (Brothers et al., 2021). At the neural level, event-related potential (ERP) evidence further indicates that high-proficiency L2 readers show neural response patterns that increasingly resemble those of native speakers (for a review, see van Hell, 2023).

1.3. L1 background as a source of variation in skilled L2 reading processes

Also consistent with the SRA principle, L2 readers' L1 backgrounds shape L2 reading through cross-linguistic transfer (MacWhinney, 2005), whereby language knowledge and procedural skills acquired in L1 facilitate or constrain reading in L2. The extent and direction of transfer depend on the degree of overlap between the L1 and L2 in their linguistic structure and writing systems. For example, because Chinese and English both make extensive use of compounding, Chinese–English speakers demonstrate compound-word awareness in English comparable with that of native English speakers. In contrast, because Chinese has relatively limited derivational morphology, these learners show weaker awareness of English derivational morphology (Ramirez et al., 2011). In addition to linguistic structure, similarity in L1–L2 writing systems also affects cross-linguistic transfer. Learning an L2 whose writing system shares print-to-speech mapping principles (e.g., mapping from writing units to phonemes, syllables or morphemes) with the learner's L1 is generally more advantageous than learning one with a different mapping system (Melby-Lervåg & Lervåg, 2014). This advantage extends beyond language

https://doi.org/10.1017/S1366728926101898 Published online by Cambridge University Press

knowledge to procedural skills, as reflected in stronger cross-language associations in word decoding when the two languages share the same writing system (Melby-Lervåg & Lervåg, 2011).

Although cross-linguistic transfer is often studied during L2 acquisition, evidence suggests that L1 influences persist even among skilled L2 readers. One manifestation of this persistent influence is that skilled L2 readers do not necessarily converge on the same level of reading performance. Nisbet et al. (2022) found that L1 background shapes the extent to which skilled L2 reading performance can approximate native levels. Using a statistical model to estimate the maximum reading performance achievable with native-like proficiency in key reading skills (i.e., spelling, vocabulary and exposure to print), they showed that when German–English speakers and Finnish–English speakers were matched on these skills, only the German readers, whose L1 shares greater structural similarity with English, could achieve reading speed comparable with native speakers.

Persistent L1 influences are evident not only in overall reading performance but also in online word recognition. Studies have shown that L1 lexical representations can remain co-activated during skilled L2 reading (Chen et al., 2021; Dijkstra & van Heuven, 2002). In one study, skilled Chinese–English bilinguals first silently read an English word for 300 ms and then performed an implicit color-judgment task on a Chinese character (Chen et al., 2021). Color judgments were faster when the character was the translation equivalent of the preceding English word than when the character was unrelated, suggesting rapid cross-language activation of shared semantic representations during L2 word recognition. Although these findings suggest that L1 influences persist even among skilled L2 readers and can affect online word recognition, how these influences extend beyond isolated word recognition to incremental reading processes during naturalistic text reading remains unclear.

1.4. Word-by-word incremental processes in reading authentic texts

When reading, individuals build a mental model of the text's meaning and incrementally update their representation as a word is processed (Ferreira & Chantavarin, 2018; Perfetti & Helder, 2021). These incremental processes are fundamental to building global comprehension (Chen et al., 2025). Much prior research on L2 reading has examined reading complexity by isolating specific components, such as lexical processing (Brysbaert et al., 2017; Diependaele et al., 2013) and syntactic parsing (Brothers et al., 2021; Román et al., 2022). Such studies typically implement factorial experimental designs in which measurements are restricted to a small number of target words. Although this approach produces informative data, it may not fully capture the continuous word-by-word processes that unfold across a text, as readers identify words, retrieve and integrate meanings and update comprehension. These processes are shaped by information at the word, sentence and passage levels, making it difficult to accommodate their combined influences within factorial designs that focus on a limited number of variables. Moreover, because both experimental and controlled variables are often defined as categorical, graded effects are difficult to detect. A further limitation is that the highly controlled experimental texts may diverge from the texts readers normally encounter. For example, some studies rely on sentences containing ambiguities or violations to probe specific processes (Gabriele et al., 2021; Grey, 2023), which may prompt reading strategies that differ somewhat from everyday reading.

These limitations highlight the need to examine L2 reading through word-by-word analyses of authentic texts to capture incremental processing during extended discourse. The Word-by-Word Processes Framework (WPF) views reading as a sequence of incremental processes that unfold continuously during naturalistic text comprehension (Chen et al., 2025). Within this framework, these processes are examined through readers' sensitivity to multiple levels of linguistic and quasi-linguistic text variables during word-by-word reading. Applying this framework, Oralova et al. (2026) examined incremental L2 reading processes using electroencephalography (EEG) during the reading of naturalistic texts. However, this work focused on readers from a single L1 background and a relatively narrow range of proficiency, leaving it unclear how incremental reading processes vary across readers with different L1 backgrounds and proficiency levels. The present study extends this line of work in two ways. First, it examines a larger and more diverse sample of skilled L2 readers from three different L1 backgrounds and a broader range of proficiency levels. Second, it applies the WPF within a self-paced reading paradigm, providing a complementary behavioral measure of incremental processing. Using this paradigm, we examine how multilevel text variables influence word-by-word reading processes during authentic text comprehension.

1.5. The present study

The present study examines reading processes in skilled L2 readers using the SRA principle as a theoretical framework. Skilled L2 reading is expected to rely largely on the same core processes as skilled L1 reading once sufficient language knowledge has been acquired and reading procedural skills have become efficient. Variation among skilled L2 readers is expected to arise primarily from differences in L2 proficiency and L1 background.

The present study addressed three questions: (1) Do skilled L2 readers rely on the same reading processes as skilled L1 readers? (2) How does L2 proficiency contribute to differences in reading processes among skilled L2 readers? and (3) How does L1 background contribute to differences in reading processes among skilled L2 readers?

Guided by the SRA principle, we tested three hypotheses. First, if skilled reading relies on the acquisition of target-language knowledge and the development of procedural skills for using this knowledge efficiently, skilled L2 readers should rely largely on the same core reading processes as skilled L1 readers and therefore show sensitivity to the same set of text variables observed in L1 readers in word-by-word reading times.

Second, among skilled L2 readers, L2 proficiency is expected to primarily modulate the effects of text variables on word-by-word reading times because it reflects the extent to which readers have acquired knowledge of the target language and developed the procedural skills needed to use this knowledge efficiently.

Third, among skilled L2 readers, L1 background is expected to influence how shared reading processes operate, thereby modulating the effects of text variables on word-by-word reading times. In the present study, L1 background is operationalized in terms of its structural similarity to English in writing system, lexical overlap and syntactic structure.

To test these hypotheses, we recruited three groups of skilled L2 readers with different L1 backgrounds (Spanish, Korean and Chinese) and varying levels of English proficiency, together with a comparison group of native English readers. These L1 backgrounds were selected to provide systematic contrasts in their structural similarity to English along the dimensions described previously.

Participants completed a self-paced reading task with *New York Times* articles, yielding reading times across each passage. Guided

https://doi.org/10.1017/S1366728926101898 Published online by Cambridge University Press

by the WPF (Chen et al., 2025), we examined how multilevel text variables, including both linguistic and quasi-linguistic factors, influence these reading times to investigate the incremental processes underlying skilled L2 reading.

2. Method

2.1. Rationale for L2 group selection

We included L2 English readers from three L1 backgrounds: Spanish, Korean and Chinese. The three L1 groups were selected to represent different degrees and types of structural similarity to English across three dimensions relevant to reading: writing system, lexical overlap and syntactic structure. (1) Writing system: Spanish and English share an alphabetic writing system with highly similar orthography–phonology mappings. Korean is also alphabetic but differs from English in script and visual layout – letters are arranged into square syllabic blocks rather than a linear sequence. Chinese, in contrast, has a morphosyllabic writing system in which each character corresponds to a syllable as well as a morpheme, resulting in minimal similarity to English in orthography–phonology mapping. (2) Lexical overlap: Spanish and English, both Indo-European languages, share many cognates that may facilitate lexical access during reading. In contrast, Korean and Chinese share fewer lexical overlaps with English, which may reduce semantic facilitation and slow meaning retrieval, although Korean contains more English-derived loanwords than Chinese. (3) Syntax: Spanish and English share the subject–verb–object (SVO) order, use verb tenses and require subject–verb agreement. Korean differs substantially, with a predominant subject–object–verb order, no subject–verb agreement and tense marked only on the final verb in a clause. Chinese shares the SVO order with English but lacks tense marking, instead using aspectual particles. These three groups provide systematic cross-linguistic contrasts for examining the influence of L1 background on L2 reading processes.

2.2. Participants

The sample consisted of three groups of L2 English college students and a comparison group of native English speakers (see Table 1 for participant characteristics). The L2 groups had Spanish, Korean and Chinese as their L1. Participants were recruited from the University of Pittsburgh or Carnegie Mellon University and provided informed consent before participation. One participant from each L2 group was excluded from the analyses because of both low scores on the standardized Nelson–Denny reading test and low accuracy on the self-paced reading task. The self-paced reading data from the native English speakers were previously reported in Chen et al. (2026a), which examined a different research question about the modulation of surprisal effects by text difficulty. In the present study, these data were reanalyzed as an L1 comparison for examining shared and differential reading processes between skilled L1 and L2 readers. The study procedure was approved by the Institutional Review Board at the University of Pittsburgh.

Table 1. Participant characteristics by language group

Group	*N* (*F*)	Age	Vocabulary test*	Comprehension test*
Spanish–English	44 (28)	19.18 (2.56)	61.65 (14.18)	70.61 (14.06)
Korean–English	29 (17)	19.69 (2.10)	60.89 (16.21)	71.73 (12.07)
Chinese–English	35 (16)	19.43 (1.40)	47.65 (16.73)	60.78 (15.81)
English native	35 (28)	18.51 (0.95)	55.66 (16.64)	62.06 (18.01)

Note: Total sample size is followed by the number of female (F) participants in parentheses. Means and standard deviations (in parentheses) are reported for age, vocabulary and comprehension tests. We initially recruited 49 Spanish–English speakers; five identified as nonnative Spanish speakers and were excluded. The table reports the final $n = 44$; *L2 readers were allocated twice as much time as L1 readers for both the vocabulary and comprehension tests (15 vs. 7.5 and 30 vs. 15 min, respectively). Therefore, scores are not directly comparable across L1 and L2 groups and are reported primarily to characterize proficiency within groups.

2.3. Reading proficiency measures

We assessed participants' L2 reading proficiency with reading-related standardized tests of English vocabulary and comprehension (Nelson & Denny, 1973). The comprehension test measured overall reading skill, whereas the vocabulary test specifically targeted word knowledge, a critical component of reading (Perfetti & Stafura, 2014) and a reliable predictor of comprehension performance, particularly among skilled readers (Landi, 2010). Accordingly, we administered the Nelson–Denny vocabulary and reading comprehension tests to both L1 and L2 readers. The vocabulary test consists of 100 multiple-choice items, and the comprehension section includes eight passages with a total of 36 multiple-choice questions.

Although originally developed for native English speakers, these tests have been validated as indicators for advanced L2 English readers and reliably capture meaningful individual differences in both L1 and L2 readers (Oralova et al., 2026; Yamasaki & Prat, 2014). For L1 readers, time limits were 7.5 min for vocabulary and 15 min for comprehension. For L2 readers, the limits were 15 and 30 min, respectively, to account for slower processing and better index individual variations in proficiency.

Participants' vocabulary and comprehension scores were calculated based on the percentage of correct answers in each test. Descriptive statistics are reported in Table 1. Because proficiency scores were intended primarily to characterize individual differences in English proficiency within the L2 groups, L2 readers were given extended time limits to reduce the influence of time constraints on their performance. Given these differences in test administration, proficiency scores should not be directly compared between L1 and L2 groups.

Among the L2 groups, Korean–English and Spanish–English speakers showed similar vocabulary and comprehension scores (Vocabulary: $\beta = -0.76$, $t = -0.20$, $p = .98$; Comprehension: $\beta = 1.12$, $t = 0.33$, $p = .94$). Chinese–English speakers showed lower scores in both vocabulary and comprehension tests compared with the other two groups (Chinese–English vs. Korean–English speakers: Vocabulary, $\beta = -13.25$, $t = -3.33$, $p < .01$; Comprehension, $\beta = -10.94$, $t = -3.03$, $p < .01$; Chinese–English vs. Spanish–English speakers: Vocabulary, $\beta = -14.00$, $t = -3.92$, $p < .001$; Comprehension, $\beta = -9.82$, $t = -3.02$, $p < .01$).

In addition to the standardized measures, we collected participants' self-rated English proficiency and information about their English language experience, including learning history, current language use and immersion experience in the United States, using the measures described in Chen et al. (2026b). Detailed descriptions of these measures and scoring procedures, together with descriptive results, are provided in Supplementary Material 3.

https://doi.org/10.1017/S1366728926101898 Published online by Cambridge University Press

2.4. Stimuli

To accommodate differences in English proficiency between L1 and L2 readers, two sets of authentic texts differing in overall difficulty were selected from the *New York Times* (published in 2018–2019). Each set contained 63 passages, with 20 passages shared across sets. The two sets were identical to those used in previous studies with L1 (Chen et al., 2025) and L2 readers (Oralova et al., 2026). All passages consisted of the first two sentences of selected articles, enabling an examination of reading processes both within and across sentence boundaries. Text difficulty was assessed using the Flesch–Kincaid Grade Level as implemented in Coh-Metrix (Graesser et al., 2011), with higher scores indicating greater difficulty. Flesh–Kincaid Grade Level scores differed significantly between the two sets, $t(124) = 2.47$, $p = .015$, Hedges' $g = 0.44$, 95% confidence interval (CI) [0.22, 2.03] (see Table 2 for descriptive statistics).

The texts spanned diverse topics without requiring specialized content knowledge. To verify that these topics were broadly accessible, we recruited 153 participants from the same population as the experiment sample (38 Spanish, 36 Korean, 29 Chinese and 50 native English speakers) to rate their familiarity with each topic on a 7-point scale (1 = *not at all familiar*, 7 = *very familiar*). Across all groups, mean familiarity ratings exceeded 5.5, confirming that the topics were generally accessible.

We also compared the linguistic characteristics of the two stimulus sets. Word frequency based on the SubtlexUS data set (Brysbaert & New, 2009), sentence length (number of words per sentence) and passage length (number of words per passage) differed significantly between the two sets: word frequency, $t(3087) = -2.88$, $p < .01$, Hedges' $g = -0.10$, 95% CI [−0.25, −0.05]; sentence length, $t(246) = 4.11$, $p < .001$, Hedges' $g = 0.52$, 95% CI [1.32, 3.73]; passage length, $t(122) = 5.12$, $p < .001$, Hedges' $g = 0.91$, 95% CI [3.10, 7.00]; in terms of lexical composition, the L1 set contained 1,727 words (1,116 open-class, 611 closed-class), and the L2 set contained 1,409 words (903 open-class, 506 closed-class). Two sample passages are shown below:

1) Polio is an infectious disease that can spread rapidly and mainly affects young children. It can cause muscle weakness, paralysis and, on rare occasions, can be fatal.
2) Tent camping should be one of the cheapest forms of travel, subject to nominal campsite and park admission fees. However, short of borrowing gear, bringing your own bedroll and shelter can be costly.

2.5. Procedure of self-paced reading

We implemented a self-paced moving-window paradigm using the experimental software PCIbex (Zehr & Schwarz, 2018). Each passage was initially displayed as a series of dashes representing each letter in a word. Pressing the spacebar revealed the first word; each subsequent press revealed the next word, and the previous word reverted to dashes. This procedure continued until all words in the passage were presented. Reading times were recorded for each word. Texts were displayed in black on a white background with a 20-point font. Punctuation marks (commas and periods) appeared together with the word to which they belonged.

Participants were instructed to read at their natural pace. To maintain engagement, a reminder appeared if no keypress occurred within 30 s. Two thirds of the passages were followed by a true/false comprehension question with immediate feedback. After each passage, participants could either take a short break or continue to the next one. The experiment began with six practice items and lasted less than 30 min in total.

Table 2. Text characteristics by set

Stimuli	Flesch–Kincaid Grade Level	Word frequency	Sentence length	Passage length
L1 set	8.19 (2.48)	4.13 (1.51)	13.71 (5.18)	27.41 (5.86)
L2 set	7.06 (2.64)	4.28 (1.39)	11.18 (4.54)	22.37 (5.18)

Note: Values are means with standard deviations in parentheses.

2.6. Text predictors

We included five linguistic and two quasi-linguistic variables to test how multilevel text factors influence reading times.

The linguistic variables included word frequency, word surprisal, word length, syntactic surprisal and phrase structure closure. Word frequency reflects how often a word occurs irrespective of the specific text context and thus serves as an index of context-independent lexical processing. Word surprisal quantifies the predictability of a word given its preceding context and serves as an index of context-dependent lexical processing. Word length (number of letters) was included as a covariate to account for potential effects of visual and form-based processing cost during word recognition. Although the word age of acquisition may explain variance beyond word frequency and length, it was not included as a predictor because values were available for only approximately 76% of the words in the materials based on the most recent large-scale norms (Kuperman et al., 2012).

To capture syntactic processing, we included syntactic surprisal and phrase structure closure. Syntactic surprisal reflects how unexpected a word's syntactic category is given the categories of preceding words, capturing sensitivity to the probabilistic distribution of syntactic categories. Phrase structure closure refers to the number of phrase structure nodes a word closes, indexing structural updating and meaning integration as syntactic structures are completed.

Quasi-linguistic factors included word position (within a sentence) and sentence position (within the text). These variables were considered quasi-linguistic to distinguish them from strictly linguistic factors (e.g., lexical and syntactic) while capturing the accumulation of contextual information as readers progress through a sentence or passage.

2.6.1. Word frequency

Word frequency affects reading time. High-frequency words develop more robust mental representations and allow faster lexical access, resulting in shorter reading times than low-frequency words (Ratcliff et al., 2004; Rayner & Duffy, 1986). In this study, we used log-transformed word frequency from the SUBTLEX corpus (Brysbaert & New, 2009), given that prior work has shown that the relationship between word frequency and reading times is approximately logarithmic (White et al., 2018).

2.6.2. Word surprisal and syntactic surprisal

We used surprisal derived from probabilistic language models to quantify the predictability of each word and to examine its effects on reading. Surprisal quantifies how unexpected a word is given its preceding contexts; it is defined as the negative logarithm of this conditional probability of the word given the preceding words (Hale, 2001; Levy, 2008). A higher surprisal value indicates lower probability. For example, in the context, *she has a pet*, a predictable continuation that includes *dog* would have low surprisal, whereas a

https://doi.org/10.1017/S1366728926101898 Published online by Cambridge University Press

less typical continuation that includes *duck* would have high surprisal. Prior work shows that surprisal is associated with cognitive effort during reading: Words with higher surprisal are read more slowly than lower-surprisal words (Chen et al., 2026a; Demberg & Keller, 2008; Roark et al., 2009) and elicit a more negative N400 response in EEG studies, indicating greater difficulty in meaning processes (Chen et al., 2025; Frank et al., 2015; Oralova et al., 2026).

In this study, we include two types of surprisal. Word surprisal reflects the probability of a word and therefore includes both lexical and syntactic information. Syntactic surprisal, in contrast, reflects the probability of a word's syntactic category, isolating the contribution of syntactic information (Demberg & Keller, 2008; Roark et al., 2009). For example, after a determiner such as *the*, a noun is highly expected (low syntactic surprisal), whereas an adverb in this position would be less expected and therefore associated with higher syntactic surprisal. Including these measures allows us to distinguish syntactic from lexical contributions to incremental processing during reading (Chen et al., 2025). Both types of surprisal were estimated using transformer-based language models (see Section 3.7 for details).

2.6.3. Phrase structure closure

We used the number of phrases (syntactic nodes) closed by a word, also known as the syntactic node count (Frazier, 1985; Miller & Chomsky, 1963), to indicate syntactic structure updating. Prior research has shown that phrase structure closure effects are evident in both neuroimaging and neurophysiological measures (Brennan et al., 2012; Chen et al., 2025). For example, Chen et al. (2025) found that increases in the number of syntactic structure closures were associated with reduced N400 negativity, suggesting that phrase structure closure facilitates meaning integration.

We counted phrase structure closures following Chen et al. (2025). First, syntactic trees for each sentence were generated by Stanford Parser (Klein & Manning, 2003) and manually reviewed and corrected before proceeding with the node counting. Figure 1 illustrates the tree for an example sentence, and Table S1 in the Supplementary Materials reports the number of syntactic nodes closed by each word. When a word closed only its own open syntactic category node, it was assigned a value of 1. To capture phrase-level closure, we counted nodes associated with the completion of the local phrase structure rather than those associated with the full syntactic tree. For example, for the word "omnivorous," nodes were counted up to the verb phrase rather than through the full sentence structure. As a robustness check, we also computed full-tree closure (nodes closed in the entire tree). For "omnivorous," this count was 4. Both measures showed similar patterns of results.

2.6.4. Word position

Word position captures the ordinal location of a word within a sentence, its effect reflecting how processing changes as contextual information accumulates. At the beginning of a sentence, the reader's primary task is to build a new structure, whether it is a sentence, clause or phrase (Calloway & Perfetti, 2017; Perfetti & Helder, 2021), typically resulting in longer reading times (Kuperman et al., 2010) and more negative N400 amplitudes (Chen et al., 2025; Payne et al., 2015), with reading times decreasing and N400 amplitudes becoming less negative as the sentence unfolds. We operationalized word position as the cardinal order of a word within its sentence.

2.6.5. Sentence position

Sentence position captures the ordinal location within a passage of the sentence in which a word occurs. The first sentence benefits least from prior context, whereas subsequent sentences are supported by preceding discourse, leading to less negative N400 amplitudes (Chen et al., 2025). Word position and sentence position also interact, so that word position effects are more robust in the first sentence (Chen et al., 2025). In this study, sentence position was treated as a categorical predictor distinguishing the first and second sentences.

2.7. Surprisal and language model training

Word and syntactic surprisal estimates were obtained from Chen et al. (2026a), using 2 two-layer transformer models: a lexicalized model for calculating word surprisal and an unlexicalized model for calculating syntactic surprisal. In the unlexicalized model, lexical items were replaced by part-of-speech (POS) tags generated using the Stanford Parser (Version 4.2.0; Klein & Manning, 2003) based on the Penn Treebank POS tagset. Following Merkx and Frank (2021), these estimates were derived from relatively simple transformer models trained on the WikiText-2 corpus (Merity et al., 2016) rather than highly complex pretrained large language models such as Generative Pre-trained Transformer (GPT). Prior research suggests that simpler models can sometimes outperform more complex ones in predicting reading times (Oh et al., 2022), likely because large models possess extensive memory capacities that exceed human memory constraints. Moreover, large-scale pretrained models are typically trained on noisy web data that may include linguistic errors. In contrast, the Wikipedia-derived WikiText-2 corpus provides well-structured content whose linguistic and stylistic features align more closely with those of our experimental materials from *the New York Times*. Such similarity is important for generating accurate surprisal estimates (Hale et al., 2019).

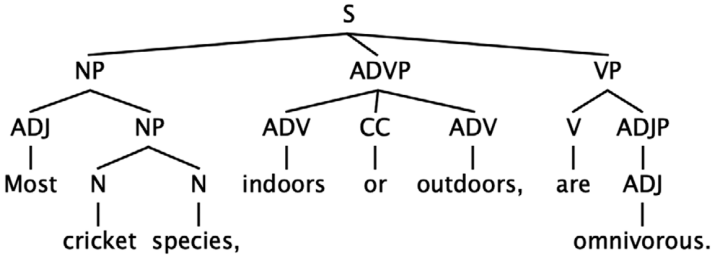


Figure 1. The syntactic tree for an example sentence in the stimuli. S = sentence; NP = noun phrase; VP = verb phrase; ADJP = adjective phrase; ADVP = adverb phrase; N = noun; V = verb; ADJ = adjective; ADV = adverb; CC = coordinating conjunction.

https://doi.org/10.1017/S1366728926101898 Published online by Cambridge University Press

Each model included a 400-dimensional embedding layer, two transformer layers (eight attention heads each) and a fully connected layer with 1,024 units. The corpus contains 720 articles (600 for training, 60 for validation and 60 for testing). After parameter tuning, we selected a learning rate of 0.005, a dropout rate of 0.1 and a batch size of 5. The resulting surprisal values for the L1 and L2 text sets are summarized in Table 3. Neither word nor syntactic surprisal differed between the two sets – word surprisal: $t(3071) = 1.64$, $p = .10$, Hedges' $g = 0.06$, 95% CI [−0.04, 0.50]; syntactic surprisal: $t(3020) = -0.20$, $p = .84$, Hedges' $g = -0.01$, 95% CI [−0.11, 0.09].

2.8. Data analysis

2.8.1. Data preprocessing

Reading times were preprocessed before statistical analysis. Following prior work, observations longer than 3,000 ms or shorter than 100 ms (Wehbe et al., 2021), as well as those more than 2.5 SD above a participant's average reading time (Tremblay et al., 2011), were excluded. Data from passages with incorrect comprehension responses were also removed.

2.8.2. Analytic strategy overview

We conducted two complementary sets of analyses to address our three research questions. First, because L1 and L2 readers were tested on different sets of stimuli tailored to their language proficiency, we analyzed each group separately to identify the pattern of effects of text variables within each group of readers to examine whether similar processing patterns emerge in skilled L1 and L2 readers. Second, we combined the three L2 groups in a cross-group analysis to evaluate the joint contributions of L2 proficiency and L1 background to individual differences in L2 reading.

2.8.3. Group-specific models

For each group, linear mixed-effects models (LMMs) were fitted, including all relevant text predictors (linguistic and quasi-linguistic variables). Linguistic variables of the immediately preceding word were also included to capture potential spillover effects on the current word's processing (Kliegl et al., 2006). An illustration of the model is provided in the Supplementary Materials. All predictors were standardized to facilitate direct comparisons of effect sizes and reduce nonessential multicollinearity (Payne et al., 2015). To facilitate interpretation of interactions, the categorical variable Sentence Position was effect-coded (−0.5 for the first sentence, 0.5 for the second). Variance inflation factors indicated no significant multicollinearity (VIF < 10; Kutner et al., 2004). Following prior studies (Frank et al., 2015; Roark et al., 2009), the first and last words of each sentence were excluded to minimize disproportionate influences of "start-up effects" and "wrap-up effects." Analyses were further restricted to open-class words (nouns, verbs, adjectives and adverbs) to minimize potential confounds from high-frequency closed-class words (Kuperman et al., 2010; Payne et al., 2015). Full model outputs are reported in Tables S2–S5 in the Supplementary Materials.

2.8.4. Combined L2 model

To examine how reading proficiency and L1 background shape L2 reading, we fit a single LMM model to the combined three L2 groups. Proficiency was indexed by standardized vocabulary and comprehension scores, which were entered as main effects and interacted with each text-level predictor. L1 background (Spanish, Korean or Chinese) was included as a categorical factor, also interacting with each text-level predictor to test modulation by language background. Contrast coding was used to compare (1) Chinese–
English readers with the combined group of Spanish– and Korean–English readers and (2) Spanish–English readers with Korean–English readers, reflecting differences between morphosyllabic and alphabetic L1s as well as differences within alphabetic L1s. To interpret interactions between the categorical variable L1 background and continuous text predictors, we used estimated marginal trends implemented in the *emmeans* package (Lenth & Piaskowski, 2025). For each interaction, we estimated the slope of the continuous predictor separately for each L1 group and then compared these slopes across groups using pairwise contrasts with Tukey's correction for multiple comparisons.

A full specification of the model is provided in the Supplementary Materials. All data and analysis code are openly available at https://doi.org/10.17605/OSF.IO/WQM8U.

Table 3. Surprisal values from transformer models

	L1 set			L2 set		
Surprisal	Range	Mean	SD	Range	Mean	SD
Word surprisal	0.005–16.64	6.99	4.01	0.003–16.58	6.77	3.78
Syntactic surprisal	0.007–9.50	2.25	1.49	0.007–8.75	2.26	1.47

3. Results

3.1. Comprehension accuracy and reading times in the self-paced reading task

Comprehension accuracy was assessed through participants' responses to true/false questions following the passages. Reading times were measured for each word across all passages. The mean comprehension accuracy and reading times are reported in Table 4.

All three groups of L2 readers showed high comprehension accuracy that was comparable with that of L1 readers. A one-way analysis of variance (ANOVA) across the L1 and three L2 groups showed no significant differences in accuracy: $F(3, 136) = 2.26$, $p = .085$, $\eta^2 = 0.05$. Post hoc comparisons also revealed no significant pairwise differences (all ps > .10).

After the reading time data preprocessing and cleaning procedures described in Section 2.8.1, 9.50%, 10.13%, 11.07% and 12.13% of the data were excluded from the native English, Spanish–English, Korean–English and Chinese–English groups, respectively. Skilled L2 readers showed comparable word-level reading times to native speakers: $F(3, 136) = 1.81$, $p = .15$, $\eta^2 = 0.04$, and no significant differences emerged in post hoc tests (all ps > .10).

Table 4. Means and standard deviations of comprehension accuracy and reading times

	Comprehension accuracy	Reading times (ms)
Spanish–English	0.93 (0.05)	367 (175)
Korean–English	0.91 (0.06)	337 (158)
Chinese–English	0.90 (0.05)	395 (200)
English native	0.93 (0.06)	348 (145)

Note: Means are followed by standard deviations in parentheses.

https://doi.org/10.1017/S1366728926101898 Published online by Cambridge University Press

3.2. Effects of text variables within each reader group

Because the L1 and L2 groups were tested on different stimulus sets, we did not directly compare the groups statistically. Instead, to address the first question of whether skilled L2 readers rely on the same core processes as skilled L1 readers, we examined whether the effects of text predictors were consistent within each group and interpreted convergence across groups as evidence of shared reading processes. Overall, similar patterns were observed in both L1 and L2 readers, suggesting that skilled L2 readers rely on reading processes comparable with those of skilled L1 readers. In the following text, we discuss the effects of each predictor.

3.2.1. Word frequency effect

Word frequency affected reading times across all groups through a spillover effect: Lower-frequency words led to longer reading times on the following word (English native speakers: $\beta = -9.04$, $t = -6.45$, $p < .001$; Spanish–English speakers: $\beta = -8.49$, $t = -4.74$, $p < .001$; Korean–English speakers: $\beta = -5.66$, $t = -3.02$, $p < .01$; Chinese–English speakers: $\beta = -11.88$, $t = -4.77$, $p < .001$). Only Chinese–English speakers showed a significant effect of word frequency on current-word reading times ($\beta = -9.31$, $t = -2.95$, $p < .01$), whereas no such effect was observed in the other three groups.

3.2.2. Word surprisal effect

Surprisal also influenced reading times through a spillover effect across all groups, with no effect on the current word: Higher surprisal increased reading times for the following word (English native speakers: $\beta = 5.02$, $t = 3.59$, $p < .001$; Spanish–English speakers: $\beta = 5.20$, $t = 2.99$, $p < .01$; Korean–English speakers: $\beta = 5.47$, $t = 3.01$, $p < .01$; Chinese–English speakers: $\beta = 9.91$, $t = 4.08$, $p < .001$).

3.2.3. Word length effect

All four groups showed significant effects of word length: Words with more letters were associated with longer reading times (English native speakers: $\beta = 7.90$, $t = 6.30$, $p < .001$; Spanish–English speakers: $\beta = 8.48$, $t = 5.59$, $p < .001$; Korean–English speakers: $\beta = 11.12$, $t = 7.01$, $p < .001$; Chinese–English speakers: $\beta = 23.91$, $t = 11.33$, $p < .001$).

3.2.4. Syntactic surprisal effect

No significant effects of syntactic surprisal were observed in any group (English native speakers: $\beta = -0.28$, $t = -0.31$, $p = .75$; Spanish–English speakers: $\beta = 0.23$, $t = 0.23$, $p = .82$; Korean–English speakers: $\beta = 0.40$, $t = 0.37$, $p = .71$; Chinese–English speakers: $\beta = -0.97$, $t = -0.69$, $p = .49$).

3.2.5. Phrase structure closure effects

Phrase structure closure did not show significant effects on reading times in any groups (English native speakers: $\beta = -1.47$, $t = -1.30$, $p = .19$; Spanish–English speakers: $\beta = -0.21$, $t = -0.15$, $p = .88$; Korean–English speakers: $\beta = -2.35$, $t = -1.60$, $p = .11$; Chinese–English speakers: $\beta = -1.36$, $t = -0.70$, $p = .49$).

3.2.6. Word position effect

Word position significantly affected reading times in all groups: Words near the beginning of a sentence were read slowly (English native speakers: $\beta = -2.41$, $t = -2.75$, $p < .01$; Spanish–English speakers: $\beta = -2.78$, $t = -2.41$, $p = .016$; Korean–English speakers: $\beta = -5.06$, $t = -4.16$, $p < .001$; Chinese–English speakers $\beta = -13.67$, $t = -8.47$, $p < .001$).

3.2.7. Sentence position effect

Reading times were longer for the first than the second sentence for English native speakers ($\beta = -5.00$, $t = -3.09$, $p < .01$), Korean–English speakers ($\beta = -8.08$, $t = -3.74$, $p < .001$) and Chinese–English speakers ($\beta = -21.28$, $t = -7.41$, $p < .001$). Spanish–English readers showed a nonsignificant trend in the same direction ($\beta = -3.39$, $t = -1.66$, $p = .097$).

3.2.8. Interactions between word position and sentence position

All groups showed significant interactions between word position and sentence position, with stronger word position effects in the first sentence (English native speakers: $\beta = 7.28$, $t = 4.00$, $p < .001$; Spanish–English speakers: $\beta = 9.44$, $t = 4.04$, $p < .001$; Korean–English speakers: $\beta = 7.44$, $t = 3.00$, $p < .01$; Chinese–English speakers: $\beta = 11.34$, $t = 3.45$, $p < .001$).

No other interactions were found, except among Chinese–English readers, who showed an interaction between word frequency and word surprisal ($\beta = -6.04$, $t = -3.55$, $p < .001$). This interaction may reflect their stronger word frequency effect, a pattern further confirmed in the cross-group comparison presented in the following paragraph.

3.3. Effects of L2 proficiency on reading times

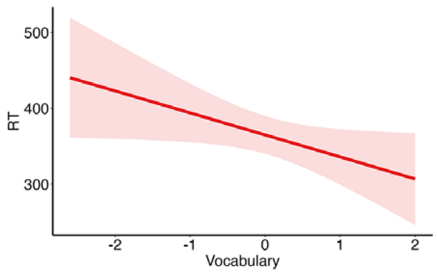
We fit a single LMM that included all three groups of L2 readers to address the second question of how L2 proficiency influences reading times. Vocabulary showed a main effect on reading times: Readers with larger vocabularies read more quickly ($\beta = -29.81$, $t = -2.15$, $p = .034$; Figure 2A). In addition, vocabulary interacted with word frequency: Readers with smaller vocabularies showed a larger word frequency effect than readers with larger vocabularies ($\beta = 2.25$, $t = 2.10$, $p = .036$; Figure 2B). Comprehension proficiency did not show a main effect on reading times ($\beta = 15.67$, $t = 1.17$, $p = .27$). However, it interacted with word frequency, such that the word frequency effect decreased as comprehension proficiency increased ($\beta = 2.53$, $t = 2.41$, $p = .016$; Figure 2C). No other interactions involving vocabulary or comprehension were significant.

3.4. Effects of L1 background on reading times

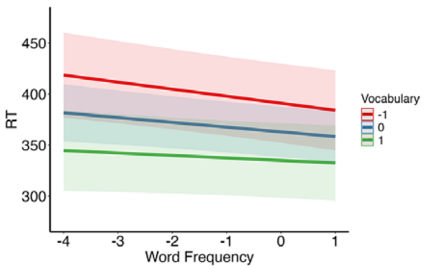
Using the same LMM, we examined the third question of whether L1 background influenced reading times. L1 background interacted with word frequency, word position and sentence position to affect reading times. Pairwise comparisons showed that Chinese–English speakers relative to the other two groups showed stronger effects of word frequency (Chinese L1 vs. Korean L1: $\beta = -20.22$, $t = -7.84$, $p < .001$; Chinese L1 vs. Spanish L1: $\beta = -17.09$, $t = -7.25$, $p < .001$; Figure 3A), word position (Chinese L1 vs. Korean L1: $\beta = -9.97$, $t = -6.25$, $p < .001$; Chinese L1 vs. Spanish L1: $\beta = -11.88$, $t = -8.31$, $p < .001$; Figure 3B) and sentence position (Chinese L1 vs. Korean L1: $\beta = -13.15$, $t = -4.70$, $p < .001$; Chinese L1 vs. Spanish L1: $\beta = -16.07$, $t = -6.40$, $p < .001$; Figure 3C). By contrast, Korean–English and Spanish–English speakers did not differ significantly from each other in the magnitude of these effects (word frequency: $\beta = 3.13$, $t = 1.37$, $p = .35$; word position: $\beta = -1.91$, $t = -1.26$, $p = .42$; sentence position: $\beta = -2.92$, $t = -1.10$, $p = .27$).

Because the Chinese–English speakers in this study had relatively lower English proficiency than the other two L2 groups (see Section 3.3), the observed L1 background effect may partly reflect a

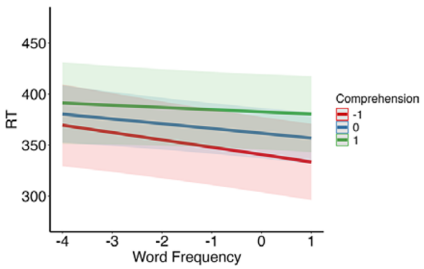
https://doi.org/10.1017/S1366728926101898 Published online by Cambridge University Press



a. The effect of vocabulary tests on reading times.

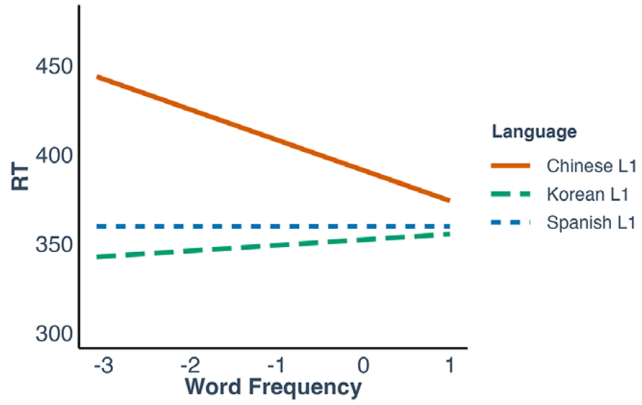


b. Interaction between vocabulary size and word frequency.

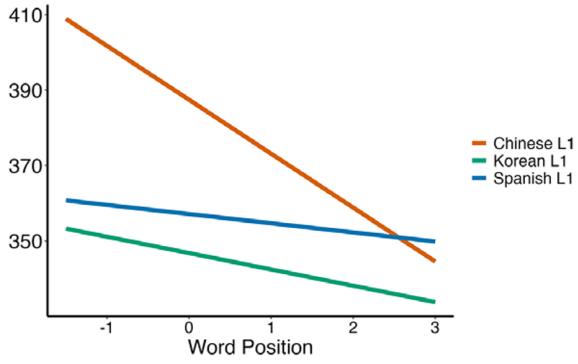


c. Interaction between comprehension tests and word frequency.

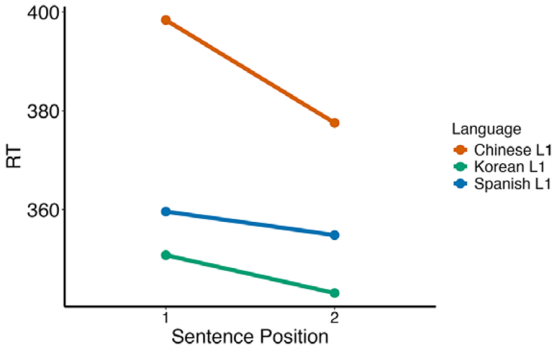
Figure 2. Reading proficiency and reading times.



a. Interaction between L1 background and word frequency.



b. Word and sentence position effects by L1 groups.



c.

Figure 3. Interactions between L1 background and text factors.

Table 5. Means and standard deviations of the accuracy of vocabulary and comprehension tests of matched subgroups

	Vocabulary	Comprehension
Spanish–English	59.72 (11.91)	71.09 (9.20)
Korean–English	58.90 (11.79)	71.69 (10.85)
Chinese–English	58.78 (10.40)	66.82 (8.55)

Note: Means are followed by standard deviations in parentheses.

proficiency imbalance, despite proficiency being included in the model. To address this concern, we formed three subgroups of participants whose vocabulary and comprehension scores were restricted to overlapping ranges (Vocabulary: 40–80; Comprehension: 50–88; see Table 5). This procedure yielded 18 Chinese–English speakers, 21 Korean–English speakers and 32 Spanish–English speakers with no significant group different in vocabulary, $F(2, 68) = 0.051$, $p = .95$, $\eta^2 = 0.0015$, or comprehension, $F(2, 68) = 1.52$, $p = .23$, $\eta^2 = 0.04$. Crucially, even after matching, the Chinese–English group continued to differ from the other two groups in word frequency (Chinese L1 vs. Korean L1: $\beta = -18.21$, $t = -6.29$, $p < .001$; Chinese L1 vs. Spanish L1: $\beta = -11.77$, $t = -4.45$, $p < .001$; Figure S1a in Supplementary Material 1), word position (Chinese L1 vs. Korean L1: $\beta = -11.36$, $t = -6.03$, $p < .001$; Chinese L1 vs. Spanish L1: $\beta = -11.75$, $t = -6.83$, $p < .001$; Figure S1b in the Supplementary Materials) and sentence position (Chinese L1 vs. Korean L1: $\beta = -11.90$, $t = -3.61$, $p < .001$; Chinese L1 vs. Spanish L1: $\beta = -15.00$, $t = -4.96$, $p < .001$; Figure S1b in Supplementary Material 1) effects. Overall, these results indicate that the Chinese–English group showed stronger sensitivity to word frequency, word position and sentence position than the other two groups, whereas Korean–English and Spanish–English readers showed largely similar patterns.

4. Discussion

This study investigated the word-by-word processes underlying skilled L2 reading of authentic texts drawn from the *New York Times*. Using self-paced reading, we examined how multiple text factors influence processing across entire passages. Three groups of skilled L2 English readers with different L1 backgrounds (Spanish, Korean and Chinese) were compared with a group of native English readers. The results show that skilled L2 reading, regardless of readers' L1 background, was influenced by the same linguistic

https://doi.org/10.1017/S1366728926101898 Published online by Cambridge University Press

(lexical and syntactic) and quasi-linguistic (word and sentence position) factors that shape L1 reading, indicating shared underlying processes. In addition, the magnitude of these effects in L2 reading varied with readers' proficiency level, with additional influences of L1 background, most notably for Chinese–English speakers. These findings support the SRA principle: Skilled L2 reading relies largely on shared processes with skilled L1 reading that emerge through the acquisition of knowledge of the target language and the development of efficient reading procedural skills. At the same time, proficiency and L1 background shape variation in how these processes operate in skilled L2 reading.

4.1. Shared processes in skilled L1 and L2 reading

Across reader groups, several text variables, including lexical, syntactic and positional factors, showed broadly similar patterns of influence on L1 and L2 reading times.

We included word frequency and word surprisal to test how context-independent and context-dependent information influences word processing, respectively. As expected, both variables affected reading times in L1 and L2 readers, indicating that skilled reading is shaped by both word-level properties and the context in which a word occurs. Word frequency reflects readers' cumulative exposure to words. Higher-frequency words support faster and more automatic retrieval of meanings because repeated encounters strengthen lexical quality (Perfetti, 2007; Perfetti & Stafura, 2014). In contrast, word surprisal captures the use of contextual information to facilitate processing. Lower surprisal indicates that a word is more predictable given its preceding context and therefore requires less processing effort.

Both word frequency and word surprisal influence self-paced reading primarily through spillover effects on the following word. Higher frequency at position n was associated with shorter reading times at $n + 1$, with a marginal effect on the current word n ($p = .052$). Higher word surprisal at position n led to longer reading times at $n + 1$, with no reliable effect on the current word. Of these effects, the spillover effect of word surprisal is particularly noteworthy, which has also been robustly observed in prior self-paced reading work (Chen et al., 2026a; Fernández Monsalve et al., 2012). These robust findings support that the processing demands generated by unpredictable words may extend to subsequent words, consistent with the incremental nature of reading.

The spillover effect may partly arise from the self-paced reading paradigm, in which participants can develop a buffering strategy by advancing quickly to the next word and postponing full semantic integration until more words are read. Such a strategy may reduce effects on the current word while increasing spillover effects on the following word. Consistent with this interpretation, ERP studies using the same materials, in which word presentation was experimenter-paced, showed robust surprisal effects on the current word (Chen et al., 2025; Oralova et al., 2026). These contrasting patterns across the self-paced reading and ERPs suggest that task characteristics can influence whether surprisal-related processing costs are more readily observed at the current word or as spillover on the subsequent word.

To examine different aspects of syntactic processing, we included two complementary measures: syntactic surprisal (the predictability of a syntactic category) and phrase structure closure (the number of syntactic constituents completed at each word). In the present self-paced reading data, neither measure reliably predicted reading times in L1 or L2 readers. Several factors may have contributed to these relatively weak behavioral effects. First, the magnitude of syntactic effects may be shaped by the characteristics of the texts. This study used authentic texts without syntactic violations, such that syntactic variation occurred continuously across words rather than being induced through categorical manipulations. These graded differences may produce subtler effects than explicit syntactic violations (Mifka-Profozic et al., 2020; Traxler, 2005). Importantly, this does not imply that syntactic effects cannot be detected during the reading of authentic texts. Syntactic surprisal effects have been reported in previous studies using naturalistic texts (Demberg & Keller, 2008; Roark et al., 2009). However, in such texts, the magnitude of these effects also depends on the text difficulty. For example, syntactic surprisal effects have been found to be stronger in more difficult texts (Chen et al., 2026a). In addition, self-paced reading times may be less sensitive than neural measures to subtle variations in syntactic processing. ERP studies using the same materials show reliable neural responses associated with both syntactic surprisal and phrase structure closure in both L1 and L2 readers (Chen et al., 2025, 2026b; Oralova et al., 2026), suggesting that graded variations in syntactic processing may be detectable at the neural level even when they are not reliably detected in self-paced reading times.

We also found that word position and sentence position affected both L1 and L2 reading, with reading times decreasing as readers progressed within and across sentences, consistent with prior work (Kuperman et al., 2010). Word position and sentence position also interacted; the word position effect was stronger in the first sentence than in the second. Although positional variables do not directly encode linguistic content, they can index the accumulation of contextual information across a sentence and a passage. The observed position effects suggest that both L1 and L2 readers use accumulated contextual information to facilitate subsequent processing.

In sum, L2 reading was shaped by the same text factors as L1 readers, suggesting that skilled L2 reading relies on incremental processes similar to those engaged in L1 reading. This pattern is consistent with the SRA principle, according to which skilled reading develops as learners acquire the essential knowledge and efficient procedural skills of the target language. As these knowledge and procedural skills develop, L2 reading increasingly engages the core processes that support skilled L1 reading. These findings therefore contribute to the long-standing debate over whether L1 and L2 reading rely on distinct or shared cognitive processes, providing evidence for shared core processes in skilled L1 and L2 reading.

4.2. Proficiency modulates incremental processing in L2 reading

Among a sample of skilled college-level L2 readers, variation in L2 reading proficiency influenced reading processes. Readers with higher vocabulary proficiency read more quickly overall, indicating greater lexical processing efficiency. More importantly, both vocabulary and comprehension proficiency modulated the word frequency effect, with lower-proficiency readers showing stronger frequency effects than higher-proficiency readers.

These proficiency effects align with the SRA principle, which proposes that becoming a skilled reader requires acquiring knowledge of the target language and developing procedural skills to use this knowledge efficiently during reading. Although participants in the present study were skilled L2 readers, variation in proficiency may reflect continued development in both language knowledge and procedural skills. The faster overall reading observed among

https://doi.org/10.1017/S1366728926101898 Published online by Cambridge University Press

higher-proficiency readers is consistent with better developed language knowledge and greater efficiency in using this developing knowledge during online reading.

The modulation of the word frequency effect by proficiency provides more specific evidence about the lexical knowledge underlying this increased processing efficiency. At the lexical level, language knowledge is encoded in representations of individual words. According to the lexical quality hypothesis (Perfetti, 2007), efficient word processing depends on high-quality lexical representations, particularly how well knowledge about a word's form and meaning is specified and bound. With increased L2 proficiency and experience, lexical representations, particularly those of less frequent words, may become more stable and readily retrievable. Thus, higher-proficiency readers may experience less of a processing disadvantage for low-frequency words, resulting in smaller word frequency effects. In contrast, relatively lower-proficiency readers may have less robust representations for lower-frequency words, resulting in larger frequency effects.

4.3. Influences of L1 background

L1 background showed group-specific influences on skilled L2 reading, with the effect primarily driven by the Chinese–English group. Chinese–English speakers showed more pronounced effects of word frequency, word position and sentence position than their Korean- and Spanish–English counterparts. These patterns persisted even in the subgroup analysis where language proficiency was matched. These findings suggest that L1 background continues to influence online reading processes even among skilled L2 readers.

The stronger word frequency effect among Chinese–English speakers may partly reflect differences in L1 writing-system experience. Chinese uses a morphosyllabic writing system in which each character corresponds to a syllable and a morpheme. Experience with such a system may encourage greater reliance on familiarity-based lexical-level orthographic processing during English word identification (Ben-Yehudah et al., 2019; Wang et al., 2003), thereby increasing sensitivity to word frequency during L2 English reading. By contrast, readers whose L1 writing systems provide more direct mapping between orthographic units and phonology, such as Korean and Spanish, may be better able to transfer sublexical phonological decoding skills to English (Ben-Yehudah et al., 2019; Wang et al., 2003), reducing reliance on stored lexical representations and leading to smaller word frequency effects.

In addition to writing-system experience, cross-linguistic lexical overlap may also contribute to these group differences. Spanish and English share many cognates, and Korean contains more English-derived loanwords than Chinese. Such lexical overlap may increase incidental exposure to English words and facilitate lexical access, even when vocabulary size is matched across groups. In contrast, because Chinese shares less lexical overlap with English, Chinese–English readers may have fewer opportunities for such cross-linguistic reinforcement, particularly for lower-frequency words, potentially contributing to the larger word frequency effects observed in this group.

The source of the stronger word- and sentence-position effects among Chinese–English readers is less clear. It is possible that these text-level differences are related to the greater lexical processing demands reflected in the stronger word frequency effect. Word-level challenges could increase start-up costs at the beginning of both the first and second sentences, where preceding context is absent (first sentence) or reduced in support (second sentence). The differences also could reflect a strategy in self-paced reading to focus on sentence beginnings to secure the foundation for structure building.

The persistent influence of L1 background on skilled L2 reading, even when differences in L2 proficiency are taken into account, supports the broader SRA account that language knowledge and procedural skills developed through prior L1 experience influence L2 reading. Such prior linguistic experience may shape how shared incremental processes operate, leading to differences in sensitivity to lexical and text-level factors across L1 backgrounds.

5. Conclusion

We found that skilled L1 and L2 reading processes are shaped by the same set of text factors, underscoring the shared processes of skilled reading. In L2, the magnitude of these effects varied with readers' proficiency, with additional group-specific influences of L1 background. This pattern is consistent with the SRA principle that skilled reading develops through the acquisition of knowledge of the target language and the development of proceduralized reading skills, supporting common processes across readers while allowing proficiency and L1 background to contribute to individual differences in how these processes are engaged during skilled L2 reading.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S1366728926101898.

Acknowledgements. We acknowledge support from the National Science Foundation (BCS-2118195; to Lin Chen and Charles Perfetti). We thank Gaisha Oralova for her helpful discussions. We thank Shannon Clark, Daniela Teodorescu, Alona Fyshe and Carrie Demmans Epp for training the models and generating the surprisal estimates reported in Chen et al. (2026) and used in the present study.

Data availability statement. All data and analysis code are openly available on the Open Science Framework (OSF) at: https://doi.org/10.17605/OSF.IO/WQM8U.

Competing interests. No potential conflict of interest was reported by the authors.

References

Ben-Yehudah, G., Hirshorn, E. A., Simcox, T., Perfetti, C. A., & Fiez, J. A. (2019). Chinese–English bilinguals transfer L1 lexical reading procedures and holistic orthographic coding to L2 English. *Journal of Neurolinguistics*, **50**, 136–148. https://doi.org/10.1016/j.jneuroling.2018.01.002

Bley-Vroman, R. (2009). The evolving context of the fundamental difference hypothesis. *Studies in Second Language Acquisition*, **31**(2), 175–198. https://doi.org/10.1017/S0272263109090275

Brennan, J., Nir, Y., Hasson, U., Malach, R., Heeger, D. J., & Pylkkänen, L. (2012). Syntactic structure building in the anterior temporal lobe during natural story listening. *Brain and Language*, **120**(2), 163–173. https://doi.org/10.1016/j.bandl.2010.04.002

Brothers, T., Hoversten, L. J., & Traxler, M. J. (2021). Bilinguals on the garden-path: Individual differences in syntactic ambiguity resolution. *Bilingualism: Language and Cognition*, **24**(4), 612–627. https://doi.org/10.1017/S1366728920000711

Brysbaert, M. (2019). How many words do we read per minute? A review and meta-analysis of reading rate. *Journal of Memory and Language*, **109**, 104047. https://doi.org/10.1016/j.jml.2019.104047

Brysbaert, M., Lagrou, E., & Stevens, M. (2017). Visual word recognition in a second language: A test of the lexical entrenchment hypothesis with lexical decision times. *Bilingualism: Language and Cognition*, **20**(3), 530–548. https://doi.org/10.1017/S1366728916000353

Brysbaert, M., & **New, B.** (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, **41**(4), 977–990. https://doi.org/10.3758/BRM.41.4.977

Calloway, R. C., & **Perfetti, C. A.** (2017). Integrative and predictive processes in text reading: The N400 across a sentence boundary. *Language, Cognition and Neuroscience*, **32**(8), 1001–1016. https://doi.org/10.1080/23273798.2017.1279340

Castles, A., **Rastle, K.**, & **Nation, K.** (2018). Ending the Reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, **19**(1), 5–51. https://doi.org/10.1177/1529100618772271

Chen, L., **Oralova, G.**, **Clark, S.**, **Teodorescu, D.**, **Fyshe, A.**, **Demmans Epp, C.**, **Helfrich, M.**, & **Perfetti, C.** (2025). Tracking the dynamic word-by-word incremental reading through multimeasures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **51**(8), 1324–1346. https://doi.org/10.1037/xlm0001438

Chen, L., **Oralova, G.**, **Fang, X.**, **Clark, S.**, **Teodorescu, D.**, **Helfrich, M.**, **Fyshe, A.**, **Demmans Epp, C.**, & **Perfetti, C.** (2026a). Text difficulty modulates the surprisal effect in self-paced reading. *Reading and Writing*. https://doi.org/10.1007/s11145-026-10768-7

Chen, L., **Oralova, G.**, **Helfrich, M.**, & **Perfetti, C.** (2026b). Syntactic surprisal effects in L2 English reading: EEG evidence during naturalistic text comprehension. *Language Learning (in press)*

Chen, L., & **Perfetti, C.** (2024). Learning Chinese as a second language: Implications of the character-word dual function model. *Language Teaching Research Quarterly*, **44**, 115–128. https://doi.org/10.32038/ltrq.2024.44.11

Chen, L., **Perfetti, C. A.**, **Fang, X.**, & **Chang, L.-Y.** (2021). Activation of L1 orthography in L2 word reading: Constraints from language and writing system. *Second Language Research*, **37**(2), 323–348. https://doi.org/10.1177/0267658320927761

Chen, L., **Xu, Y.**, & **Perfetti, C.** (2025). Acquiring the structure of a writing system is important in learning to read: A test of the character-word dual-focus approach in learning Chinese as a second language. *Bilingualism: Language and Cognition*, **28**(4), 1056–1067. https://doi.org/10.1017/S1366728924000920

Clahsen, H., & **Felser, C.** (2006). How native-like is non-native language processing? *Trends in Cognitive Sciences*, **10**(12), 564–570. https://doi.org/10.1016/J.TICS.2006.10.002

Clahsen, H., & **Felser, C.** (2018). Some notes on the shallow structure hypothesis. *Studies in Second Language Acquisition* **40**(3), 693–706). https://doi.org/10.1017/S0272263117000250

Cop, U., **Keuleers, E.**, **Drieghe, D.**, & **Duyck, W.** (2015). Frequency effects in monolingual and bilingual natural reading. *Psychonomic Bulletin and Review*, **22**(5), 1216–1234. https://doi.org/10.3758/s13423-015-0819-2

Demberg, V., & **Keller, F.** (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, **109**(2), 193–210. https://doi.org/10.1016/j.cognition.2008.07.008

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

Dijkstra, T., & **van Heuven, W. J. B.** (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, **5**(3), 175–197. https://doi.org/10.1017/S1366728902003012

Droop, M., & **Verhoeven, L.** (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly*, **38**(1), 78–103. https://doi.org/10.1598/rrq.38.1.4

Fernández Monsalve, I., **Frank, S. L.**, & **Vigliocco, G.** (2012). Lexical surprisal as a general predictor of reading time. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, 398–408. Association for Computational Linguistics

Ferreira, F., & **Chantavarin, S.** (2018). Integration and prediction in language processing: A synthesis of old and New. *Current Directions in Psychological Science*, **27**(6), 443–448. https://doi.org/10.1177/0963721418794491

Frank, S. L., **Otten, L. J.**, **Galli, G.**, & **Vigliocco, G.** (2015). The ERP response to the amount of information conveyed by words in sentences. *Brain and Language*, **140**, 1–11. https://doi.org/10.1016/j.bandl.2014.10.006

Frazier, L. (1985). Syntactic complexity. In D. Dowty, L. Karttunen, & A. M. Zwicky (Eds.), *Natural language parsing: Psychological, computational, and theoretical perspectives* (pp. 129–189). Cambridge University Press. https://doi.org/10.1017/CBO9780511597855.005

Gabriele, A., **Alemán Bañón, J.**, **Hoffman, L.**, **Covey, L.**, **Rossomondo, A.**, & **Fiorentino, R.** (2021). Examining variability in the processing of agreement in novice learners: Evidence from event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **47**(7), 1106–1140. https://doi.org/10.1037/xlm0000983

Graesser, A. C., **McNamara, D. S.**, & **Kulikowich, J. M.** (2011). Coh-Metrix: Providing multi-level analyses of text characteristics. *Educational Researcher*, **40**(5), 223–234. https://doi.org/10.3102/0013189X11413260

Grey, S. (2023). Variability in native and nonnative language: An ERP study of semantic and grammar processing. *Studies in Second Language Acquisition*, **45**(1), 137–166. https://doi.org/10.1017/S0272263122000055

Gu, C., **Nastase, S. A.**, **Zada, Z.**, & **Li, P.** (2025). Reading comprehension in L1 and L2 readers: Neurocomputational mechanisms revealed through large language models. *npj Science of Learning*, **10**(1), 46. https://doi.org/10.1038/s41539-025-00337-y

Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Second Meeting of the North American Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics

Hale, J., **Kuncoro, A.**, **London, D.**, **Dyer, C.**, & **Brennan, J. R.** (2019). Text Genre and Training Data Size in Human-Like Parsing. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing* (pp. 5846–5852). Association for Computational Linguistics

Hopp, H. (2022). Annual review of linguistics second language sentence processing. *Annual Review of Linguistics*, **8**, 235–256. https://doi.org/10.1146/annurev-linguistics-030821-054113

Klein, D., & **Manning, C.** (2003). Accurate unlexicalized parsing. In *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics* (pp. 423–430). Association for Computational Linguistics

Kliegl, R., **Nuthmann, A.**, & **Engbert, R.** (2006). Tracking the mind during Reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General2006*, **135**(1), 12–35. https://doi.org/10.1037/0096-3445.135.1.12

Koedinger, K. R., **Corbett, A. T.**, & **Perfetti, C.** (2012). The knowledge-learning-instruction framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive Science*, **36**(5), 757–798. https://doi.org/10.1111/J.1551-6709.2012.01245.X

Kuperman, V., **Dambacher, M.**, **Nuthmann, A.**, & **Kliegl, R.** (2010). The effect of word position on eye-movements in sentence and paragraph reading. *Quarterly Journal of Experimental Psychology*, **63**(9), 1838–1857. https://doi.org/10.1080/17470211003602412

Kuperman, V., **Stadthagen-Gonzalez, H.**, & **Brysbaert, M.** (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, **44**(4), 978–990. https://doi.org/10.3758/s13428-012-0210-4

Kutner, M., **Nachtsheim, C.**, & **Neter, J.** (2004). *Applied linear regression models* (4th ed.). McGraw Hill.

Landi, N. (2010). An examination of the relationship between reading comprehension, higher-level and lower-level reading sub-skills in adults. *Reading and Writing*, **23**(6), 701–717. https://doi.org/10.1007/s11145-009-9180-z

Lenth, R. V., & **Piaskowski, J.** (2025). *emmeans: Estimated marginal means, aka least-squares means (Version 2.0.0) [R package]*. Comprehensive R Archive Network (CRAN). https://CRAN.R-Project.Org/Package=emmeans

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, **106**(3), 1126–1177. https://doi.org/10.1016/j.cognition.2007.05.006

MacWhinney, B. (2005). A unified model of language acquisition. In P. Robinson & N. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 341–371). Routledge/Taylor & Francis Group.

Melby-Lervåg, M., & **Lervåg, A.** (2011). Cross-linguistic transfer of oral language, decoding, phonological awareness and reading comprehension: A meta-analysis of the correlational evidence. *Journal of Research in Reading*, **34**(1), 114–135. https://doi.org/10.1111/j.1467-9817.2010.01477.x

Melby-Lervåg, M., & **Lervåg, A.** (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of

https://doi.org/10.1017/S1366728926101898 Published online by Cambridge University Press

studies comparing first- and second-language learners. *Psychological Bulletin*, **140**(2), 409–433. https://doi.org/10.1037/a0033890

Merkx, D., & **Frank, S. L.** (2021). Human sentence processing: Recurrence or attention? In Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics (pp. 12–22). *Association for Computational Linguistics*. https://doi.org/10.18653/v1/2021.cmcl-1.2

Merity, S., **Xiong, C.**, **Bradbury, J.**, & **Socher, R.** (2016). Pointer Sentinel Mixture Models. *ArXiv Preprint ArXiv:1609.07843*. http://arxiv.org/abs/1609.07843

Mifka-Profozic, N., **O'Reilly, D.**, & **Guo, J.** (2020). Sensitivity to syntactic violation and semantic ambiguity in English modal verbs: A self-paced reading study. *Applied Psycholinguistics*, **41**(5), 1017–1043. https://doi.org/10.1017/S0142716420000338

Miller, G. A., & **Chomsky, N.** (1963). Finitary models of language users. In D. Luce (Ed.), *Handbook of mathematical psychology* (pp. 2–419)

Mor, B., & **Prior, A.** (2022). Frequency and predictability effects in first and second language of different script bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **48**(9), 1363–1383. https://doi.org/10.1037/xlm0000927

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development.

Nelson, M. J., & **Denny, E. C.** (1973). *The Nelson–Denny reading test*. MA: Houghton Mifflin Company.

Nisbet, K., **Bertram, R.**, **Erlinghagen, C.**, **Pieczykolan, A.**, & **Kuperman, V.** (2022). Quantifying the difference in reading fluency between L1 and L2 readers of English. *Studies in Second Language Acquisition*, **44**(2), 407–434. https://doi.org/10.1017/S0272263121000279

Oh, B. D., **Clark, C.**, & **Schuler, W.** (2022). Comparison of structural parsers and neural language models as surprisal estimators. *Frontiers in Artificial Intelligence*, **5**. https://doi.org/10.3389/frai.2022.777963

Oralova, G., **Chen, L.**, **Clark, S.**, **Teodorescu, D.**, **Helfrich, M.**, **Fyshe, A.**, **Demmans Epp, C.**, & **Perfetti, C.** (2026). Surprisal in Reading: Language models predict the N400 for L2 readers. *Language, Cognition, and Neuroscience*, **41**(1), 136–155. https://doi.org/10.1080/23273798.2025.2585303

Payne, B. R., **Lee, C. L.**, & **Federmeier, K. D.** (2015). Revisiting the incremental effects of context on word processing: Evidence from single-word event-related brain potentials. *Psychophysiology*, **52**(11), 1456–1469. https://doi.org/10.1111/psyp.12515

Pérez, A., **Hansen, L.**, & **Bajo, T.** (2019). The nature of first and second language processing: The role of cognitive control and L2 proficiency during text-level comprehension. *Bilingualism: Language and Cognition*, **22**(5), 930–948. https://doi.org/10.1017/S1366728918000846

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, **11**(4), 357–383. https://doi.org/10.1080/10888430701530730

Perfetti, C. A., & **Helder, A.** (2021). Incremental comprehension examined in event-related potentials: Word-to-text integration and structure building. *Discourse Processes*, **58**(1), 2–21. https://doi.org/10.1080/0163853X.2020.1743806

Perfetti, C. A., & **Stafura, J.** (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, **18**(1), 22–37. https://doi.org/10.1080/10888438.2013.827687

Ramirez, G., **Chen, X.**, **Geva, E.**, & **Luo, Y.** (2011). Morphological awareness and word reading in English language learners: Evidence from Spanish- and Chinese-speaking children. *Applied PsychoLinguistics*, **32**(3), 601–618. https://doi.org/10.1017/S0142716411000233

Ratcliff, R., **Gomez, P.**, & **McKoon, G.** (2004). A diffusion model account of the lexical decision task. *Psychological Review*, **111**(1), 159–182. https://doi.org/10.1037/0033-295X.111.1.159

Rayner, K., & **Duffy, S. A.** (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, **14**(3), 191–201. https://doi.org/10.3758/BF03197692

Rayner, K., **Foorman, B. R.**, **Perfetti, C. A.**, **Pesetsky, D.**, & **Seidenberg, M. S.** (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest*, **2**(2), 31–74. https://doi.org/10.1111/1529-1006.00004

Roark, B., **Bachrach, A.**, **Cardenas, C.**, & **Pallier, C.** (2009). Deriving lexical and syntactic expectation-based measures for psycholinguistic modeling via incremental top-down parsing. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing* (pp. 324–333). Association for Computational Linguistics

Román, P., **Kaan, E.**, & **Dussias, P. E.** (2022). Access to verb bias and plausibility information during syntactic processing in adult Spanish–English bilinguals. *Bilingualism: Language and Cognition*, **25**(3), 417–429. https://doi.org/10.1017/S1366728921000924

Traxler, M. J. (2005). Plausibility and verb subcategorization in temporarily ambiguous sentences: Evidence from self-paced reading. *Journal of Psycholinguistic Research*, **34**(1), 1–30. https://doi.org/10.1007/s10936-005-3629-2

Tremblay, A., **Derwing, B.**, **Libben, G.**, & **Westbury, C.** (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning*, **61**(2), 569–613. https://doi.org/10.1111/j.1467-9922.2010.00622.x

van Hell, J. G. (2023). The neurocognitive underpinnings of second language processing: Knowledge gains from the past and future outlook. *Language Learning*, **73**(2), 95–138. https://doi.org/10.1111/lang.12601

van Hell, J. G., & **Tokowicz, N.** (2010). Event-related brain potentials and second language learning: Syntactic processing in late L2 learners at different L2 proficiency levels. *Second Language Research*, **26**(1), 43–74. https://doi.org/10.1177/0267658309337637

Verhoeven, L., & **Perfetti, C. A.** (2017). *Learning to read across languages and writing systems*. Cambridge University Press. https://doi.org/10.1017/9781316155752

Wang, M., **Koda, K.**, & **Perfetti, C. A.** (2003). Alphabetic and nonalphabetic L1 effects in English word identification: A comparison of Korean and Chinese English L2 learners. *Cognition*, **87**(2), 129–149. https://doi.org/10.1016/s0010-0277(02)00232-9

Wehbe, L., **Blank, I. A.**, **Shain, C.**, **Futrell, R.**, **Levy, R.**, **der Malsburg, T.**, … **Fedorenko, E.** (2021). Incremental language comprehension difficulty predicts activity in the language network but not the multiple-demand network. *Cerebral Cortex*, **31**(9), 4006–4023. https://doi.org/10.1093/cercor/bhab065

White, S. J., **Drieghe, D.**, **Liversedge, S. P.**, & **Staub, A.** (2018). The word frequency effect during sentence reading: A linear or nonlinear effect of log frequency? *Quarterly Journal of Experimental Psychology*, **71**(1), 46–55. https://doi.org/10.1080/17470218.2016.1240813

Yamasaki, B. L., & **Prat, C. S.** (2014). The importance of managing interference for second language Reading ability: An individual differences investigation. *Discourse Processes*, **51**(5–6), 445–467. https://doi.org/10.1080/0163853X.2014.913415

Zehr, J., & **Schwarz, F.** (2018). PennController for Internet Based Experiments (IBEX). https://Osf.Io/Md832/

https://doi.org/10.1017/S1366728926101898 Published online by Cambridge University Press