Journal of Experimental Psychology: Learning, Memory, and Cognition

Frequency and Predictability Effects in First and Second Language of Different Script Bilinguals

Billy Mor and Anat Prior

Online First Publication, September 9, 2021. http://dx.doi.org/10.1037/xlm0000927

CITATION

Mor, B., & Prior, A. (2021, September 9). Frequency and Predictability Effects in First and Second Language of Different Script Bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. http://dx.doi.org/10.1037/xlm0000927

Journal of Experimental Psychology:
Learning, Memory, and Cognition

ISSN: 0278-7393

https://doi.org/10.1037/xlm0000927

Frequency and Predictability Effects in First and Second Language of Different Script Bilinguals

Billy Mor and Anat Prior
Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, Department of Learning Disabilities, Faculty of Education, University of Haifa

Reading efficiently in a second language (L2) is a crucial skill, but it is not universally achieved. Here we ask whether L2 reading efficiency is better captured as a language specific skill or whether it is mostly shared across L1 and L2, relying on general language abilities. To this end, we examined word frequency and predictability effects in sentence reading, and tested the same readers in L1 and L2, recording participants' eye-movements. Participants were 57 undergraduate bilingual speakers of Hebrew and English, languages that use different scripts, allowing for a clearer distinction between L1 and L2 processing. Both word frequency and word predictability effects were more pronounced in participants' L2 than in the L1, suggesting that both lower level and higher-order processes in reading are sensitive to language proficiency. Further, frequency effects in the L2 were linked with L2 proficiency but not general language abilities, and L2 predictability effects were not associated with either variable. Finally, readers' frequency and predictability effects in L1 and L2 were not associated with each other. Taken together, these results suggest that for different-script bilinguals, efficient reading in the L2 is a highly specific skill, dependent upon proficiency in that language, and drawing less on L1 and general language ability.

Keywords: bilingualism, eye movements, frequency, predictability, reading

In today's increasingly globalized era, foreign language skills are a key component for participating and succeeding in educational, social, and professional environments (Commission of the European Communities [EC], 2003). This is especially true when it comes to acquiring literacy in English, which is the lingua franca of global communication (Lee & Fradd, 1998; Lee et al., 2013). However, acquiring literacy in a nonnative language is different from native literacy acquisition, and the outcomes are highly variable, even among individuals who share the same environments and exposure rates (Bialystok & Hakuta, 1994; Miyake & Friedman, 1998). In the current study, we examine efficient L2 reading, as reflected in both lexical access, which is related to lower-level language skills, and using context to predict upcoming words, which is considered a higher order skill. Specifically, we ask whether L2 reading efficiency is best understood as a result of L2 specific linguistic knowledge, or whether reading efficiency is mostly shared across L1 and L2 and relies on learners' general linguistic abilities.

Billy Mor https://orcid.org/0000-0002-5045-9122
Anat Prior https://orcid.org/0000-0001-8026-2825
Correspondence concerning this article should be addressed to Anat Prior, Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, Department of Learning Disabilities, Faculty of Education, University of Haifa, 199 Abba Hushi Road, 31905 Haifa, Israel. Email: aprior@edu.haifa.ac.il

Reading Efficiency: Word Frequency

Words that appear more often in the written language, meaning high-frequency words (e.g., *star*), are recognized more quickly than low-frequency words (e.g., *tent*; e.g., Rayner & Duffy, 1986; Scarborough et al., 1977; for a review, see Brysbaert et al., 2011). In eye movements studies of adult skilled readers, *frequency effects* are evident in fixation durations, number of fixations and skipping rates (e.g., Gollan et al., 2011; Inhoff & Rayner, 1986; Just & Carpenter, 1980; Kretzschmar et al., 2015; Rayner & Duffy, 1986). High frequency words have high-quality mental representations, namely accessible knowledge about the form (phonological, morphological, syntactic, and orthographic knowledge) and meaning of a word (Perfetti, 2007; Perfetti & Hart, 2002) and thus enjoy easy lexical access (Balota & Chumbley, 1990). Frequency effects have been demonstrated across different native languages, alphabetic and nonalphabetic (for example, Cop et al., 2015; Gollan et al., 2011; Li et al., 2014; Liu et al., 2015, 2016; Pivneva et al., 2014; Rau et al., 2015; Whitford & Titone, 2012), including Hebrew (Frost, 1994; Koriat, 1984).

Word frequency effects are a signature of lexical access, and their *magnitude* (namely the difference in reading measures between low and high frequency words) reflects reading efficiency. Thus, skilled readers demonstrate smaller word frequency effects than do less skilled readers (Ashby et al., 2005) or children who are still acquiring literacy (Kuperman & Van Dyke, 2013). This relation has also been found when comparing reading in L1 versus L2, such that word frequency effects are larger in the L2 than in the L1 (Cop et al., 2015; Duyck et al., 2008; Gollan et al., 2011; Mor & Prior, 2020; Whitford & Titone, 2012, 2017). This negative relation between reading efficiency and the magnitude of frequency effects can be explained by the logarithmic nature of the function of frequency, and a ceiling effect of exposure (e.g., Monsell, 1991; Morton, 1970). Consequently, as overall exposure to written language increases in skilled readers, efficiency for processing high-frequency words stops improving, whereas additional exposure to low-frequency words leads to continuing improvements in efficiency.

Reading Efficiency: Word Predictability

Context can play an important role in lexical access, by providing preliminary semantic and syntactic cues allowing the reader to predict upcoming words. This high-order process relies on readers' constructing an online sentence representation and integrating it with existing knowledge (e.g., Morris, 2006). Word predictability supports efficient lexical access as evidenced by robust *predictability effects*, where higher-predictability words are recognized faster than lower-predictability words (e.g., Brothers & Kuperberg, 2021; Ehrlich & Rayner, 1981; Rayner & Well, 1996), reflected in fixation durations and skipping rates (e.g., Drieghe et al., 2005; Ehrlich & Rayner, 1981; Kretzschmar et al., 2015; Rayner et al., 2004; Rayner & Well, 1996; Staub, 2011). Thus, efficient lexical access benefits from both bottom-up and top-down processes. Word predictability effects have mostly been investigated in native readers of English, but were also replicated in native readers of other languages and orthographies (e.g., Kliegl et al., 2004; Miellet et al., 2007; Rayner et al., 2005). Yet, only a small number of studies have investigated word predictability effects in bilinguals reading in the L2 (Gollan et al., 2011; Libben & Titone, 2009).

Similar to the frequency effect, there is some evidence that the magnitude of the predictability effect is also negatively correlated with reading proficiency. For example, larger word predictability effects were found among average relative to highly skilled readers in the native language (Ashby et al., 2005) and in readers with mild reading deficits attributable to aphasia compared with healthy skilled readers (Huck et al., 2017). These findings seem perhaps counterintuitive, because one might suppose that skilled readers would be more adept at generating expectations for upcoming words. However, Huck and colleagues (2017) suggest a possible explanation, derived from the theoretical framework of compensatory processing (Stanovich, 1986), which predicts that readers who have weaker lexical representations, resulting in lower word recognition skill, tend to rely more on other sources of knowledge, such as context, as a compensatory strategy of the cognitive system. Along similar lines, it is possible that when online integration during reading is less efficient, the presence of preliminary semantic cues can be particularly helpful in reducing overall cognitive effort and supporting comprehension, and thus the difference in reading times between high- and low-predictability words is greater. In other words, we suggest that low-skilled readers' lexical access benefits more both from supporting context and from word frequency even though the former is a top-down process and the latter is a bottom-up process.

This line of reasoning would lead to the hypothesis that bilinguals will show larger word predictability effects when reading in their L2 compared with their L1 or compared with monolinguals. However, the few studies that have investigated this issue have mostly found comparable effects across languages (Foucart et al., 2014; Gollan et al., 2011; Whitford & Titone, 2017). Importantly, the bilingual participants in these studies were highly proficient in their L2, and most were immersed in an L2 speaking environment at the time of testing. One interesting exception are the Dutch-English bilinguals tested by Gollan and colleagues (2011), who were unbalanced (less proficient in English, their L2) and indeed showed larger predictability effects in the L2 than did readers of the language as an L1 (but only when required to perform overt naming).

Thus, it might be that above a certain level of proficiency, predictability effects in the L2 are stabilized and similar to those observed in L1, but that differences are detectable in lower proficiency readers. This issue will be investigated in the current study, by probing word predictability effects among intermediate bilingual speakers, who have lower quality lexical representations (Perfetti, 2007) and are less skilled readers in their L2 than in their L1.

Is Reading Efficiency Language Specific?

Not all learners achieve efficient reading in the L2, even after many years of study (Prior et al., 2020). One important factor to examine is whether efficient reading in the L2 can recruit underlying skills that support L1 reading efficiency, or whether it is better understood as mostly language specific. Previous research has examined several relevant factors.

Specific Language Factors: L2 Proficiency

Second language proficiency, often measured by vocabulary knowledge, has been linked to reading comprehension (e.g., Qian, 1999) and reading efficiency. Specifically, bilinguals with wider vocabulary knowledge in the L2 have smaller word frequency effects in isolated word reading (Brysbaert et al., 2017; Diependaele et al., 2013; Mor & Prior, 2020). However, a possible link between L2 proficiency and readers' ability to generate expectations of upcoming words in a sentence has yet to be directly examined. This gap is addressed in the current study.

Following the positive relation between vocabulary size and the quality of lexical representations (e.g., Baayen, 2001) and the notion that the quality of lexical representations is negatively related to the size of word predictability effects (Stanovich, 1986), we hypothesize a negative relation between L2 vocabulary knowledge and the magnitude of L2 word predictability effects. Namely, readers with smaller vocabulary knowledge will show larger word predictability effects. However, we acknowledge that it is also theoretically possible that readers with very limited vocabulary knowledge will find it more difficult, or even impossible, to make use of semantic context in real-time, and would therefore show smaller word predictability effects when compared with higher-proficiency second language readers. In other words, at the beginning stages of foreign language learning, there might be a positive relation between vocabulary knowledge and word predictability effects. Nevertheless, we suggest that the intermediate L2 population tested in the current study most likely have vocabulary knowledge that is sufficient to support ongoing sentence comprehension and prediction, and therefore that a negative relation between proficiency and predictability is the more likely outcome. In the current study, we also include a measure of L2 reading fluency, to achieve a fuller characterization of participants' skills.

In addition to proficiency, L2 exposure has also been found to modulate L2 reading and processing (Cop et al., 2015; Diependaele et al., 2013; Duyck et al., 2008; Gollan et al., 2011; Whitford & Titone, 2012). However, in practice, exposure is difficult to measure accurately, and often relies on self-reports, which are less reliable (Tomoschuk et al., 2019). Although some previous studies do report correlations between self-reported L2 exposure and objective L2 proficiency for individuals living in bilingual environments (e.g., Gilbert et al., 2020), in our own previous research (Mor & Prior, 2020) we did not find a correlation between self-reported exposure and objective proficiency in L2. This might be because the two constructs might be less well aligned under conditions of foreign language study. Moreover, in that study, self-rated exposure did not relate to reading efficiency of single words. Therefore, although we describe participants' self-reported exposure rates, we do not include them in the analyses.

General Language Factors: L1 Ability

There is considerable variation in the language performance of adults in their native language, again as measured by vocabulary knowledge. Such variability has recently been linked to education level as well as to general language aptitude (Daôbrowska, 2018). Indeed, readers with larger vocabulary knowledge in L1 demonstrate more efficient reading in that language, as evidenced in smaller frequency (Brysbaert et al., 2017; Diependaele et al., 2013) and predictability (Stanovich, 1986) effects.

If indeed L1 proficiency and vocabulary capture individuals' general language aptitude, such measures might arguably also predict performance in the L2. This notion relies on the idea of shared mechanisms supporting L1 and L2 processing. Indeed, in a recent study by Cop et al. (2015), vocabulary knowledge in the L1 (but, surprisingly, not in the L2) predicted frequency effects in both languages of Dutch-English adult bilinguals. However, the typological and orthographic similarity between the readers' L1 and L2 make it difficult to clearly disentangle vocabulary knowledge in the two languages. Specifically, in typologically close languages there is inevitable overlap in some vocabulary items due to cognates and interlingual homographs (Diependaele et al., 2013; Lemhöfer et al., 2008), making it difficult to identify unique contributions of L1 proficiency.

In contrast, participants in the current study are bilingual speakers of typologically and orthographically different languages, allowing us to assume very little overlap between L1 and L2 oral and written vocabulary knowledge, and to isolate the general language aptitude, as reflected in L1 vocabulary knowledge, from specific L2 knowledge.

L1 Word Frequency and Predictability Effects

The current study goes beyond previous research, by investigating a possible link between efficient reading processes in the L1 and efficient reading processes in the L2 within the same individual. These will allow us to ask whether the ability to build high-quality lexical representations (Perfetti, 2007) and/or use preceding context to predict upcoming words, might be a characteristic of an individual, which can then be expressed in the various languages she uses. Specifically, individuals who more easily construct word representations, following exposure to a given word, will have smaller word frequency effects in the L1 than individuals who require a larger number of exposures to create a durable lexical representation for low-frequency words. The question we wish to examine here is to what degree such efficiency might generalize across languages. In other words, would individuals with smaller frequency effects in L1 also demonstrate smaller effects in the L2? In a previous study we examined this possibility in a lexical-decision task (Mor & Prior, 2020) and did not find an association between frequency effects in L1 and L2. In the current study we reexamine this possibility in a naturalistic reading task, which in contrast to the lexical-decision task, does not include decision components.

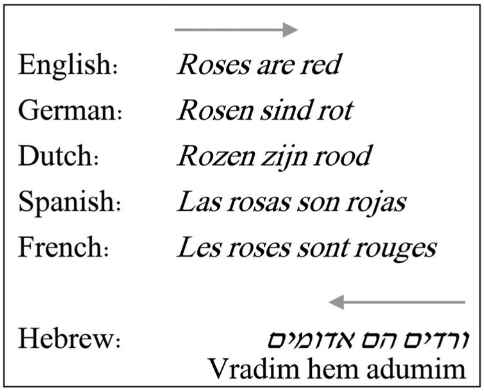
Along similar lines, we investigate the possibility that the ability to generate expectations for upcoming words might also be shared across languages of bilinguals. Indeed, this higher-order skill, which involves top-down processes such as integration, might be even more likely to reflect a characteristic of the individual and thus be utilized by both L1 and L2. Such prediction goes beyond successfully extracting meaning from preceding words but also requires integrating them with the reader's existing knowledge. This existing knowledge is in part language specific knowledge of sentence structures and phrases but is also to a large degree general world and semantic knowledge, that is independent of a specific language. Therefore, we wish to examine whether individuals with larger predictability effects in L1 also demonstrate larger effects in the L2, an issue that has not yet been investigated.

Same- Versus Different-Script Bilinguals

Previous studies examining frequency and predictability effects in bilinguals have almost exclusively examined bilinguals of typologically similar languages (Dutch-English, French-English, German-English, Spanish-English) who share a similar writing system that uses the Roman alphabet (Duyck et al., 2008), whereas many bilinguals in the world are speakers of typologically and orthographically different languages. This issue is of great importance, because in same-script bilinguals there might be significant links between L1 and L2 which are derived from specific language exposure. This relative familiarity between the languages affects word recognition in the L2, due to inevitable overlap in some vocabulary items, such as cognates and interlingual homographs (Diependaele et al., 2013; Lemhöfer et al., 2008). Also, similarity between languages in syntactic rules (for a review see Tolentino & Tokowicz, 2011) might affect high order reading processes, and even ease cognitive load. Consequently, sharing of underlying skills and similarities in reading processes might be more evident in same-script bilinguals but will not necessarily generalize to wider bilingual populations.

Thus, to allow for a better differentiated examination of L2 processing, in the current study we examine bilingual speakers of typologically different languages that do not share a writing system, assuming very little overlap between L1 and L2 oral and written knowledge (Mor & Prior, 2020). Specifically, we examine Hebrew-English bilinguals living in Israel, where English has a high status, is learned as a mandatory foreign language in a school setting from the third grade, and is widely present in the environment (for a review, see Shohamy, 2014). However, recent findings demonstrate that Israeli adolescents achieve English reading skills that are comparable with those achieved by native English speakers who are 6–8 years younger (Prior et al., 2020), a finding which might be at least partly explained by the typologically and orthographic distance between English and Hebrew. Hebrew is a

Figure 1
An Illustration of Distance Between Languages



Note. The same sentence translation, written in different languages. Italics were used to differentiate between the example sentence and the name of the language.

Semitic language, characterized by a rich morphology, which consists of derivations of mostly tri-consonantal roots, as opposed to the linear and sequential morphology that characterizes English (Bick et al., 2011; Frost et al., 1997; Prior & Markus, 2014). Further, The Hebrew script is mostly a consonantal system, and is written from right to left, as opposed to languages that use the Roman alphabet (see Figure 1).

Most Hebrew sentences have an SVO syntactic structure, but it is far from obligatory, as it is in English. The rich morphology of Hebrew allows for great flexibility in syntactic structure, but in the current study all Hebrew stimuli were SVO sentences. Another difference between the languages is in adjective placement, which is postnominal in Hebrew (but prenominal in English). Here is an example Hebrew sentence with a literal gloss in English ("3" = 3rd person, "SG" = singular, "M" = masculine, "PST" = past tense, "F" = feminine):

ha-dayag yašav betox sira ktana asuya me-ec

the-fisherman sit.3SG.M.PST inside boat small.SG.F made.SG.F from-wood

'The fisherman sat inside a small boat made of wood'

Although examination of eye movements in Hebrew reading in and of itself was not a central motivation of the current study, it is still worth noting that this is the first eye movement study to examine word frequency and word predictability effects in readers of Hebrew.

As stated above, previous studies have investigated mainly highly proficient L2 speakers, or balanced bilinguals (Foucart et al., 2014; Gollan et al., 2011; Whitford & Titone, 2017). Because reading efficiency, and its expression in word frequency and predictability effects, might differ across proficiency levels, in the current study we chose to examine a population of intermediate-advanced L2 readers, to see how well previous results might generalize. Of great importance, the current study investigates L1 and L2 reading within the same reader, whereas most previous studies compared bilinguals' L2 reading to monolingual native readers of the same language (e.g., Gollan et al., 2011). This design was chosen so that we could address the question of skill sharing across languages.

The Current Study

The current study examines whether specific language proficiency or general language abilities might better capture variance in efficient L2 reading, as reflected in word frequency and predictability effects. Reading is measured by tracking the readers' eye movements during sentence reading. L2 proficiency was measured by a vocabulary test and a single word reading fluency test. General language ability (as expressed in L1) was indexed by the readers' L1 vocabulary knowledge and reading fluency, alongside L1 word frequency effects (as a predictor of L2 frequency effects), and L1 word predictability effects (as a predictor of these effects in the L2).

In contrast to most previous studies, the current study explores L1 and L2 within the same reader and focuses on intermediate proficiency different script bilinguals. Thus, the population of the current study is unbalanced Hebrew-English bilinguals, who have been studying the second language for at least nine years and have reached intermediate proficiency. As noted, the significant implication of examining bilinguals of two typologically different languages that do not share a writing system, beyond generalization, is that it allows a clearer examination of L2 processing, being highly distinct from the L1.

Hypotheses

Word Frequency Effects. As a baseline, we expect to replicate the well-established findings of word frequency effects in both languages, with larger effects in the L2. As opposed to most previous research that used between-participants designs, the current study examines L1 and L2 effect sizes within the same reader. Next, we hypothesize a negative relation between L2 proficiency and the magnitude of L2 word frequency effects, supported by findings from previous research (Cop et al., 2015; Diependaele et al., 2013; Duyck et al., 2008; Gollan et al., 2011; Monaghan et al., 2017; Whitford & Titone, 2012). We also wish to examine whether the pattern of results found in our previous study of single word reading, within the same population as examined in the current study, of different-script intermediate L2 speakers (Mor & Prior, 2020) also holds for reading in context. Further, we hypothesize a possible negative relation between general language abilities and the magnitude of L2 word frequency effects (found in Cop et al., 2015), within the same reader, assuming general linguistic mechanisms that underly reading efficiency across languages. This suggestion did not reach significance in our single word study; however, it is possible that the language abilities captured by L1 proficiency might relate more strongly to L2 frequency effects as manifest in contextual reading, as in Cop et al. (2015), than in isolated, less natural, single word reading. Finally, following the assumption regarding general linguistic mechanism, we examine whether the ability to build high-quality lexical representations (Perfetti, 2007) might be a characteristic of an individual and therefore manifest in a possible relation between frequency effects in L1 and L2, although such a relation did not emerge in our previous study of single word reading (Mor & Prior, 2020).

Word Predictability Effects. As a baseline, we expect to replicate previous findings of word predictability effects in L1 and L2 reading. We also examine whether reading proficiency modulates word predictability effects. Following findings from monolingual studies, which found larger word predictability effects for less skilled- than skilled- readers (Ashby et al., 2005; Huck et al., 2017), we hypothesize larger word predictability effects in L2, participants' less proficient language, than in their L1. This stems from the assumption that readers will have a stronger need to rely on context in L2, as a compensation for weaker lexical representations (Stanovich, 1986). Previous bilingual studies examining highly proficient L2 speakers have reported equal effects size across languages (Foucart et al., 2014; Gollan et al., 2011; Whitford & Titone, 2012), but we hypothesize that in the currently examined population of intermediate L2 speakers, such cross-language differences might emerge. Next, we tentatively hypothesize a possible negative relation between L2 proficiency and the magnitude of L2 word predictability effects, although the limited research available on this topic makes it difficult to formulate specific predictions. However, following the positive relation between exposure/proficiency and reading skill (e.g., Baayen, 2001; Payne et al., 2012) and the notion that the quality of lexical representations is negatively related to the size of word predictability effect (Stanovich, 1986), we assume the above. Finally, assuming possible general linguistic mechanisms which underlie efficiency across languages, we also hypothesize that the magnitude of word predictability effects in L1 and L2 might be correlated.

Method

Participants

Sixty-six university students (41 females; mean age 24.89 [2.9], range 18–31; mean years of education 13.63 [1.87], range 12–19) were recruited through advertisements offering payment or course credit for participating. Data from 9 participants were excluded due to technical problems resulting in partially lost data, leading to a final sample of 57 participants.[1] All participants were unbalanced Hebrew-English bilinguals, with Hebrew as their native and dominant language (L1), and English as their second language (L2), learned as a foreign language in a school setting, from third grade. Participants were not proficient in any other language. All participants had normal or corrected-to-normal vision (however participants wearing contact lenses were not recruited), without reported reading disability, attention disorders, or language impairment. All participants provided written informed consent. The study was approved by the Institutional Review Board of the University of Haifa.

Materials

Language History Questionnaire (LEAP-Q)

A Hebrew translation (Prior & Beznos, 2009) of the *Language Experience and Proficiency Questionnaire* was used (*LEAP-Q*; Marian et al., 2007; Cronbach's alpha for L2 = .88, for L1 = .92). The questionnaire includes questions regarding language exposure, use, and proficiency, in the context of oral and written language, and yields scores for Hebrew and English proficiency and exposure on a 1–10 scale.

Language Proficiency

Language proficiency in each language was assessed using a vocabulary knowledge test and a reading fluency test.

Vocabulary Knowledge Test (Shipley). Two versions of the vocabulary test, in Hebrew and English (English version: Shipley, 1946; split half internal consistency reported as .87 by Zachary, 1991; Hebrew version: Gilboa, 2008) were administered. The test consists of 40 multiple-choice questions, presented in writing, in which participants are asked to choose which of four words is closest in meaning to a target word, with no time limit. The score is the number of correct responses.

Reading Fluency Test (TOWRE). An English version and a Hebrew version of the test of Word Reading Efficiency (TOWRE; English version: Torgesen et al., 1999; average test–retest reliability $> .9$; Hebrew version: Katzir et al., 2012; Cronbach's $\alpha = .95$) were used to measure word reading efficiency in participants' L1 and L2. In each version, the test contains 104 words, ordered in increasing level of difficulty, arranged in four columns. The participant is required to read aloud as many words as possible within 45 seconds. Separate scores were calculated for reading fluency in Hebrew and English. The Hebrew version presented words without diacritics.

Sentence Reading Tasks (Eye Movements)

Participants performed two sentence reading tasks, one in their L1 (Hebrew) and one in their L2 (English). Each task included 60 target words, ranging in frequency between 1.55 and 420.61 per million. Words in the two languages were not translation equivalents of each other. Each target word was embedded in two different sentences, one where the target word was highly predictable and the other where it had low predictability (see Materials description below). The high-predictability condition was set to prompt the target word, whereas the low-predictability condition was set as neutral, allowing the target word as a possible completion, but also many other words. Each participant was presented with each target word only once, in either high or low predictability condition. Thus, even though there were 120 sentences in each language, each participant read only 60 sentences in each language. For counterbalancing purposes, the list of 60 target words was divided in two parts—high- and low-frequency words—and each subset included an equal number of high- and low-predictability sentences, such that each participant read an equal number of sentences in each of the four conditions (15 sentences of *high*-frequency-*high*-predictability words, 15 sentences of *high*-frequency-*low*-predictability words, and so on).

Each sentence was presented separately and appeared as one line on the screen. The target word was always followed by at least

[1] Following Brysbaert and Stevens (2018), we ran a power analysis for the current experiment using a utility developed by Westfall et al. (2014; https://jakewestfall.shinyapps.io/two_factor_power/), which took into account the number of participants and stimuli as random factors. Because we do not have reliable estimates from previous research regarding the expected effect sizes for some of the variables investigated in the current study, we chose conservatively small effect sizes for the power analysis. For effect sizes of 0.3 the original sample of 66 participants yielded a power estimate of 0.86, and the final sample of 57 participants yielded a power estimate of 0.823. For a slightly larger, although still small, effect size of 0.4 the power estimates are 0.98 and 0.97, respectively. These estimates demonstrate that the current design was adequately powered.

two words before the sentence end, since sentence-final words might be subjected to biased reading patterns (Raney et al., 2014).

Target Words. English target words were taken from the stimuli in Gollan et al. (2011), which originally included 90 nouns as target words, half high-frequency and half low-frequency, based on a standard frequency corpus (*Time* Magazine Corpus; Davies, 2007). This corpus includes 100 million tokens taken from written magazine text. To adapt the stimulus set to the current study population, we eliminated Hebrew-English cognates from the list and pretested the remaining list in a lexical-decision task to make sure that participants, who are L2 English speakers, were indeed familiar with the experimental items. This pretest lexical-decision task was completed by 18 undergraduates from the University of Haifa (who did not participate in the main experiment, nor in the other pretest and norming tasks, see below). Only words that were correctly identified by at least 80% of participants were retained.

The Hebrew target words were not translation equivalents of the English target words, but they were selected to match the English nouns by semantic categories. The nouns in the two languages were also matched in terms of frequency (and length, see below). Frequency counts for the Hebrew nouns were based on the heTenTen 2014 corpus via Sketch Engine (Kilgarriff et al., 2004). This corpus includes 890 million words extracted from the Internet. The Hebrew nouns were also matched to the English list in length in phonemes (see Table 1 for target words characteristics), but not length in letters, because most vowel letters are omitted in the written form of Hebrew. Because length in letters has an influence on word recognition reaction time (RT; e.g., New et al., 2006), this factor was treated as a control variable in all analyses.

Sentences. The English Sentences for the high- and low- predictability conditions were also taken from Gollan et al. (2011), who determined predictability in two rounds of a cloze task, each with a sample of 20 American university students. The Hebrew sentences were not translations of the English sentences and were created for the current study by constructing two sentences to each target word, one each for the high- and low- predictability conditions. Word predictability norms in Hebrew were established using a cloze task in two samples (two rounds) of 20 undergraduates from the University of Haifa, who did not participate in the target words pretest nor in the main experiment. Participants were presented with the beginning of each sentence up to the target word and asked to complete the sentence with the very first word that comes to their mind. Sentences that did not meet criterion (target produced by at least 60% of the participant in the high-predictability condition, and by no more than 30% of the participants in the low predictability condition) in the first round, were rewritten and retested in the second round.

These pretest participants, who completed the Hebrew sentence norming test, also performed a cloze test on all sentences from the original Gollan et al. (2011) study, for two purposes: First, to gauge the degree to which Israeli participants were able to achieve basic understanding of the English materials; second, to ascertain that the cloze probabilities of the target items within this new population were comparable to those reported from native English speakers in the original study. In line with the first goal, only sentences that were completed by at least 50% or pretest participants with a plausible word (even if it was not the intended target word) were retained in the stimulus list. With regard to the second goal, results showed that the targets retained differences in predictability, as reported by Gollan and colleagues (2011), but the actual cloze probabilities themselves were not all identical. Specifically, the cloze probability of target words in the low predictability condition collected from the Hebrew-English bilinguals was comparable with those collected from native English monolinguals. However, the cloze probability of target words in the high-predictability condition collected in the current pretest from bilinguals was lower than those completed by native speakers in Gollan et al. (2011; 66.4% and 89.9%, respectively). This difference is expected and reflects the fact that the participants in the current study were of intermediate proficiency. Importantly, because the low-predictability condition maintained low values, there was still a robust difference in predictability across conditions (see Table 2 for examples of sentences in the different conditions). In all the analyses, the percent of participants who responded with the target word given the preceding sentence context (cloze probability) was entered into the models as the predictability value of that word.

Task Procedure. In each trial of the sentence reading task, a fixation point appeared on the screen, in the location where the first letter of the upcoming sentence would later appear. After the participant pressed a button, and only while her eye fixated on the fixation point, the sentence appeared on the screen, and remained on the screen until the participant fixated on a square at the bottom of the screen, below the end of the sentence. Participants were instructed to read silently to comprehend. The sentences were presented in a random order, and 20% of the sentences (also randomly) were followed by a yes/no comprehension question, requiring the

Table 1
Target Word and Sentence Characteristics in L1 (Hebrew) and L2 (English)

	Target frequency per million		Target length in phonemes		Target length in letters[a]		Sentence length in words[a]		Target predictability – high condition		Target predictability – low condition	
Measure	Hebrew	English	Hebrew	English	Hebrew	English	Hebrew	English	Hebrew	English	Hebrew	English
M	49.91	49.65	4.7	4.38	3.92	5.22	9.58	12.9	91.33	89.92	2	3.08
SD	76.92	72.55	1.18	1.09	.89	.98	2.5	2.55	11.71	11.41	6.32	5.9
Max	409.8	420.61	8	9	6	9	17	18	100	100	30	20
Min	2	1.55	3	3	2	4	5	7	60	50	0	0

Note. Frequency counts for the English target words were taken from Gollan et al. (2011), based on time Magazine Corpus (Davies, 2007). Frequency counts for the Hebrew target words were based on the heTenTen 2014 corpus via Sketch Engine (Kilgarriff et al., 2004). *M* = means; *SD* = standard deviations; *Max* = maximum; *Min* = minimum.
[a] Means of L1 and L2 target word length in letters and sentence length are significantly different, $p < .001$.

Table 2
Examples of Sentences in Different Predictability Conditions Across Languages

Measure	Hebrew	English
High Predictability	לכבוד יום הולדתי, אמי אפתה עוגה בטעם שוקולד. *English literal translation: For my birthday, my mother baked a* ***cake*** *chocolate flavor.*	She read the story of Goldilocks and the three **bears** to her daughter.
Low Predictability	בתוך הקופסה תמצא עוגה בטעם שוקולד. *English literal translation: In the box, you will find a* ***cake*** *chocolate flavor.*	She was frightened that she would run into some **bears** but she was safe.

Note. Each experimental sentence was presented separately and appeared on the screen as one line. Italics signify the literal translation; bold signifies the target word in the sentence.

participants to respond by button press. These questions were included to make sure that participants remained engaged in the task (see Gollan et al., 2011). Six practice items preceded the experimental list.

Apparatus

Eye movements were recorded using a SR Research EyeLink 1000 eye tracker, tower-mounted to reduce head movements. Participants were seated 55 cm from a 24-in. LCD monitor (Benq XL2411) with 1,024 × 768 pixel resolution and a refresh rate of 60 Hz. Text was presented in black on a white background, in 14-point Courier New font, with an average of 3 letters equaling 1° of visual angle. Reading was binocular, but eye movements were recorded from the right eye. The testing room was dimly lit.

Procedure

Participants were tested individually in a single session lasting one hour. Participants first completed the sentence reading tasks in both languages, order counterbalanced across participants. Next, participants were tested in the reading fluency tests (TOWRE) and the vocabulary knowledge test (Shipley). Finally, participants completed the language proficiency and exposure questionnaire (LEAP-Q). Participants were offered breaks between tasks.

Results

Participant characteristics are presented in Table 3. As expected, there were significant differences between participants' L1 and L2 abilities, indicating higher exposure and proficiency scores for Hebrew (L1) than for English (L2; first order correlations between these variables are presented in Table A1, in the Appendix).

Comprehension during sentence reading was highly accurate in both languages (96% in Hebrew, 86% in English), but accuracy rates still reflected differences in proficiency between L1 and L2. Performance was analyzed using linear mixed-effects (LME) models (Baayen et al., 2008) in R (R Core Team, 2018), with the lme4 library (Version 1.1–7, Bates et al., 2015). LME models are preferred in the current design over Analyses of Variance, because they are less vulnerable to missing data points, and also retain full information of performance, without the need for averaging across items or participants (Baayen et al., 2008). Plots were created using the ggplot2 package (Version 2.3.00; Wickham, 2016). The p values were derived using Satterthwaite approximations of degrees of freedom in the lmerTest function, an approach found to produce acceptable type I error rates (Luke, 2017).

Eye Movement Measures

Reading performance was analyzed for the target words embedded within the sentences. Fixations shorter than 80 ms were excluded from analyses or merged with a longer fixation if they were located within 10 pixels or less from the longer fixation (4.8% of all data points; e.g., Fitzsimmons & Drieghe, 2013; Hermena et al., 2019). Eye movement measures were calculated using the Get Reading Measures application available through SR Research. The analyses include four typical standard eye movements measure in reading research (e.g., Cop et al., 2015; Gollan

Table 3
Participant Characteristics: Means (Standard Deviations) and Ranges

Characteristic	Hebrew	English
Age of language acquisition	—	8.42 (1.78), 3–16
Reading fluency[a]	82.14 (10.16), 60–99	73.61 (9.86), 50–99
Vocabulary knowledge[a]	70.8 (14.73), 35–92.5	49.42 (10.95), 27.5–75
Self-rated language exposure[a]	8.44 (1.23), 6–10	5.08 (1.73), .1.4–8.6
Self-rated oral language proficiency[a]	9.52 (0.61), 8–10	6.88 (1.16), 3.5–9.5

Note. Reading fluency = % of words correctly read in 45 seconds (Word Reading Efficiency [TOWRE]); Self-rated oral language proficiency = an average score of two measures: oral language comprehension and production, on a 1–10 scale (Language Experience and Proficiency Questionnaire [LEAP-Q]); Vocabulary knowledge = % correct (Shipley test); Self-rated language exposure = an average score of five measures: current exposure to Audio, TV, reading, family setting, and social setting, on a 1–10 scale (LEAP-Q).
[a] Means of L1 and L2 are significantly different, $p < .001$.

et al., 2011; for a review, see Rayner, 2009) and a fifth alternative measure. The four typical measures were *first fixation duration* (the duration of the first fixation on the target word, in first pass reading), *gaze duration* (the sum of all fixation durations on the target word, in first pass reading), *total reading time* (the sum of all fixations durations on the target word in a given trial), and *Skipping rate* (the probability that the target word was skipped). First fixation duration, gaze duration, and skipping rate are considered as early-stage measures of lexical access, whereas total reading time represents late-stage measures of higher-order reading processes (Rayner, 2009). Several previous studies have also used alternative measures to account for trade-offs between skipping rate and reading times in early-stage measures, and thus to allow a full appreciation of reading performance (e.g., Gollan et al., 2011; Rayner et al., 2011). These studies included skipped words into the averaging of gaze durations, with a fixation value zero. In the current study, the trade-off between skipping rate and reading times was especially prominent in early-stage measures in Hebrew. As a first step, we calculated the alternative gaze duration measure, following previous research, but analyses using this measure did not differ from those of the standard gaze duration measure. We therefore created an *alternative first fixation duration*, by including skipped words as having a first fixation duration of zero. In this alternative measure we found a significant L1 word frequency effect, which was important for predicting individual differences (see models' description below). Therefore, we chose to include the alternative measure of first fixation durations.

The Effects of Word Frequency and Word Predictability in L1 and L2

Model Descriptions

Similar LME models were applied to all fixation duration measures. The models included the following fixed factors: language (Hebrew, English; a categorical factor with English set as the reference), word frequency (a continuous factor, log transformed to normalize the distribution and centered), word predictability (continuous), and word length in letters as a control variable (continuous). The model also included interactions between language, word frequency and word predictability. The model included random effect terms, to capture variance associated with specific participants (e.g., overall reading speed) or items, beyond that associated with the investigated factors. The random effect structure included intercepts by items and by participants and random slopes of language by participant (e.g., Friesen et al., 2020). For skipping rate, we applied a binomial model, but an interactive model similar to the model described above failed to converge, hence we constructed an additive model, with the fixed factors of language (Hebrew, English, categorical), word frequency (continuous), word predictability (continuous), and word length in letters (continuous). The random effect structure included intercepts by item and by participant.

The analyses of the interactive models yielded significant interactions with language (as will be detailed below). Thus, we also computed separate models for each language, namely, a model including only L1 words and another including only L2 words, with the fixed factors of word frequency (continuous), word predictability (continuous), and word length in letters as a control variable (continuous). The random effect structure included intercepts by items and by participants and random slopes of word frequency by participant. The models predicting skipping rates separately in L1 and L2 failed to converge with the full random structure and therefore included only intercepts for participants and items. The importance of adding word length in letters to all models as a control variable stems from the difference in length between L1 and L2 words (Hebrew words are shorter than English words) and from the fact that this variable has a strong influence on eye-movement measures. Compared with shorter words, longer words receive longer reading durations and are skipped less often (for a review, see Rayner, 2009).

Language

Across all reading time measures (first fixation duration, gaze duration, total reading time, and alternative first fixation duration), reading in L2 was slower than in L1, a main effect of language (see Table 4 for effect sizes, standard errors, and *t* values).

Word Frequency

Reading times (across all measures, first fixation duration, gaze duration, total reading time, skipping rates, and alternative first fixation duration) were slower and skipping rates were lower for lower-frequency than for high-frequency words, a significant main effect of word frequency (see Table 4). These frequency effects were more pronounced for L2 than for L1, as shown by a significant interaction between language and word frequency (even after controlling for word length) for first fixation duration, gaze duration, and total reading time. The single-language models (see Table 5 for L1 and Table 6 for L2) revealed that the effect of word frequency was significant in the L2 for all measures (first fixation duration, gaze duration, total reading time, skipping rates, and alternative first fixation duration; see Figure 2 for total reading times), but in the L1 only for skipping rates and alternative first fixation duration.

Word Predictability

Less predictable words were read more slowly than more predictable words, but this was evident on in a significant main effect of word predictability for total reading time, (see Table 4). For this same measure, we also found that predictability effects were more pronounced in the L2 than in the L1, as evidenced by a significant interaction between language and word predictability (see Figure 3). The single-language models indicated that the effects of word predictability in total reading times were significant in both languages.

There was also a significant three-way interaction between language, word frequency, and word predictability for gaze duration and total reading time. The single language models revealed significant frequency-predictability interactions in L1 for skipping rates and alternative first fixation duration, and in L2 for total reading time and alternative first fixation duration, due to longer reading times/lower skipping rates for words that are of lower-frequency and are less predictable.

Finally, averages for fixation durations and skipping rates across conditions are presented in Table 7. Word frequency and word predictability were treated as continuous variables in the analyses presented above, but here we offer a categorical presentation for ease of interpretation.

Table 4
Language, Word Frequency, and Word Predictability Effects

	FFD			GD			TT			SR			Alternative FFD		
Fixed effects	*b*	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*z*	*b*	*SE*	*t*
(intercept)	268	12.1	22.2***	302.21	32.63	9.26***	420.3	58	7.13***	.5	.42	1.17	193	15.4	12.51***
Language	−52.1	6.28	−8.3***	−120.03	17.95	−6.69***	−195.23	30.73	−6.4***	−1.37	.18	−7.8***	−69.6	8.25	−8.44***
Word frequency	−31.5	5.46	−5.77***	−87.85	13.27	−6.62***	−155.62	24.1	−6.46***	−.4	.13	−3.11**	−31.3	7.14	−4.38***
Word predictability	.01	.03	.4	−.08	.06	−1.37	−.46	.09	−4.9***	.0	.0	.59	0	.04	−.02
Language × Word Frequency	31.2	7.88	3.97***	96.31	18.94	5.08***	165.39	34.07	4.85***	—	—	—	11.8	10	1.18
Language × Word Predictability	−0.05	.05	−.99	.07	.09	.8	.33	.14	2.43*	—	—	—	0	.05	−.03
Word Frequency × Word Predictability	.14	.07	2.14*	.27	.12	2.24*	.61	.19	3.19**	—	—	—	.17	.08	2.16*
Language × Word Frequency × Word Predictability	−.12	.1	−1.21	−.42	.18	−2.33*	−.78	.28	−2.83**	—	—	—	.01	.11	.13
Control variable															
Word length in letters	.01	2	.01	15.83	5.26	3.01**	17.76	9.72	1.83	.67	.08	8.94***	12.9	2.64	4.89***
Random effects	Variance	*SD*		Variance	*SD*		Variance	*SD*		Variance	*SD*		Variance	*SD*	
Word (intercept)	163.1	12.77		1,648	40.6		6,176	78.59		.14	.38		364.7	19.1	
Participant (intercept)	1,215.7	34.87		12,260	110.7		33,125	182		.6	.78		1,312.5	36.23	
Language	948.5	30.8		10,309	101.5		27,080	164.56		—	—		1,689.2	41.1	

Note. FFD = first fixation durations; GD = gaze durations; TT = total reading times; SR = skipping rates; Alternative FFD = alternative first fixation durations. *b* = effect sizes; *SE* = standard errors; *t* = *t* values; *z* = *z* values.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 5
Word Frequency and Word Predictability Effects in Participants' L1

	FFD			GD			TT			SR			Alternative FFD		
Fixed effects	*b*	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*z*	*b*	*SE*	*t*
(intercept)	209.87	8.7	24.1***	197.5	16.07	12.3***	225.53	30.66	7.36***	−.94	.35	−2.67**	94.80	13.80	6.88***
Word frequency	.21	4.63	.04	6.7	8.6	.8	10.34	15.38	.67	−.63	.19	−3.3***	−16.50	7.07	−2.34*
Word predictability	−.03	.03	−1.2	−.01	.04	−.03	−.13	.06	−2.1*	.0	.0	.62	0.00	0.04	−.08
Word Frequency × Word Predictability	.03	.06	.47	−.14	.08	−1.67	−.18	.12	−1.46	.01	.0	2.42*	0.19	0.08	2.47*
Control variable															
Word length in letters	1.54	1.99	.77	11	3.74	3.21**	17.62	7.21	2.44*	.69	.08	8.2***	20.10	3.22	6.24***
Random effects	Variance	*SD*		Variance	*SD*		Variance	*SD*		Variance	*SD*		Variance	*SD*	
Word (intercept)	101.1	10.05		492.1	22.18		2,151	46.38		.19	.43		360.33	18.98	
Participant (intercept)	363.4	19.06		1,014.9	31.86		3,302	57.46		.67	.82		777.48	27.88	
Word frequency	56.8	7.54		331.7	18.21		1,039	32.24		-	-		11.42	3.38	

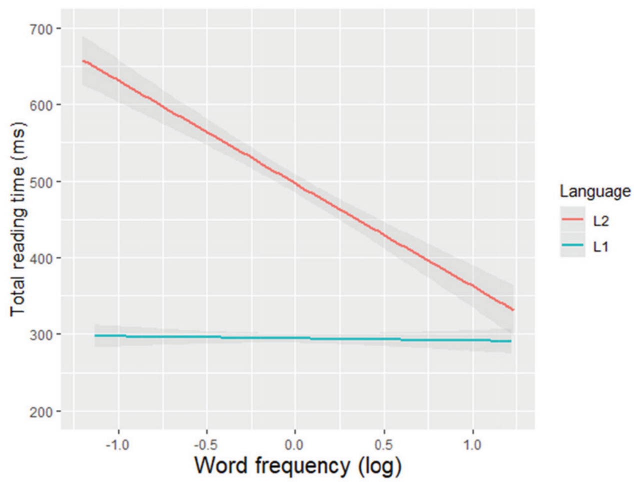
Note. FFD = first fixation durations; GD = gaze durations; TT = total reading times; SR = skipping rates; Alternative FFD = alternative first fixation durations; *b* = effect sizes; *SE* = standard errors; *t* = *t* values; *z* = *z* values.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 6
Word Frequency and Word Predictability Effects in Participants' L2

	FFD			GD			TT			SR			Alternative FFD		
Fixed effects	*b*	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*z*	*b*	*SE*	*t*
(intercept)	281.42	21.24	13.25***	265.29	62.82	4.22***	419.5	116.42	3.6***	1.57	1.32	1.19	266.00	22.40	11.88***
Word frequency	−32.36	6.43	−5.04***	−85.94	18.76	−4.58***	−155.68	33.25	−4.68***	−.8	.35	−2.26*	−35.50	6.74	−5.27***
Word predictability	.01	.04	.33	−.08	.07	−1.1	−.45	.12	−3.94***	−.0	.0	−1.4	0.00	0.04	−.09
Word Frequency × Word Predictability	.15	.08	1.91	.28	.15	1.9	.62	.24	2.64**	.0	.01	1.37	0.18	0.08	2.19*
Control variable															
Word length in letters	−2.18	3.8	−.64	22.67	11.2	2.02*	17.89	20.9	.86	.52	.25	2.11*	−0.53	3.99	−.13
Random effects	Variance	*SD*		Variance	*SD*		Variance	*SD*		Variance	*SD*		Variance	*SD*	
Word (intercept)	224.92	15		2,804	52		10,339	101.68		.0	.0		246.99	15.72	
participant (intercept)	1,176.46	34.3		12,028	109.7		32,474	180.2		.88	.94		1,301.76	36.08	
Word frequency	11.09	3.33		3,342	57.81		7,549	86.88		—	—		2.6	1.61	

Note. FFD = first fixation durations; GD = gaze durations; TT = total reading times; SR = skipping rates; Alternative FFD = alternative first fixation durations; *b* = effect sizes; *SE* = standard errors; *t* = *t* values; *z* = *z* values.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Figure 2
Word Frequency Effects in Participants' L1 and L2 Total Reading Times



Note. Word frequency (log-transformed and centered) ranges from lower (−1) to higher (1). Shaded areas represent the 95% confidence intervals. See the online article for the color version of this figure.

Modulation of L2 Reading Efficiency

Model Description – L2 Word Frequency

To analyze whether L2 frequency effects are language specific or recruit general language skills, we ran an LME model to predict reading of L2 target words within sentences. Similar models were applied to all fixation duration measures, which yielded a significant word frequency effect in L2 as in the previous analyses. We examined the possible predictive role of the following fixed factors (all continuous): L2 proficiency, L1 proficiency, and the participant's sensitivity to frequency in L1. Participants' language proficiency scores, in each language, were derived from the vocabulary knowledge test and reading fluency test and were computed by applying a principal component analysis to the tests, resulting in a single score for each participant in each language. To assess each participant's sensitivity to frequency in L1, we calculated the first-order correlation between word frequency and alternative first fixation durations in L1 for each individual participant. In this measure, lower values represent lower sensitivity to L1 word frequency, and higher values represent higher sensitivity. The correlation value for each participant was then entered into the model. We chose to use the alternative first fixation duration, because this was the only measure where significant L1 word frequency effects were apparent in the former analyses. Because we were interested in whether the investigated variables might modulate the word frequency effects across participants, the effects of interest in the current model were possible interactions between each predictor variable and word frequency in the L2. Word predictability and word length were included as control variables. The random effect structure included intercepts by item and by participant (fuller random structures failed to converge).

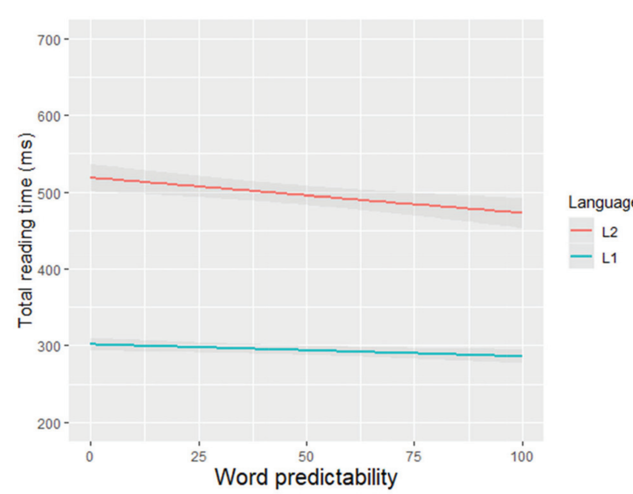
L2 Word Frequency Effects

Effect sizes, standard errors, and *t* values for reading measures are presented in Table 8. Participants with higher L2 proficiency had shorter reading times across all reading time measures (first fixation duration, gaze duration, total reading time, and alternative first fixation duration), a significant main effect of L2 proficiency. Further, participants with higher L2 proficiency demonstrated smaller L2 frequency effects, namely significant interactions between L2 proficiency and word frequency in first fixation duration, gaze duration, and total reading time (see Figure 4) but not in alternative first fixation duration. The main effects of L1 proficiency and of sensitivity to frequency in the L1 were not significant, and they did not interact significantly with L2 word frequency in the analyses of gaze duration, total reading time, skipping rates, and alternative first fixation duration. The interaction between L1 proficiency and L2 word frequency was significant only in the analysis of first-fixation duration, such that participants with higher L1 proficiency demonstrated larger L2 frequency effects. Because this effect was only evident in a single reading measure, we are hesitant to offer any interpretation of it before it can be replicated in future studies. See Table A2 in the Appendix for correlation matrix of the predictor variables.

Model Description – L2 Word Predictability

To examine whether L2 proficiency or general language ability might modulate predictability effects in the L2, we ran an LME model which was similar to the one constructed to examine possible predictors of frequency effects in the L2, as detailed above, with three changes: First, instead of participants' sensitivity to frequency in L1, here we assessed each individual participant's sensitivity to word predictability in L1, by calculating the first order correlation between word predictability and total reading time in L1. Here as well, lower values represent lower sensitivity to L1 word predictability, and higher values represent higher sensitivity. We based our

Figure 3
Word Predictability Effects in Participants' L1 and L2 Total Reading Times



Note. Cloze probability ranges from lower (0) to higher (100). Shaded areas represent the 95% confidence intervals. See the online article for the color version of this figure.

Table 7

A Categorical Presentation of Word Frequency and Word Predictability Performance Across All Eye Movements Measures, During L1 (Hebrew) and L2 (English) Sentence Reading: Means (Standard Errors)

Measure	Hebrew		English	
	Low frequency	High frequency	Low frequency	High frequency
First fixation durations (ms)				
Low predictability	218 (10)	215 (9)	284 (15)	255 (13)
High predictability	212 (9)	215 (9)	280 (15)	259 (12)
Gaze durations (ms)				
Low predictability	247 (13)	248 (14)	427 (33)	351 (23)
High predictability	252 (15)	243 (13)	419 (35)	344 (24)
Total reading times (ms)				
Low predictability	300 (22)	300 (24)	598 (56)	436 (36)
High predictability	299 (22)	278 (19)	540 (57)	415 (38)
Skipping rates (%)				
Low predictability	15 (5)	24 (6)	2 (2)	3 (2)
High predictability	16 (5)	19 (5)	2 (2)	3 (2)
Alternative first fixation durations (ms)				
Low predictability	186 (14)	163 (14)	279 (15)	248 (14)
High predictability	178 (13)	173 (14)	273 (16)	251 (14)

Note. Word frequency and word predictability are treated as *continuous* variables in the current study; however, here we offer a categorical presentation of the data for the ease of interpretation. For this purpose, target words list was halved to high- and low-frequency levels, in each language. Differentiation between high and low word predictability levels was set in the preliminary norming study.

calculation on total reading times because this was the only measure that revealed a significant main effect for predictability in L1 in the first stage of analysis.

Second, the effects of interest were possible interactions between each predictor and L2 word predictability. Finally, word frequency was included as a control variable. We ran the model for participants' total reading time, which was the only eye movement measure revealing a significant main effect for predictability in L2.

L2 Word Predictability Effects

Effect sizes, standard errors, and *t* values for gaze durations are presented in Table 9. Here again, participants with higher L2 proficiency had faster reading times, a significant main effect of L2 proficiency. However, the interaction between L2 proficiency and word predictability was not significant. The main effects of L1 proficiency and sensitivity to predictability in L1 were not significant, nor were the interactions between word predictability and all other variables.

Discussion

In the current study, we examined whether L2 reading efficiency in different script bilinguals can be better understood as a language specific skill, or whether it also recruits general language abilities, that are shared with L1. Specifically, we examined L2 lexical access through the manifestation of word frequency effects, and readers' ability to generate expectations for the upcoming word, a higher-order reading process which affects lexical access, through word predictability effects in sentence reading. We chose to investigate this issue in a population of unbalanced Hebrew–English bilinguals, who speak typologically different languages that do not share a writing system, to examine how well patterns previously reported for bilingual speakers of closely related languages might generalize. The first and main finding demonstrated that only language specific vocabulary knowledge modulated the efficiency of L2 lexical access, as will be discussed in detail below. In addition, the current findings support the notion that the magnitude of word frequency effects is a marker of efficient word reading, whereas the magnitude of word predictability effects might be a marker of higher order reading efficiency. However, whereas the magnitude of word frequency effects is discriminative even between high levels of language proficiency, the word predictability effect seems less sensitive to differences in proficiency at the higher end of the scale.

Lexical Access in L1 and L2

Replicating previous studies (e.g., Duyck et al., 2008; Gollan et al., 2011; Whitford & Titone, 2017), we found significant word frequency effects in both L1 and L2, which were larger in the latter. Word frequency effects interacted with language proficiency also within the same language, as participants with larger vocabulary knowledge in the L2 had smaller L2 frequency effects. Taken together, these findings support the notion that the magnitude of word frequency effects is a marker of reading efficiency, which is highly associated with language use and proficiency, as predicted by the frequency-lag hypothesis (Gollan et al., 2011) and suggested by the lexical entrenchment approach (Brysbaert et al., 2017; Diependaele et al., 2013). The frequency-lag hypothesis addresses the difference in the amount of language exposure and use between bilinguals and monolinguals, claiming that larger frequency effects for bilinguals are a result of the fact that they use each language only part of the time. Bilinguals, then, have less exposure to word forms in each language compared with monolinguals, leading to less efficient lexical retrieval, especially for low-frequency words, and thus resulting in larger word frequency effects for bilinguals (Gollan et al., 2011). The lexical entrenchment theory (Brysbaert et al., 2017; Diependaele et al., 2013) focuses on the difference in language exposure and use between native/dominant and nonnative/nondominant languages, as in the case with unbalanced bilinguals. Relatively lower language exposure for non-native language leads to lower lexical accessibility, and eventually to larger word frequency effects, compared with the native language. Namely, both approaches indicate the amount of language exposure and experience as the explanatory factor for larger

Table 8
Predicting Participants' L2 Word Frequency Effects

Fixed effects	FFD			GD			TT			SR			Alternative FFD		
	b	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*z*	*b*	*SE*	*t*
(intercept)	280.22	21.09	13.29***	263.79	62.6	4.21***	421.94	117.06	3.6***	1.48	1.31	1.13	263.86	22.16	11.91***
L2 word frequency	−26.11	5.31	−4.92***	−74.31	15.81	−4.7***	−129.19	29.65	−4.36***	−.37	.24	−1.55	−27.64	5.56	−4.97***
L2 proficiency	−27.08	5.79	−4.68***	−89.14	15.28	−5.83***	−131.95	25.79	−5.12***	−.03	.25	−.12	−27.15	6.17	−4.4***
L1 proficiency	10.23	5.82	1.76	14.84	15.38	.97	10.15	25.95	.39	−.18	.27	−.67	9.3	6.21	1.5
L1 sensitivity to frequency	45.4	31.2	1.46	149.35	82.35	1.81	268.15	138.97	1.93	.54	1.35	.4	49.33	33.27	1.48
L2 Word Frequency × L2 Proficiency	10.88	4.82	2.26*	47.45	9.32	5.09***	49.43	14.82	3.33***	−.14	.33	−.41	9.69	5.1	1.9
L2 Word Frequency × L1 Proficiency	−9.7	4.84	−2.01*	−12.8	9.34	−1.37	10.48	14.9	.7	.18	.39	.48	−9.28	5.14	−1.81
L2 Word Frequency × L1 Sensitivity to Frequency	280.22	21.09	13.29	18.85	50.19	.38	−56.64	79.79	−.71	2.62	1.81	1.44	−21.02	27.49	−.76
Control variable															
L2 word predictability	.01	.04	.15	−.08	.07	−1.01	−.44	.12	−3.77***	.0	.0	−1.12	−.01	.04	−.24
L2 word length in letters	−1.99	3.79	−0.53	23.49	11.32	2.08*	18.55	21.23	.87	.53	.25	2.14*	.01	3.98	.0
Random effects	Variance	*SD*		Variance	*SD*		Variance	*SD*		Variance	*SD*		Variance	*SD*	
Word (intercept)	221.48	14.88		240.18	15.5		10,646.87	103.18		.0	.0		2,848.65	53.37	
Participant (intercept)	779.63	27.92		890.16	29.84		17,124.93	130.86		.79	.89		5,946.21	77.11	

Note. FFD = first fixation durations; GD = gaze durations; TT = total reading times; SR = skipping rates; Alternative FFD = alternative first fixation durations; *b* = effect sizes; *SE* = standard errors; *t* = *t* values; *z* = *z* values.
* $p < .05$. *** $p < .001$.

word frequency effects in bilinguals' L2, whether it is compared with monolinguals or bilinguals' L1.

The current study also revealed an unexpected finding. Whereas L2 word frequency effects were evident in both early and late stages of lexical access, L1 word frequency effects were identified only in an early stage of lexical access, namely in skipping rates and alternative fixation duration (which is a measure which also accounts for skipping). This finding is surprising, because many previous studies of highly proficient adults reading in their native language, have reported robust frequency effects in both early and late stages of lexical access (e.g., Ashby et al., 2005; Jared et al., 1999). Further, in our own previous work we found robust frequency effects in Hebrew (L1) for a mostly overlapping stimulus set and a comparable sample size, using a lexical-decision task (Mor & Prior, 2020).

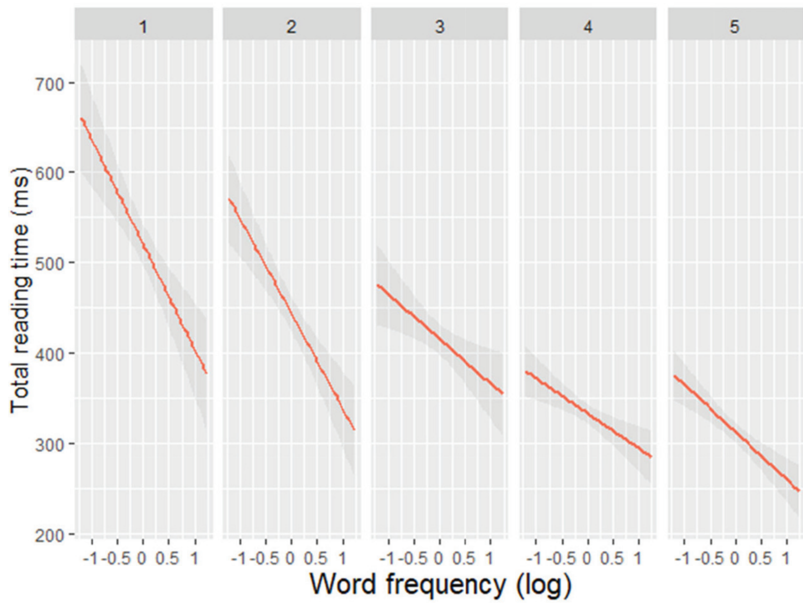
It is possible that the L1 pattern found in the current study is related to the specific orthographic characteristics of Hebrew, which is more densely packed than English. This feature of the Hebrew writing system might facilitate the extraction of information from the parafovea, namely before the reader's eye is fixated on the target word (Deutsch et al., 2003), and thus allow higher skipping rates when reading in Hebrew. Although an earlier study suggested that the difference between readers of Hebrew and English is reflected in shorter saccades and not in fixation durations (Pollatsek et al., 1981), the current results showed relatively, though not extremely, high skipping rates in the L1. There is some variability in the reported skipping rates of content words in the literature. Rayner (2009) states that the expected skipping rate is about 15% of content words. Whitford and Titone (2014) also reported skipping rates of 15% for high-frequency words but only 5% for low-frequency words, for monolinguals reading in English. However, Gollan et al. (2011) reported somewhat higher skipping rates, 19.7% for high-frequency words and 17.7% for low-frequency words. The current skipping rates in L1 reading are again on the high side for both high frequency (21.77%) and low frequency (15.61%) words. It is important to note that our participants demonstrated relatively low skipping rates in English (2.87% for high-frequency words and 1.99% for low-frequency words), which is expected and reflects their intermediate proficiency in the L2. This also suggests that the high skipping rate observed in Hebrew is not a result of lost data points or other experimental error, but rather reflects participants' natural behavior. Because the current study is the first to use eye movement measures to investigate word frequency effects in Hebrew, this issue should be further investigated.

The Ability to Generate Expectations in L1 and L2

The intermediate-advanced L2 speakers tested in the current study were able to efficiently use preceding context to predict upcoming words in both their L1 and L2. This finding of a stable word predictability effect, in Hebrew-English bilinguals, adds to a small number of studies of fairly highly proficient L2 speakers who showed word predictability effects in the L2 (Foucart et al., 2014; Gollan et al., 2011; Libben & Titone, 2009; Martin et al., 2013; Whitford & Titone, 2017). Our finding demonstrates that even less skilled L2 readers can use context to facilitate word reading fairly efficiently.

These previous studies demonstrated comparable word predictability effects in the two languages of same-script bilinguals, and thus concluded that the ability to efficiently use preceding context to

Figure 4
Second Language (L2) Frequency Effects by L2 Proficiency in Participants' Total Reading Times



Note. The level of participants' language proficiency ranges from lower (1) to higher (5), word frequency (log-transformed and centered) ranges from lower (−1) to higher (1). Shaded areas represent the 95% confidence intervals. See the online article for the color version of this figure.

predict upcoming words is language-invariant (Whitford & Titone, 2017; see also Foucart et al., 2014; Gollan et al., 2011). In contrast, participants in the current study demonstrated larger word predictability effects in their L2 compared with their L1. We suggest that the theoretical framework of compensatory processing (Stanovich, 1986) might offer an explanation for the variability in findings across studies, by focusing on reading skill as a key factor. Specifically, less skilled L2 readers might rely more strongly on contextual cues, as a way of enlisting available supporting resource to reduce effort, which is a compensatory cognitive mechanism. Namely, when the overall reading process is less efficient, syntactic and semantic cues might be more useful, resulting in stronger predictability effects. This pattern has been reported in L1 readers, in studies that compared average and highly skilled readers (Ashby et al., 2005) and readers with mild aphasia with healthy readers (Huck et al., 2017). As described above, the bilingual populations investigated in previous studies

Table 9
Predicting Participants' L2 Word Predictability Effects in Total Reading Times

Fixed effects	*b*	*SE*	*t*
(intercept)	407.21	115.65	3.52***
L2 word predictability	−.45	.12	−3.79***
L2 proficiency	−137.21	26.97	−5.09***
L1 proficiency	3.14	27.11	.12
L1 sensitivity to predictability	−163.84	142.37	−1.15
L2 Word Predictability × L2 Proficiency	.22	.16	1.33
L2 Word Predictability × L1 Proficiency	.14	.16	.85
L2 Word Predictability × L1 Sensitivity to Predictability	−.27	.86	−.31
Control variable			
L2 word frequency	−126.72	29.25	−4.33***
L2 word length in letters	19.33	20.95	.92
Random effects	**Variance**	***SD***	
Word (intercept)	10,364.68	101.81	
Participant (intercept)	17,972.02	134.06	

Note. b = effect sizes; *SE* = standard errors; t = t values.
*** $p < .001$.

were mostly highly proficient in their L2, some were immersed in the L2 at the time of testing, in addition to sharing similar writing systems. Thus, they might have been sufficiently efficient readers in L2 (even if not fully equal to L1) that the magnitude of word predictability effect was no longer discriminative. In contrast, the bilinguals in the current study were intermediate proficiency readers of the L2, studied as a foreign language.

Finally, it should be noted that the predictability effects in the current study were evident only in a late-stage measure of lexical access, total reading time, a pattern that was stable across both languages. This finding aligns well with the theoretical characterization of late-stage eye movement measures as reflecting higher order processes, such as syntactic and semantic integration and generating expectations for upcoming words. However, this pattern of results is inconsistent with some previous studies, which reported word predictability effects in both early- and late- stage measures, among monolingual (e.g., Drieghe et al., 2005; Ehrlich & Rayner, 1981; Kretzschmar et al., 2015; Rayner et al., 2004; Rayner & Well, 1996; Staub, 2011) and bilingual readers (e.g., Gollan et al., 2011; Whitford & Titone, 2017). A possible explanation for the pattern of our results is offered following Kliegl et al. (2004), who studied monolingual readers, and found word predictability effects for both early- and late- stage measures, but the effects were more pronounced in late stage measures. Specifically, the effect of word predictability was larger on total reading time than on gaze duration, a difference that reflects the additional time the readers spent refixating low-predictability words compared with high-predictability words. Thus, it is possible that whereas in previous studies predictability effects were evident in first and second pass reading and were more pronounced in the latter, in the current study they were significant only in second pass reading but failed to reach significance in earlier measures.

Lexical Access and the Ability to Generate Expectations: A Possible Interaction?

In the current study we also found a frequency-predictability interaction, although most previous studies indicated that the effects of word predictability and of word frequency do not interact but are rather additive. Some have suggested that the few studies reporting a significant interaction should be understood as sporadic findings, which may have originated in bias (Staub, 2015). In the current study, we found larger word frequency effects for low-predictability words than for high-predictability words, but this was manifested differently across language. In L1 reading, the interaction was evident only in an early stage of lexical access that accounts for skipping (skipping rates and alternative first fixation durations), aligning with the findings from Gollan et al. (2011). This leads us to cautiously suggest that rather than a bias (Staub, 2015), this finding may indicate a valid pattern. However, it is clear that this issue warrants further investigation. In contrast to the L1, findings from the L2 revealed significant interactions in both early- and late- stage measures of lexical access (alternative first fixation durations and total reading times), a pattern that may arguably align with the results from Huck et al. (2017), who found a trend of frequency-predictability interaction in readers with mild reading deficits, but not in their control group of skilled readers. Specifically, the significant interaction in that study indicated larger word predictability effects for low frequency words than for high-frequency words. The authors explained their findings using the framework of compensatory processing (Stanovich, 1986), which we suggest can also explain the current findings, because it predicts that readers with weaker lexical representations tend to rely more on sources other than the orthographic input, such as context, as a compensatory strategy. In the current study as well, predictability effects in L2 were stronger for low-frequency than for high-frequency items, suggesting recruitment of contextual information to compensate for lower quality lexical representations.

Modulators of L2 Lexical Access

Language Specific Factors

Results showed that L2 proficiency proved to be important for understanding variability in L2 lexical access; frequency effects in L2 word recognition among readers with lower L2 proficiency were larger compared with readers with higher L2 skills. This finding aligns well with relatively recent studies reporting larger vocabulary knowledge or exposure to be associated with smaller L2 frequency effects, both in single word (Brysbaert et al., 2017; Diependaele et al., 2013; Mor & Prior, 2020) and contextual reading times (Whitford & Titone, 2012). The current results reinforce and support this pattern and demonstrate that L2 proficiency modulates frequency effects in both early and late stages of lexical access.

These results align with the theoretical suggestions made by both the frequency-lag hypothesis (Gollan et al., 2011) and the lexical entrenchment approach (Brysbaert et al., 2017; Diependaele et al., 2013), which emphasize the importance of specific language use in establishing lexical representations, expressed in the magnitude of word frequency effects. Our findings, therefore, reinforce the notion that lower proficiency, but not cross language competition, is most likely the reason for increased L2 frequency effects in general.

General Language Abilities

In addition to examining the role of L2 proficiency in supporting efficient reading, the current study also examined whether efficient L2 reading is associated with general language abilities, as reflected in L1 proficiency and L1 frequency effects.

L1 proficiency. Participants' proficiency in their *native* language is assumed to reflect their general language aptitude. A previous study examining Dutch-English bilinguals reported L1 vocabulary knowledge as a significant predictor of L2 frequency effects in eye-movements (Cop et al., 2015), but the current study examining Hebrew-English bilinguals does not replicate this finding. This divergence across studies might be explained by the specific bilingual population tested in each of the studies. Whereas the current study examined bilingual speakers of two typologically and orthographically distant languages, Cop et al. (2015) tested speakers of typologically similar languages, which share an alphabet and have unavoidable overlap in some vocabulary items (Diependaele et al., 2013; Lemhöfer et al., 2008; Schepens et al., 2012). Under these circumstances, knowledge of L1 Dutch can play a stronger role in predicting performance of L2 English than knowledge of L1 Hebrew can (see similar results in Mor & Prior, 2020). Therefore, future research should further investigate the possible contribution of general language abilities, as reflected in L1 vocabulary knowledge, to the quality of L2 lexical representations. The current results also stress the importance of examining bilinguals of diverse profiles especially in regard to typological distance and overlap of orthographic systems or lack thereof, when investigating commonalities in processing across languages (as also recently suggested by Jiang, 2018; Mishra, 2019; Van Heuven & Wen, 2019).

L1 Lexical Representations. In the current study we hypothesized that the magnitude of L1 frequency effects, as a measure of readers' general ability to create high-quality lexical representations (Perfetti, 2007), might explain variance in L2 frequency effects, but this was not borne out by the data. Specifically, individual participants' correlation between word frequency and fixation durations in L1 did not interact significantly with word frequency in predicting reading times of L2 words. In other words, participants who were less sensitive to word frequency in L1, were not necessarily less sensitive to word frequency in L2. This pattern replicates findings from a previous study of isolated word reading (Mor & Prior, 2020).

Although null effects can only be interpreted with some caution, the current results suggest that the mechanisms supporting the creation of lexical representations in L1 and L2 might only partially overlap with each other. Further, the relation between frequency effects in L1 and L2 may depend on typological and orthographic similarity between the languages, as suggested above for the effect of L1 language proficiency more generally. Thus, especially for speakers of Hebrew and English, who show qualitative differences in the principles governing lexical organization (e.g., the importance of morphology for Hebrew, see Frost et al., 1997; Kolan et al., 2011; Prior & Markus, 2014), the process of creating high-quality lexical representations (Perfetti, 2007) that affect lexical access may be relatively language specific. Certainly, this is a speculative suggestion, and again requires further investigation of different and varied bilingual populations to gain a fuller understanding of this issue.

Modulators of Generating Expectations in L2

Because our goal was to examine efficient reading processes, in addition to lexical access we also probed a higher and more complex reading process that modulates lexical access, namely readers' ability to generate expectations of the upcoming word, as manifested in word predictability effects. As in lexical access, we asked whether specific versus general language variables are related to participants' ability to generate expectations in the L2. We also examined whether the ability to generate expectations is related across L1 and L2 of the same reader.

Language-Specific Factors

Results showed that whereas L2 proficiency did explain variance in L2 lexical access, it was not a significant predictor of the ability to generate expectations in the L2. Although this issue has not been directly examined in previous research, we did expect to find a relation, as explained in the introduction. One possible explanation for the lack of a significant correlation, again with the caution of interpreting null effects, is that L2 vocabulary knowledge and reading fluency do not capture the relevant aspect of L2 proficiency. Indeed, the ability to efficiently use preceding context takes more than adequate representations in the mental lexicon, that is, lexical knowledge is probably necessary but may not be sufficient to support predictive processes in reading. Obviously, familiarity with the target word is necessary for being able to predict it from context, but readers must also be able to integrate linguistic information, have good syntactic understanding, and retrieve relevant schemas. Although vocabulary knowledge is a strong predictor of reading comprehension in general (e.g., Cromley & Azevedo, 2007; Ouellette, 2006; Perfetti & Stafura, 2014; Qian, 1999; Wagner et al., 1997), comprehension and prediction also rely on top-down processing. Certainly, it would be interesting to examine in future research whether more holistic measures of L2 proficiency might be related to readers' ability to efficiently use preceding context and make predictions at the word level.

General Language Abilities

L1 Proficiency. As was the case for lexical access, L1 proficiency did not explain unique variance in L2 word predictability effects. A possible explanation is that domain-general abilities are not the main and most essential factor in capturing individual differences in L2 reading processes, whether these processes are lexical access or of higher order. Specifically, regarding vocabulary knowledge and reading fluency, it might be argued that these variables do not fully capture the aspects of general language ability that would be most relevant for predictive processing.

L1 Ability to Generate Expectations for Upcoming Words. We examined here whether the ability to generate expectations for upcoming words is a general characteristic of the individual, independent of specific language proficiency, and therefore might manifest similarly across L1 and L2. Compared with the ability to create high-quality lexical representations (Perfetti, 2007), the ability to generate expectations for upcoming words seems at first to be a better candidate for reflecting a general capability, because it might involve top-down process more prominently. Somewhat surprisingly, this hypothesized relation was not evident in the current data. Namely, predictability effects of the same participants in L1 and L2 were unrelated. The current study is the first to directly test this relation, and thus we can only offer a speculative explanation at this point. We suggest that the fact that participants' ability to predict upcoming words in L1 was not related to their performance in L2 might be a consequence of the specific characteristics of the tested population. Two aspects are important here. First, because of the great typological distance between the L1 and the L2 of the participants, in terms of lexicon, morpho-syntactic structure and writing system, the predictive processes developed through the L1 might not be fully applicable to L2 sentences, thus limiting similarities in performance across the two languages. Second, participants in the current study had achieved only intermediate proficiency in the L2, which limited their prediction efficiency, as evidenced by the overall larger predictability effects in L2. This limited prediction efficiency might have then masked more subtle differences in the application of higher-order predictive processes, those which can also support prediction in the L1. This explanation has clearly testable predictions—namely, that predictability effects in L1 and L2 should be better aligned in speakers of typologically close languages, and in speakers who have more balanced levels of proficiency in the two languages. Future research examining these issues is highly desirable.

Limitations and Future Directions

In the current study, we probed second language reading in a population of different script bilinguals who are intermediate-advanced L2 speakers. This choice was motivated by the fact that many bilinguals world-wide share this profile of unbalanced different script bilingualism (Jiang, 2018; Mishra, 2019). Further, we argue that both factors, namely L2 proficiency and the distance between L1 and L2, are important for understanding second language reading efficiency. However, because previous studies focused mostly on highly proficient bilinguals, who also read languages that used the same script, it is difficult to draw a firm conclusion regarding the relative role of each factor in the currently observed reading behavior. Although we based our conclusions on theory and previous evidence (e.g., the magnitude of word predictability effects as a marker of reading proficiency), future research can further elaborate these initial findings to investigate readers with different L2 proficiency levels, within the same bilingual population. It would be especially valuable to investigate the development of reading efficiency in low proficient L2 speakers, about whom less is known. Alternatively, and more challenging, would be to investigate readers with similar L2 proficiency levels, but sampled from various different-script bilingual populations (for a recent example, see Fadlon et al., 2019).

Comparing different languages within the same reader, which we see as a strength of the current study, is also a challenge in and of itself. Using reading materials in different languages may cause unexpected and possibly confounding differences between conditions. In the current study, we extended significant effort to adequately matching stimuli across languages and to controlling relevant variables, yet this examination is still not as precise as comparing bilinguals reading in their L2 with monolinguals reading in the same language. However, the within participant design has an outstanding advantage, namely the possibility to compare L1 and L2 reading abilities within the same reader, and to probe possible shared mechanism, which we would expect to be evident in both L1 and L2 reading of the same reader. Thus, future research can overcome the issue of potential bias due to matching challenges by providing cumulative evidence on L1 and L2 reading behavior within the same individual, again across various bilingual populations.

Such research from a wide range of bilingual populations is also necessary for further elaborating the initial findings of the current study regarding the modulators of efficient L2 reading. It seems that more holistic language-specific abilities should be included as possible predictors for higher-order reading process, such as the ability to generate expectations in the L2. Also, it would be interesting to further develop the notion of reading abilities as a characteristic of the individual, namely the ability to create high-quality lexical representations (Perfetti, 2007) and the ability to generate expectations. Finally, the findings from the current study, which support the notion that efficient reading in the L2 is language-specific, raise the question of whether most L2 literacy skills and processes are, essentially, language specific.

Conclusion

The results of the current study demonstrated less efficient reading in L2 than in L1 for intermediate proficiency bilinguals—both in bottom-up processing (lexical access) and in top-down processing (the ability to generate expectation). Importantly, this was the case even though the participants we tested had been studying their L2 for more than 10 years and were exposed to it on a daily basis, especially in the written form (see also Prior et al., 2020). This pattern suggests that the magnitude of word frequency effects is a marker of reading efficiency, associated with language use and proficiency (Brysbaert et al., 2017; Diependaele et al., 2013; Gollan et al., 2011) and that the magnitude of word predictability effects might be a marker of higher order reading efficiency, indexing compensatory processing (Stanovich, 1986).

In addition to the overall difference between L1 and L2, bilinguals' proficiency in the L2 proved to be important for understanding variability in the efficiency of L2 lexical access. In contrast, the specific measures of L2 proficiency implemented here were not associated with the higher-order process of prediction in reading. Importantly, efficiency of lexical access was not associated across the two languages of the bilingual participants, suggesting that the mechanisms supporting the establishment of high-quality lexical representations (Perfetti, 2007) may only partially overlap in readers of different-script and typologically distant languages (Mor & Prior, 2020). Similarly, higher-order prediction efficiency was also not associated across the languages of bilinguals, possibly owing to the gaps in proficiency and the typological distance between the two languages used by the current bilingual population. Taken together, these results suggest that for the different-script bilinguals tested here, efficient reading in the L2 is a specialized skill, distinct from efficient L1 reading (see also Mor & Prior, 2020).

More generally, the current results, which highlight the distinction between efficient reading in the L1 and L2, have important implications. On the downside, even readers who have achieved a high level of skill in the L1 are not guaranteed to reach equivalent levels of skill when studying a second language, especially if the L2 is very different from the L1 (Prior et al., 2020). On a more encouraging note, however, the current results suggest that to achieve efficient reading in an L2, individuals should strive to broaden their vocabulary knowledge and the rapid accessibility of mappings from form to meaning in that language. Thus, even individuals who do not have exceptional language abilities in the L1 should be able to achieve efficient reading in the L2, with sufficient exposure and practice. Given the increasingly central role of nonnative language skills for education and economic opportunities, both for immigrants and for foreign language learners, this is an important and optimistic finding.

References

Ashby, J., Rayner, K., & Clifton, C., Jr. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *The Quarterly Journal of Experimental Psychology Section A*, *58*(6), 1065–1086. https://doi.org/10.1080/02724980443000476

Baayen, R. H. (2001). *Word frequency distributions*. Springer Science & Business Media. https://doi.org/10.1007/978-94-010-0844-0

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Balota, D. A., & Chumbley, J. I. (1990). Where are the effects of frequency in visual word recognition tasks? Right where we said they were! Comment on Monsell, Doyle, and Haggard (1989). *Journal of Experimental Psychology: General*, *119*(2), 231–237. https://doi.org/10.1037//0096-3445.119.2.231

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). lme4: Linear mixed-effects models using Eigen and S4 (R Package Version 1, 1–7).

Bialystok, E., & Hakuta, K. (1994). *In other words: The science and psychology of second-language acquisition*. Basic Books.

Bick, A. S., Goelman, G., & Frost, R. (2011). Hebrew brain vs. English brain: Language modulates the way it is processed. *Journal of Cognitive Neuroscience*, *23*(9), 2280–2290. https://doi.org/10.1162/jocn.2010.21583

Brothers, T., & Kuperberg, G. (2021). Word predictability effects are linear, not logarithmic: Implications for probabilistic models of sentence comprehension. *Journal of Memory and Language*, *116*, Article 104174. https://doi.org/10.1016/j.jml.2020.104174

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology*, *58*(5), 412–424. https://doi.org/10.1027/1618-3169/a000123

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

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

Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, *99*(2), 311–325. https://doi.org/10.1037/0022-0663.99.2.311

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

Davies, M. (2007). Time Magazine corpus: 100 million words, 1920s–2000s. http://corpus.byu.edu/time

Deutsch, A., Frost, R., Pelleg, S., Pollatsek, A., & Rayner, K. (2003). Early morphological effects in reading: Evidence from parafoveal preview benefit in Hebrew. *Psychonomic Bulletin & Review*, *10*(2), 415–422. https://doi.org/10.3758/BF03196500

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: Human Experimental Psychology*, *66*(5), 843–863. https://doi.org/10.1080/17470218.2012.720994

Drieghe, D., Rayner, K., & Pollatsek, A. (2005). Eye movements and word skipping during reading revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(5), 954–959. https://doi.org/10.1037/0096-1523.31.5.954

Duyck, W., Vanderelst, D., Desmet, T., & Hartsuiker, R. J. (2008). The frequency effect in second-language visual word recognition. *Psychonomic Bulletin & Review*, *15*(4), 850–855. https://doi.org/10.3758/PBR.15.4.850

EC. (2003). *Promoting language learning and linguistic diversity: An action plan 2004–2006*.

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior*, *20*(6), 641–655. https://doi.org/10.1016/S0022-5371(81)90220-6

Fadlon, J., Li, C., Prior, A., & Gollan, T. H. (2019). Using what's there: Bilinguals adaptively rely on orthographic and color cues to achieve language control. *Cognition*, *191*, 103990. https://doi.org/10.1016/j.cognition.2019.06.002

Fitzsimmons, G., & Drieghe, D. (2013). How fast can predictability influence word skipping during reading? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(4), 1054–1063. https://doi.org/10.1037/a0030909

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(5), 1461–1469. https://doi.org/10.1037/a0036756

Friesen, D., Whitford, V., Titone, D., & Jared, D. (2020). The impact of individual differences on cross-language activation of meaning by phonology. *Bilingualism: Language and Cognition*, *23*(2), 323–343. https://doi.org/10.1017/S1366728919000142

Frost, R. (1994). Prelexical and postlexical strategies in reading: Evidence from a deep and a shallow orthography. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(1), 116–129. https://doi.org/10.1037/0278-7393.20.1.116

Frost, R., Forster, K. I., & Deutsch, A. (1997). What can we learn from the morphology of Hebrew? A masked-priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(4), 829–856. https://doi.org/10.1037/0278-7393.23.4.829

Gilbert, A. C., Cousineau-Perusse, M., & Titone, D. (2020). L2 exposure modulates the scope of planning during first and second language production. *Bilingualism: Language and Cognition*, *23*(5), 1093–1113. https://doi.org/10.1017/S1366728920000115

Gilboa, A. (2008). *Hebrew version of the Shipley Vocabulary Scale*.

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General*, *140*(2), 186–209. https://doi.org/10.1037/a0022256

Hermena, E. W., Liversedge, S. P., Bouamama, S., & Drieghe, D. (2019). Orthographic and root frequency effects in Arabic: Evidence from eye movements and lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(5), 934–954. https://doi.org/10.1037/xlm0000626

Huck, A., Thompson, R. L., Cruice, M., & Marshall, J. (2017). Effects of word frequency and contextual predictability on sentence reading in aphasia: An eye movement analysis. *Aphasiology*, *31*(11), 1307–1332. https://doi.org/10.1080/02687038.2017.1278741

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics*, *40*(6), 431–439. https://doi.org/10.3758/BF03208203

Jared, D., Levy, B. A., & Rayner, K. (1999). The role of phonology in the activation of word meanings during reading: Evidence from proofreading and eye movements. *Journal of Experimental Psychology: General*, *128*(3), 219–264. https://doi.org/10.1037/0096-3445.128.3.219

Jiang, N. (2018). Phonology based bilingual activation among different-script bilinguals? *Bilingualism: Language and Cognition*, *22*(4), 693–694. https://doi.org/10.1017/S1366728918000664

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329–354. https://doi.org/10.1037/0033-295X.87.4.329

Katzir, T., Schiff, R., & Kim, Y. S. (2012). The effects of orthographic consistency on reading development: A within and between cross-linguistic study of fluency and accuracy among fourth grade English-and Hebrew-speaking children. *Learning and Individual Differences*, *22*(6), 673–679. https://doi.org/10.1016/j.lindif.2012.07.002

Kilgarriff, A., Rychly, P., Smrz, P., & Tugwell, D. (2004). Itri-04-08 the sketch engine. *Información Tecnológica*, *105*, 116

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *The European Journal of Cognitive Psychology*, *16*(1-2), 262–284. https://doi.org/10.1080/09541440340000213

Kolan, L., Leikin, M., & Zwitserlood, P. (2011). Morphological processing and lexical access in speech production in Hebrew: Evidence from picture–word interference. *Journal of Memory and Language*, *65*(3), 286–298. https://doi.org/10.1016/j.jml.2011.06.004

Koriat, A. (1984). Reading without vowels: Lexical access in Hebrew. In H. Bouma, & D. G. Bouwhuis (Eds.), *Attention and Performance X: Control of language processes* (pp. 227–242). Lawrence Erlbaum Associates.

Kretzschmar F., Schlesewsky M., Staub A. (2015). Dissociating word frequency and predictability effects in reading: Evidence from coregistration of eye movements and EEG. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *41*(6), 1648–1662. https://doi.org/10.1037/xlm0000128

Kuperman, V., & Van Dyke, J. A. (2013). Reassessing word frequency as a determinant of word recognition for skilled and unskilled readers. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 802–823. https://doi.org/10.1037/a0030859

Lee, O., Quinn, H., & Valdés, G. (2013). Science and language for English language learners in relation to Next Generation Science Standards and with implications for Common Core State Standards for English language arts and mathematics. *Educational Researcher*, *42*(4), 223–233. https://doi.org/10.3102/0013189X13480524

Lee, O., & Fradd, S. H. (1998). Science for all, including students from non-English-language backgrounds. *Educational Researcher*, *27*(4), 12–21. https://doi.org/10.3102/0013189X027004012

Lemhöfer, K., Dijkstra, T., Schriefers, H., Baayen, R. H., Grainger, J., & Zwitserlood, P. (2008). Native language influences on word recognition in a second language: A megastudy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 12–31. https://doi.org/10.1037/0278-7393.34.1.12

Li, X., Bicknell, K., Liu, P., Wei, W., & Rayner, K. (2014). Reading is fundamentally similar across disparate writing systems: A systematic characterization of how words and characters influence eye movements in Chinese reading. *Journal of Experimental Psychology: General*, *143*(2), 895–913. https://doi.org/10.1037/a0033580

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(2), 381–390. https://doi.org/10.1037/a0014875

Liu, P., Li, X., & Han, B. (2015). Additive effects of stimulus quality and word frequency on eye movements during Chinese reading. *Reading and Writing*, *28*(2), 199–215. https://doi.org/10.1007/s11145-014-9521-4

Liu, Y., Reichle, E. D., & Li, X. (2016). The effect of word frequency and parafoveal preview on saccade length during the reading of Chinese. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(7), 1008–1025. https://doi.org/10.1037/xhp0000190

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods*, *49*(4), 1494–1502. https://doi.org/10.3758/s13428-016-0809-y

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

Martin, C. D., Thierry, G., Kuipers, J. R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, *69*(4), 574–588.

Miellet, S., Sparrow, L., & Sereno, S. C. (2007). Word frequency and predictability effects in reading French: An evaluation of the E-Z Reader model. *Psychonomic Bulletin & Review*, *14*(4), 762–769. https://doi.org/10.3758/BF03196834

Mishra, R. (2019). A few suggestions on broadening the cross-linguistic relevance of the Multilink model. *Bilingualism: Language and Cognition*, *22*(4), 691–692. https://doi.org/10.1017/S1366728918000834

Miyake, A. & Friedman, N. P. (1998). Individual differences in second language proficiency: The role of working memory. In A. F. Healy & L. E. Bourne Jr. (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339–364). Erlbaum.

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

Monsell, S. (1991). The nature and locus of word frequency effects in reading. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 148–197). Erlbaum.

Mor, B., & Prior, A. (2020). Individual differences in L2 frequency effects in different script bilinguals. *International Journal of Bilingualism*, *24*(4), 672–690. https://doi.org/10.1177/1367006919876356

Morris, R. K. (2006). Lexical processing and sentence context effects. In M. J. Traxler In M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 277–401). Academic Press. https://doi.org/10.1016/B978-012369374-7/50011-0

Morton, J. (1970). A functional model of human memory. In D. A. Norman (Ed.), *Models of human memory* (pp. 203–254). Academic Press. https://doi.org/10.1016/B978-0-12-521350-9.50012-7

New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review*, *13*(1), 45–52. https://doi.org/10.3758/BF03193811

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*, *98*(3), 554–566. https://doi.org/10.1037/0022-0663.98.3.554

Payne, B. R., Gao, X., Noh, S. R., Anderson, C. J., & Stine–Morrow, E. A. (2012). The effects of print exposure on sentence processing and memory in older adults: Evidence for efficiency and reserve. *Aging, Neuropsychology, and Cognition*, *19*(1–2), 122–149.

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

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

Perfetti, C., & 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

Pivneva, I., Mercier, J., & Titone, D. (2014). Executive control modulates cross-language lexical activation during L2 reading: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 787–796. https://doi.org/10.1037/a0035583

Pollatsek, A., Bolozky, S., Well, A. D., & Rayner, K. (1981). Asymmetries in the perceptual span for Israeli readers. *Brain and Language*, *14*(1), 174–180. https://doi.org/10.1016/0093-934X(81)90073-0

Prior, A., Zeltsman-Kulick, R., & Katzir, T. (2020). Adolescent word reading in English as a foreign language. *Journal of Research in Reading*, *43*(1), 116–139. https://doi.org/10.1111/1467-9817.12293

Prior, A., & Beznos, M. (2009). *Hebrew version of the Language Experience and Proficiency Questionnaire*. https://www.iris-database.org/iris/app/home/detail?id=york%3a822288&ref=search

Prior, A., & Markus, E. (2014). Morphological activation in sentence context: When the root prevails over the meaning. *Language, Cognition and Neuroscience*, *29*(9), 1180–1188. https://doi.org/10.1080/23273798.2014.920511

Qian, D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review*, *56*(2), 282–308. https://doi.org/10.3138/cmlr.56.2.282

R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Raney, G. E., Campbell, S. J., & Bovee, J. C. (2014). Using eye movements to evaluate the cognitive processes involved in text comprehension. *Journal of Visualized Experiments*, (83), e50780. https://doi.org/10.3791/50780

Rau, A. K., Moll, K., Snowling, M. J., & Landerl, K. (2015). Effects of orthographic consistency on eye movement behavior: German and

English children and adults process the same words differently. *Journal of Experimental Child Psychology*, *130*, 92–105. https://doi.org/10.1016/j.jecp.2014.09.012

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *62*(8), 1457–1506. https://doi.org/10.1080/17470210902816461

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(4), 720–732. https://doi.org/10.1037/0096-1523.30.4.720

Rayner, K., Li, X., Juhasz, B. J., & Yan, G. (2005). The effect of word predictability on the eye movements of Chinese readers. *Psychonomic Bulletin & Review*, *12*(6), 1089–1093. https://doi.org/10.3758/BF03206448

Rayner, K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(2), 514–528. https://doi.org/10.1037/a0020990

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., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, *3*(4), 504–509. https://doi.org/10.3758/BF03214555

Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Human Perception and Performance*, *3*(1), 1–12. https://doi.org/10.1037/0096-1523.3.1.1

Schepens, J., Dijkstra, T., & Grootjen, F. (2012). Distributions of cognates in Europe as based on Levenshtein distance. *Bilingualism: Language and Cognition*, *15*(1), 157–166. https://doi.org/10.1017/S1366728910000623

Shipley, W. C. (1946). *Institute of living scale*. Western Psychological Services.

Shohamy, E. (2014). The weight of English in global perspective: The role of English in Israel. *Review of Research in Education*, *38*(1), 273–289. https://doi.org/10.3102/0091732X13509773

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*(4), 360–407. https://doi.org/10.1598/RRQ.21.4.1

Staub, A. (2011). The effect of lexical predictability on distributions of eye fixation durations. *Psychonomic Bulletin & Review*, *18*(2), 371–376. https://doi.org/10.3758/s13423-010-0046-9

Staub, A. (2015). The effect of lexical predictability on eye movements in reading: Critical review and theoretical interpretation. *Language and Linguistics Compass*, *9*(8), 311–327. https://doi.org/10.1111/lnc3.12151

Tolentino, L. C., & Tokowicz, N. (2011). Across languages, space, and time: A review of the role of cross-language similarity in L2 (morpho) syntactic processing as revealed by fMRI and ERP methods. *Studies in Second Language Acquisition*, *33*(1), 91–125. https://doi.org/10.1017/S0272263110000549

Tomoschuk, B., Ferreira, V. S., & Gollan, T. H. (2019). When a seven is not a seven: Self-ratings of bilingual language proficiency differ between and within language populations. *Bilingualism: Language and Cognition*, *22*(3), 516–536. https://doi.org/10.1017/S1366728918000421

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *TOWRE: Test of Word Reading Efficiency – Examiner's manual*. PRO-ED.

Van Heuven, W., & Wen, Y. (2019). The need for a universal computational model of bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, *22*(4), 695–696. https://doi.org/10.1017/S1366728918000688

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., Donahue, J., & Garon, T. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. *Developmental Psychology*, *33*(3), 468–479. https://doi.org/10.1037/0012-1649.33.3.468

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, *143*(5), 2020–2045. https://doi.org/10.1037/xge0000014

Whitford, V., & Titone, D. (2012). Second-language experience modulates first- and second-language word frequency effects: Evidence from eye movement measures of natural paragraph reading. *Psychonomic Bulletin & Review*, *19*(1), 73–80. https://doi.org/10.3758/s13423-011-0179-5

Whitford, V., & Titone, D. (2014). The effects of reading comprehension and launch site on frequency-predictability interactions during paragraph reading. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *67*(6), 1151–1165. https://doi.org/10.1080/17470218.2013.848216

Whitford, V., & Titone, D. (2017). The effects of word frequency and word predictability during first- and second-language paragraph reading in bilingual older and younger adults. *Psychology and Aging*, *32*(2), 158–177. https://doi.org/10.1037/pag0000151

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer.

Zachary, R. A. (1991). *The manual of the Shipley Institute of Living Scale*. Western Psychological Services.

(*Appendix follows*)

Appendix

Table A1
Correlation Matrix of the Predictor Variables

Variable	2.	3.	4.
1. L2 proficiency	.694**	.055	.093
2. L1 proficiency		−.079	.140
3. L1 sensitivity to frequency			−.038
4. L1 sensitivity to predictability			

Note. Following Bonferroni correction for multiple correlations, significance level was set at .008.
** $p < .01$.

Table A2
Correlation Matrix of Predictor Variables, With Separate Values for Participants' Vocabulary Knowledge and Reading Fluency

Variable	2.	3.	4.	5.	6.
1. L2 vocabulary knowledge (Shipley)	.587**	.556**	.32	.099	.181
2. L2 reading fluency (TOWRE)		.521**	.625**	−.001	−.016
3. L1 vocabulary knowledge (Shipley)			.334	−.109	.164
4. L1 reading fluency (TOWRE)				−.02	.65
5. L1 sensitivity to frequency					−.038
6. L1 sensitivity to predictability					

Note. TOWRE = test of Word Reading Efficiency. Following Bonferroni correction for multiple correlations, significance level was set at .003.
** $p < .01$.

Received May 19, 2020
Revision received April 18, 2021
Accepted May 9, 2021 ■