When sentence meaning biases another language: an eye-tracking investigation of cross-language activation during second language reading

Karla Tarin[1,2] ⬥, Esteban Heranadez-Rivera[1,2], Antonio Iniesta[1,2], Pauline Palma[1] ⬥, Veronica Whitford[3] and Debra Titone[1,2] ⬥

[1]Department of Psychology, McGill University, Montréal, QC, Canada; [2]Department of Psychology, Centre for Research on Brain, Language, & Music, McGill University, Montreal, Quebec, Canada and [3]Department of Psychology, University of New Brunswick, Fredericton, NB, Canada

![CAMBRIDGE UNIVERSITY PRESS]

Abstract

Bilingual adults use semantic context to manage cross-language activation while reading. An open question is how lexical, contextual and individual differences simultaneously constrain this process. We used eye-tracking to investigate how 83 French–English bilinguals read L2-English sentences containing interlingual homographs (*chat*) and control words (*pact*). Between subjects, sentences biased target language or non-target language meanings (English = conversation; French = feline). Both conditions contained unbiased control sentences. We examined the impact of word- and participant-level factors (cross-language frequency and L2 age of acquisition/AoA and reading entropy, respectively). There were three key results: (1) L2 readers showed global homograph interference in late-stage reading (total reading times) when English sentence contexts biased non-target French homograph meanings; (2) interference increased as homographs' non-target language frequency increased and L2 AoA decreased; (3) increased reading entropy globally facilitated early-stage reading (gaze durations) in the non-target language bias condition. Thus, cross-language activation during L2 reading is constrained by multiple factors.

Highlights

- Second language (L2) readers co-activate languages as a function of context
- We tested if context can both reduce and enhance co-activation using eye-tracking
- English sentences biasing French versus English meanings led to more co-activation
- Co-activation varied with the relative frequency in English/French of target words
- Co-activation varied with the bilingual experience of L2 readers

Bilingualism is often conceptualized as the ability to switch between two separate language systems. However, bilinguals' languages are intricately connected and frequently interact, even in monolingual contexts (Bergmann et al., 2015; Giezen & Emmorey, 2016; reviewed in Bailey et al., 2024). Language co-activation is particularly evident during reading, which requires rapid access to lexical representations from semantic memory (Duyck et al., 2007; Libben & Titone, 2009; Palma & Titone, 2020). This phenomenon is often studied through interlingual homographs – words that share orthographic form but differ in meaning across languages (e.g., **chat** = "informal conversation" in English and "cat" in French). When presented in isolation, homographs often elicit processing delays and higher error rates compared to language-unique words matched for length and frequency (i.e., interlingual homograph interference, e.g., Beauvillain & Grainger, 1987; Gerard & Scarborough, 1989; von Studnitz & Green, 2002). Models such as the Bilingual Interactive Activation Plus (BIA+; Dijkstra & van Heuven, 2002, 2013; Jared & Szucs, 2002) and Multilink (Dijkstra et al., 2019), suggest this interference stems from competition between a word's multiple meanings due to shared orthographic and phonological features (e.g., De Groot & Nas, 1991; Dijkstra et al., 1998, 2010).

Notably, bilingual reading demands both within- and cross-language lexical ambiguity resolution (e.g., interlingual homographs, **chat**). To account for this, Degani and Tokowicz (2010) extended unilingual approaches such as the Reordered Access Model (Duffy et al., 1988) into a Three-Factor Framework, which emphasizes how context can bias interlingual homograph interpretation toward target language meanings (**chat** in an English sentence about conversation) or non-target language meanings (**chat** in an English sentence about veterinarians). The idea of contexts semantically biasing a *non-target language* meaning is particularly interesting, as it should produce *greater* rather than lesser cross-language activation for specific sentences or globally for an entire task session. This idea recalls Grosjean's Language Mode Hypothesis (2001),

according to which, bilinguals operate along a continuum of activation, from a "monolingual mode" (with the non-target language suppressed) to a "bilingual mode" (with both languages active). When context biases target language meanings, bilinguals may shift toward a monolingual mode, reducing interference. Conversely, when context biases non-target language meanings, bilinguals may shift toward a "bilingual mode," amplifying interlingual homograph interference.

In addition to its emphasis on context, another key aspect of the Three-Factor Framework is its integration of word-level and participant-level factors during bilingual reading. Within this view, resolving cross-language ambiguity in context is shaped by word-level properties (e.g., cross-language status, relative frequency) as well as participant-level factors (e.g., L2 AOA, language experiences). Of relevance here, while these factors have been studied in isolation, their simultaneous, interactive effects are less understood. As will be seen, the current study addresses this gap by examining how these factors interact to shape bilingual reading. In building to this study, we first review relevant psycholinguistic research on cross-language competition during sentence reading, focusing on how context can create bias at both the target and non-target language levels, the role of word-level properties and readers' unique language experiences.

1. Cross-language activation: semantic bias toward target language and non-target language meanings

Many studies have investigated how semantic context constrains interlingual homograph comprehension (reviewed in Lauro & Schwartz, 2017). Schwartz and Kroll (2006) first investigated this with Spanish–English bilingual adults of varying English proficiency. Participants read English sentences one word at a time, using rapid serial visual presentation (RSVP). Interlingual homographs (e.g., *fin* meaning "end" in Spanish) appeared in a different-colored font, and participants named the colored word aloud as quickly and accurately as possible. Critically, sentences preceding the appearance of the interlingual homograph were either semantically biased (e.g., "*From the beach we could see the shark's fin pass through the water*") or unbiased (e.g., "*We felt a bit nervous when we saw the fin of the shark in the distance*"). Results showed that cross-language activation occurred for all bilingual readers in that they made more naming errors for interlingual homographs compared to control words when the sentence context was neutral. However, when the same words appeared in highly constraining sentences, bilinguals with lower English proficiency showed reduced interlingual homograph interference, suggesting that the sentence context helped them to inhibit contextually inappropriate cross-language meanings.

Recent studies have employed more naturalistic methodology, such as eye-tracking (e.g., Libben & Titone, 2009; Pivneva et al., 2011; Titone et al., 2011; reviewed in Whitford et al., 2016), enabling real-time examination of bilingual reading (reviewed in Paterson et al., 2007; Rayner, 1998). In this method, measures such as first fixation duration (FFD; duration of the initial fixation on a word), gaze duration (GD; sum of all fixation durations on a word during its first-pass reading), and word skipping (probability of not fixating a word during its first-pass reading) reflect early-stage lexical access. In contrast, measures such as go-past time (GPT; sum of all fixation durations on a word during first-pass reading and any refixation durations on earlier occurring words, until a saccade is made to a later occurring word) and total reading time (TRT; grand sum of all fixation durations on a word) reflect later stage processes like semantic integration and ambiguity resolution.

For example, Libben and Titone (2009) used eye-tracking to investigate French–English bilingual adults reading L2 (English) sentences containing interlingual homographs (e.g., *chat*) or matched control words (e.g., *pact*) in high-semantic versus low-semantic constraint sentence contexts. Because their interest was to examine how context dampened cross-language activation, they selected interlingual homographs for which the English word form was less frequent than the French word form. This is to maximize the likelihood of cross-language activation when people read English sentences. Similar to Schwartz and Kroll (2006), during early-stage reading (FFD, GD, skipping), bilinguals showed robust interlingual homograph interference regardless of contextual constraint. However, during late-stage reading (GPT, TRT), bilinguals showed reduced interlingual homograph interference for sentences with high-semantic but not low-semantic constraints. Subsequent work using the same materials found evidence of interlingual homograph interference during first language (L1) reading, though to a lesser extent (Titone et al., 2011), and that greater executive control capacity patterned with less cross-language activation for L2 readers (see also, Pivneva et al., 2014).

Semantic context can also bias interlingual homographs toward non-target language meanings. For example, the English sentence "*When the veterinarian noticed the sickly animal, she had a thorough chat with the owners*" biases the French meaning of the word *chat* (i.e., cat). Jouravlev and Jared (2014) addressed this issue by examining contexts biasing the non-target language meaning. In an event-related potential (ERP) study, Russian–English bilingual adults read L2 (English) sentences containing Russian/English interlingual homographs (e.g., **рот**/mouth in Russian) and the English translations of the interlingual homographs' Russian meaning (e.g., mouth). The semantic context always biased the non-target L1 (Russian) meaning of the interlingual homograph (e.g., "*To see Tom's throat, the doctor asked Tom to open his рот /mouth*"). Results showed that bilinguals exhibited comparable N400 amplitudes for interlingual homographs and semantically plausible English words, suggesting they effectively used the context to integrate the non-target L1 (Russian) meaning. However, the study design did not clarify whether a prior context biasing the non-target language meaning could interfere with comprehension if the target language meaning was intended (e.g., a prior context mentioning mouth, while the kitchen-related meaning of **рот** was intended).

Hoversten and Traxler (2016) examined contexts that biased both target and non-target language meanings using eye-tracking. Spanish–English bilingual and English monolingual adults read English sentences containing Spanish–English interlingual homographs (e.g., *pie* meaning "foot" in Spanish), where the context could bias either the target English meaning (e.g., "*While eating dessert, the diner crushed his pie accidentally with his elbow*") or the non-target Spanish meaning (e.g., "*While carrying bricks, the worker crushed his pie accidentally with the load*"). The authors observed no interlingual homograph interference during early-stage reading (GD), regardless of language bias condition or participant group. Thus, bilinguals did not automatically activate the interlingual homographs' L1 (Spanish) meaning. However, the authors observed group differences during late-stage reading (TRT), with bilinguals experiencing less overall slowing than monolinguals integrating interlingual homographs into incoherent conditions. Thus, bilinguals spent less time reading English sentences that required the integration of interlingual homographs' L1 (Spanish) meaning.

Collectively, these studies suggest that bilingual adult readers can adeptly use the semantic bias of sentences to modulate cross-language activation of interlingual homograph meanings. We now turn to factors that potentially modulate this process, such as word frequency and language experience.

2. Cross-language activation: effects of word frequency and language experience

According to the Three-Factor Framework, word-level properties such as word frequency should also influence language co-activation and how context may be used to resolve ambiguity (Degani & Tokowicz, 2010). Word frequency has two key dimensions: a word's frequency *within* a given language and its relative frequency *across* multiple languages. Lower-frequency words (e.g., *endemic, outbreak*) are generally processed more slowly and less accurately than higher-frequency words (e.g., *people, house*), especially in the L2 (reviewed in Rayner, 1998; Rayner et al., 2004; Whitford et al., 2016). Importantly, relative frequency also matters. For instance, although **chat** is common in English, French–English bilinguals might default to its French meaning ("cat") when the word is presented without context, due to its higher frequency in French.

A review of past work suggests that cross-language word frequency varies across studies. For instance, Hoversten and Traxler (2016) used interlingual homographs with balanced word frequencies across both languages (English, Spanish), which may have limited early-stage non-target activation (moreover, their Spanish–English bilinguals may have been reverse dominant, being more efficient L2 versus L1 readers). In contrast, Jouravlev and Jared (2014) examined words more frequent in the target language (English), possibly dampening cross-language activation effects. Finally, Libben and Titone (2009) deliberately selected interlingual homographs for which the target language (English) meaning was lower in frequency compared to the non-target (French) meaning (e.g., **manger** is less frequent in English "animal feeding trough" than in French "to eat"). Their goal was to determine whether a biased semantic context would reduce interlingual homograph interference; thus, their materials, by design, maximized such interference in semantically neutral conditions. These differences underscore the need for more systematic investigation of how frequency – within and across languages – modulates cross-language activation.

The Three-Factor Framework (Degani & Tokowicz, 2010) suggests that bilingual experience should also impact reading behavior. For example, readers who have very early L2 AoA or communicate in highly bilingual social contexts may be more sensitive to language context than those who read in a less dominant or less practiced L2. The BIA+ (e.g., Dijkstra & van Heuven, 2002, 2013; Jared & Szucs, 2002) and Multilink (Dijkstra et al., 2019) models emphasize the importance of historical factors such as L2 AoA in modulating language co-activation. Specifically, L2 AoA influences the degree and nature of cross-language activation, as early L2 acquisition is linked to more integrated lexical and semantic representations, leading to higher cross-language activation (reviewed in Palma & Titone, 2020; Titone et al., 2011; Whitford et al., 2016). Consistent with this, Titone et al. (2011) found greater interlingual homograph interference during L1 reading for bilinguals who acquired their L2 early in life. This coheres with other studies showing how individual differences in linguistic background (Schwartz & Kroll, 2006) and cognitive capacity (Pivneva et al., 2014) modulate semantic bias effects in interlingual homograph processing.

In addition to historical bilingual experiences, bilingual experience can vary multidimensionally (e.g., Baum & Titone, 2014; DeLuca et al., 2019; Gullifer and Titone, 2020; Gullifer et al., 2021; Luk & Bialystok, 2013; Titone & Tiv, 2023). For example, bilinguals can vary in how they distribute their language use even when holding L2 AoA constant. Some bilinguals compartmentalize the use of their L1 and L2 across specific social settings and/or times of the day. Other bilinguals adopt a more integrated approach, using both languages throughout the day and across social contexts (e.g., Gullifer & Titone, 2020; Luk & Bialystok, 2013; Titone & Tiv, 2023; Tiv et al., 2021). Such differences can be captured by measures such as *language entropy* (Gullifer & Titone, 2020; Kalamala et al., 2022; Li et al., 2021; see Iniesta et al., 2025, for evidence validating this measure as capturing socially realistic language use). Lower language entropy (pertaining to reading specifically) indicates predictable use of multiple languages (i.e., language compartmentalization), while higher entropy suggests less predictable, more fluid language use (i.e., language integration). Importantly, language entropy predicts language proficiency (Gullifer & Titone, 2020), executive function (Gullifer & Titone, 2021a; Li et al., 2021) and cognitive adaptations related to language fluency (Gullifer & Titone, 2021b). However, few studies have examined how entropy affects cross-language activation during bilingual reading.

3. The current study

As proposed by the Three-Factor Framework (Degani & Tokowicz, 2010) and supported by prior work reviewed above, lexical, contextual, and individual factors all modulate interlingual homograph processing during bilingual reading. However, little is known about how these factors simultaneously interact to impact eye movement measures of reading among bilinguals.

To address this gap, we examined how French–English bilinguals process interlingual homographs in L2 (English) sentences, with sentence contexts biasing either the target (English) or non-target (French) meaning. We also investigated the roles of word-level properties (e.g., cross-language word frequency) and participant-level factors (e.g., L2 AoA, reading-related entropy). Our research questions were threefold. First, how do French–English bilinguals process interlingual homographs embedded in L2 (English) sentences that bias the target language meanings? For example, when an English sentence biases the meaning of **chat** as "conversation" versus non-target language meaning "cat." This question examines whether semantic context influences the readers' language mode. That is, shifting from a monolingual mode (where one language is highly active and the other is suppressed) to a bilingual mode (where both languages are active to varying degrees) and vice versa, as shown by differential cross-language activation during reading (Grosjean, 1998, 2001). Second, how do word-level properties, such as cross-language word frequency, modulate interlingual homograph processing? For example, **chat** is more frequent in French versus English. Third, how do individual differences, such as L2 AoA and current reading habits (i.e., reading entropy), interact with the above?

To address the first question, L1-dominant French–English bilinguals read L2 (English) sentences containing interlingual homographs or matched language-unique control words. Contextual bias for each sentence was manipulated in two ways: first, each sentence was manipulated for *semantic* bias (biased versus neutral context toward the target word). Within the biased condition, we further manipulated the context to constrain interpretation toward

either the target language (L2-English) or the non-target language (L1-French) meaning of the interlingual homograph. Participants were exposed to only one of the two language-biasing conditions to prevent shifts in language mode due to the varying contextual cues, ensuring that language mode remained consistent within each group.

We hypothesized that semantic bias and language bias would jointly affect interlingual homograph interference, with larger effects in late-stage (TRT) reading compared to early-stage (GD), as TRT reflects semantic integration (Hoversten & Traxler, 2016; Libben & Titone, 2009). Specifically, we expected sentence context to either reduce or amplify interlingual homograph interference, depending on whether it biased target or non-target language meanings. We predicted heightened interference for non-target language bias contexts as it would lead to pre-selection of the unintended L1 (French) meaning, enabling co-activation of both languages and eliciting a bilingual mode. In contrast, when sentences biased target language (L2-English) meanings of the interlingual homographs, we expected reduced interference, as the context would facilitate the pre-selection of the intended L2 (English) meaning, suppressing the non-target L1 and eliciting an L2-only mode.

To address the second question concerning the impact of word-level properties, we included interlingual homographs with a broad range of cross-language word frequencies (e.g., **"*note*"**, with a higher English frequency score compared to French versus **"*chat*"** with a lower English frequency score compared to French), rather than only those items that had higher form frequency in the non-target language (e.g., Libben & Titone, 2009). We hypothesized that interlingual homographs with higher non-target (L1-French) word frequencies would produce longer reading times, as a higher frequency in French could make the non-target L1 meaning more difficult to inhibit during L2 reading. Additionally, we predicted this effect would be especially evident when the sentence context biased the L2 (English) meaning, as activation of the intended L2 meaning would conflict with the strong, highly frequent L1 (French) meaning, thereby increasing processing difficulty.

To address the third question, we examined whether individual differences in historical language use and current exposure modulated L2 (English) interlingual homograph processing. Our first measure was L2 (English) AoA, as it indicates people's historical language use, reflecting when a person first began acquiring their L2 (de Leeuw & Bogulski, 2015; Gullifer et al., 2018; see Elsherif et al., 2023 for a review). As a measure of language history, L2 AoA offers a window into long-term language experience but does not capture present-day language patterns. To account for this, we examined *reading entropy*, a measure of *current* reading habits across languages.

We hypothesized that bilingual adults with earlier L2 AoAs would exhibit shorter overall reading times and more pronounced interlingual homograph interference across semantic conditions, evidenced by greater differences in fixation durations between interlingual homographs and language-unique control words. This would suggest more integrated lexical and semantic representations across languages, resulting in higher levels of language co-activation (Palma & Titone, 2020; Whitford et al., 2016; see also Titone et al., 2011). Additionally, we hypothesized that bilingual adults who frequently read in two languages, as indicated by higher reading entropy, would exhibit reduced interlingual homograph interference when context biased the non-target language (L1-French) meanings of interlingual homographs. We hypothesized this would arise because higher entropy readers would be more practiced at flexibly and proactively managing

cross-language activation during bilingual reading than lower entropy readers (e.g., Gullifer & Titone, 2021a).

4. Methods

4.1. Participants

We tested 83 L1-dominant French–English bilingual adults (M_{age} = 22 ± 3.87 years). All had normal or corrected-to-normal vision and no self-reported hearing, language, neurological or psychiatric disorders. Participants received either $10/hour or course credit for their time.

Participants were recruited from Montreal, Quebec, a linguistically and culturally diverse city, where bilingualism is common. Most residents navigate both French and English in everyday life across a variety of domains (Kircher, 2014; Lamarre, 2013; Titone & Tiv, 2023). All participants completed a self-report language history questionnaire to confirm French as their L1 and language exposure and proficiency across groups (all *p* values > .05). Participants were randomly assigned to either the target (i.e., English) or non-target language (i.e., French) biasing condition (see Section 4.2). Slight differences in writing proficiency were noted between groups (see Table 1). No differences were found in speaking or reading proficiency. These are addressed in Section 6.

Additionally, we computed a *reading-based language entropy* score for each participant from the language questionnaire data, reflecting current patterns of language use in reading. Reading entropy was calculated as follows: first, the questionnaire asked participants to estimate the percentage of time spent using each language in various contexts, including reading (e.g., "What percentage of the time do you use English/French/other languages to read?") across a variety of domains (e.g., "when at home?", "school/work?", "socially?"). We then used the *languageEntropy* R package (Gullifer & Titone, 2020) to calculate reading entropy. The calculation of reading entropy is determined by the equation $H = -\sum_{i=1}^{n} P_i \log_2(P_i)$. Here, *n* represents the total number of possible languages (in this case, *n* = 3 for English, French, Spanish), and P_i represents the proportion of language usage (PEnglish = 0.9, PFrench = 0.05, PSpanish = 0.05). The sum is multiplied by −1 to yield a positive entropy value. The lower bound of entropy is zero, indicating that one language is used exclusively 100% of the time (functional monolingual). When there are only two possible languages, the upper bound of entropy is one, representing maximal language entropy (50% usage of each language). As the total number of possible languages increases (e.g., trilingual), the upper bound of entropy also increases.

4.2. Procedure

Participants' eye movements were recorded using an EyeLink1000 desktop-mounted system (right-eye recording, 1-kHz sampling rate, spatial resolution of 0.01°, and mean accuracy of 0.25°; SR-Research, Ontario, Canada). Before the task, calibration was performed using a 5-point cross (average fixation error was less than 0.5° of visual angle). A drift correction point was presented before the onset of each sentence to ensure tracking accuracy. A padded headrest minimized head movements.

Experimental sentences were presented in yellow 10-point Monaco font (due to equidistant character spacing) on a black background using EyeTrack software developed at the University of Massachusetts Amherst). Sentences were displayed in a single line (maximum characters = 75, with 3.2 characters subtending 1°

Table 1. Participant's characteristics

	Target language biasing task English (*n* = 42)		Non-target language biasing task French (*n* = 41)		
Sex (female–male) ratio	29:13		30:11		
Age (years)	25		27		
	Range	**Mean (SD)**	**Range**	**Mean (SD)**	***t (p)***
L2 AoA (years)	0–24	9.11 (4.95)	0–13	7.28 (3.96)	1.81 (.074)
Language exposure (current%)					
French	10–100	52.76 (20.21)	15–90	53.94 (17.67)	0.32 (.752)
English	0–90	44.31 (20.54)	10–85	44.11 (17.23)	−0.08 (.933)
Reading entropy	0.00–1.22	0.69 (0.35)	0.00–1.36	0.82	1.73 (.088)
French proficiency					
Reading	6–7	6.93 (0.26)	5–7	6.87 (0.40)	−0.72 (.476)
Writing	5–7	6.93 (0.33)	4–7	6.62 (0.83)	−2.12 (.037)
Speaking	6–7	6.95 (0.21)	5–7	6.90 (0.38)	−0.77 (.441)
Fluency	6–7	6.95 (0.21)	5–7	6.80 (0.52)	−1.73 (.099)
Overall Fr. proficiency	6–7	6.92 (0.26)	5–7	6.77 (0.58)	−1.57 (.121)
Reading	3–7	5.64 (1.06)	3–7	6.03 (1.14)	1.45 (.151)
Writing	3–7	4.97 (1.01)	2–7	5.73 (1.01)	2.34 (.022)
Speaking	3–7	5.30 (1.14)	3–7	5.84 (1.18)	1.63 (.108)
Fluency	3–7	4.98 (1.12)	2–7	5.56 (1.57)	1.85 (.084)
Overall Eng. proficiency	3–7	5.26 (0.94)	2–7	5.66 (1.30)	1.60 (.114)

of visual angle) against a black background using a 21-inch View-Sonic CRT monitor (positioned 57 cm from participants). Participants read a total of 72 sentences (36 containing interlingual homographs and 36 containing control words). The onset of each trial was initiated by fixating on a yellow, gaze-contingent box located before the first word of each sentence.

Participants were instructed to press a button on a controller after reading each sentence. To ensure attention throughout the task, a simple yes/no comprehension question appeared on 25% of the experimental trials. Following the sentence reading task, participants completed a language history questionnaire. The study was approved by McGill University's Research Ethics Board (112-0909).

4.3. Materials

Target words were 36 French–English interlingual homographs and their corresponding language-unique control words. English and French word frequencies (occurrences per million words) were obtained from CELEX (Baayen et al., 1995) and Lexique (New et al., 2001), respectively. English control words were generated using WordGen (Duyck et al., 2004) and matched the interlingual homographs for length, neighborhood density and frequency (see Table 2). The word frequency difference score was calculated by subtracting English word frequency from French frequency for each target word. Negative values reflect lower English frequency, positive values reflect higher English frequency and zero reflects balanced frequency (M = −49.17, SD = 112.61, range = −368.44 to 156.42).

Half of the participants (*n* = 42) were assigned to the target language bias condition (i.e., bias toward English homograph meanings), while the other half (*n* = 41) were assigned to the non-target

Table 2. Word length, neighborhood density and frequency (occurrences per million words) for all interlingual homographs and control words

	Length	Neighborhood density	English frequency	French frequency
Interlingual homographs	4.94 (1.49)	7.33 (5.26)	38.39 (48.40)	50.53 (57.97)
Control words	4.92 (1.50)	6.86 (5.00)	35.47 (40.65)	N/A

Note: Means and standard deviations are included in parentheses.

language bias condition (i.e., bias toward French homograph meanings). Thus, a separate set of experimental sentences was created for each condition. All sentences contained two coordinated clauses, with target words (interlingual homographs and control words) presented in a non-sentence final position. Target words were preceded by a word of at least five characters to minimize skipping, kept identical across conditions.

For participants in the target language bias condition, sentences either biased the English meaning of the target word or biased no specific meaning ("neutral"). For instance, a biased sentence would read: "*When Jane finished talking to the real-estate agent, she had a thorough **chat** with the owner.*" Here, "talking" creates a semantic bias toward the English meaning of ***chat*** ("informal conversation"). A neutral sentence was also presented and would read: "*When the teacher found the missing wallet, she had a thorough **chat** with the owner.*" Here, the sentence context preceding the interlingual homograph does not bias any meaning of the target word.

In contrast, in the non-target language (French) bias condition, sentences either biased the French meaning of the interlingual

Table 3. Sample stimuli for experimental conditions

Semantically biased *(within subjects)*		Semantically unbiased *(within subjects)*
target word = INTERLINGUAL HOMOGRAPH – ILH		
Target language bias *(between subjects)*	**Non-target language bias** *(between subjects)*	
When Jane finished talking to the real-estate agent, she had a thorough **chat** with the owner.	When the veterinarian noticed the sickly animal, she had a thorough **chat** with the owner.	When the teacher found the missing wallet, she had a thorough **chat** with the owner.
target word = LANGUAGE-UNIQUE CONTROL WORDS		
When the feuding nations ceased their fighting, they made a formal **pact** with each other.		When the women had finally calmed down, they made a formal **pact** with each other.

homograph or were neutral. An example of the biasing condition would be: "*When the veterinarian noticed the sickly animal, she had a thorough **chat** with the owner.*" Here, "veterinarian" and "animal" create a semantic bias toward the French meaning of **chat** ("cat"), although the English meaning ("informal conversation") is intended. Again, a neutral sentence was presented. This could read: "*When the teacher found the missing wallet, she had a thorough **chat** with the owner.*" Here, the context does not bias any specific meaning.

Finally, for each sentence containing interlingual homographs, a set of two sentences containing a matched control word was created. One sentence included the control word in a semantically biased context, and the other sentence was semantically neutral (see Table 3 for a summary of experimental conditions and sample stimuli). This ensured that each participant saw a control sentence matched for both context and target word.

5. Results

5.1. Sentence comprehension

Participants' sentence comprehension scores were high ($M = 96.25\%$, $SD = 7.40$), indicating good task attention.

5.2. Sentence Eye movement data

The region of interest was defined as the target word of each sentence (i.e., interlingual homograph or language-unique control word). We analyzed two eye-tracking measures reflecting early and late reading processes.[1] For early stages, we extracted gaze duration (GD) due to its substantial overlap with other typical early measures such as first fixation duration (FFD). Indeed, the two variables were strongly correlated (r (81) = .822, $p < .001$). For late-stage reading, we extracted total reading time (TRT). We applied a lower cutoff of

80 milliseconds to all fixations, and both GD and TRT were log-transformed to correct for normality violations (skewness).

Data were analyzed using linear mixed-effects models (Baayen, 2008) in R (version 4.4.1; R Development Core Team, 2025) using the *lme4* package (Bates et al., 2015). First, we checked if our number of observations was suitable for our analysis. Brysbaert and Stevens (2018) recommend "at least 1,600-word observations per condition (e.g., 40 participants, 40 stimuli)." In this study, observations from 83 participants (41 in the non-target language bias condition, 42 in the target language bias condition) and from 72 words (36 interlingual homographs, 36 control words) were included. This resulted in 5,976 observations for target words, with 3,024 observations for the target language bias condition and 2,952 observations for the non-target language bias condition.

Maximal random effects by subjects and by items were included as warranted by the data (Barr, 2013; Bates et al., 2015; Hohenstein, 2017) using the *buildmer* package (Voeten, 2023). This R function automates the process of stepwise model selection through backward elimination, identifying the best model based on all possible predictors and random effects. Maximal random effects structure for our model included a random intercept by participant, a random intercept by item, a correlated random slope for target word by participant and a correlated random slope for target language versus non-target language bias by item. We created plots of predicted data using *ggplot2* package (Version 2.3.0.0), which represent the effects estimated by the linear mixed-effects models we constructed.

We conducted three sets of analyses, across our early (GD) and late (TRT) eye movement reading measures, to address our three research questions. First, how do L2 (English) reading patterns differ when interlingual homographs are embedded in sentential contexts that semantically bias the target language (English) versus non-target language (French) meaning? Second, how does cross-language word frequency (difference score) modulate the above reading patterns? Third, how do individual difference factors, namely, L2 AoA and reading entropy scores, modulate the above reading patterns? Data analytic scripts are available on the Open Science Framework, OSF (https://osf.io/4tg3u/).

5.2.1. Overall L2 interlingual homograph reading: semantic and language bias in early and late reading measures:

We examined whether L2 (English) reading patterns differed when interlingual homographs appeared in contexts semantically biasing the target language (English) versus non-target language (French) meanings. We fit two linear mixed-effects models (one for GD, one for TRT) with fixed effects for: target word (interlingual homograph versus control), semantic bias (bias versus unbiased), target language versus non-target language bias (English versus French bias toward interlingual homograph) and their interactions. Categorical predictors were deviation-coded (control: −.5 versus critical: +.5) so that intercepts reflected the grand mean across all conditions, and the beta coefficients could be interpreted as main effects. A word frequency difference score (continuous) was included as a covariate to control for differences in word processing that may arise from English and French word frequency differences. This variable was centered and scaled to aid in model interpretation.

For our early reading measure (GD), we fit a Linear mixed effect (LME) model with the following structure: gaze duration ~ target word * semantic bias * target versus non-target language bias + scale (cross-language word frequency) + (1 + target word | subject) + (1 + target versus non-target language bias | item). There were no significant main effects of target word ($\beta = .017$,

---

[1]Based on reviewer's suggestion, we replicated the models for GD and TRT using *Regressions Out* as the dependent variable. However, since our area of interest consisted of only one word, very few regressions were captured. The overall proportion of regressions across participants was approximately 0.2% of the total data. This low percentage of data points prevented models from converging.

$SE = .014$, $t = 1.205$, $p = .232$), semantic bias ($\beta = .005$, $SE = .014$, $t = .378$, $p = .795$) or target language versus non-target language bias ($\beta = -.005$, $SE = .037$, $t = -0.128$, $p = .899$). Thus, during early-stage L2 reading (for homographs that varied greatly in terms of relative target versus non-target language frequency), the processing of interlingual homographs and control words was comparable, regardless of semantic and target language versus non-target language bias conditions.

There was, however, a marginally significant interaction between target word and target language versus non-target language bias conditions ($\beta = .054$, $SE = .028$, $t = 1.909$, $p = .060$). As shown in the left panel of Figure 1, this reflected a trend toward slower GDs for interlingual homographs (i.e., interlingual homograph interference) for participants in the non-target language (French) bias condition, but not for participants in the target language (English) bias condition. This effect did not interact with sentential semantic bias, suggesting a global slowing of all interlingual homographs versus control words as a function of target language bias (with the non-target language bias group showing overall greater interlingual homograph interference than the target language bias group).

For our late reading measure (TRT), we fit a model with the following LME structure: total reading time ~ target word * semantic bias * target versus non-target language bias + scale (cross-language word frequency) + (1 + semantic bias| subject) + (1 + target versus non-target language bias | item). There was a significant main effect of target word ($\beta = .116$, $SE = .016$, $t = 7.048$, $p < .001$), where interlingual homographs had slower TRTs than control words. There was a significant interaction between target word and target language versus non-target language bias condition ($\beta = .151$, $SE = .033$, $t = 4.589$, $p < .001$), where TRTs were slower for interlingual homographs for participants in the non-target language (French) bias condition (see Figure 1, right panel). Again, this interaction was a global effect across each block that did not interact with sentential semantic bias.

5.2.2. Interactions with cross-language word frequency scores. Results from early and late reading measures.

We subsequently examined how item differences in cross-language word frequency modulated the results. Two linear mixed-effects models were fitted (one for GD, one for TRT) with the following fixed effects: target word, semantic bias, target language versus non-target language bias condition, cross-language frequency scores and their interactions. As a reminder, cross-language word frequency scores refer to the English word frequency subtracted from the French word frequency for each target word. This variable was centered and scaled in the model. The maximal random effect structure included random intercepts for participants and items. Below, we only describe findings that involved cross-language word frequency.

For our early reading measure (GD), we used a model with the following LME structure: gaze duration ~ target word * semantic bias * target versus non-target language bias * scale (cross-language word frequency) + (1 | subject) + (1 | item). There was no significant main effect or interactions involving the cross-language word frequency scores (all p values > .05).

For our late reading measure (TRT) we used a model with the following LME structure: total reading time ~ target word * semantic bias * target versus non-target language bias* scale (cross-language word frequency) + (1 | subject) + (1 | item). We found a significant interaction between target word and cross-language word frequency ($\beta = -.038$, $SE = .017$, $t = -2.323$, $p = .021$), where interlingual homograph interference was greater when the target word was more frequent in the non-target language (French) than in the target language (English) (see Figure 2, left panel).

Additionally, there was a significant interaction between cross-language word frequency and target language versus non-target language bias ($\beta = .033$, $SE = .017$, $t = 2.027$, $p = .043$). Here, TRTs were slower when the target word was more frequent in the non-target language (French) and occurred in the block where sentences
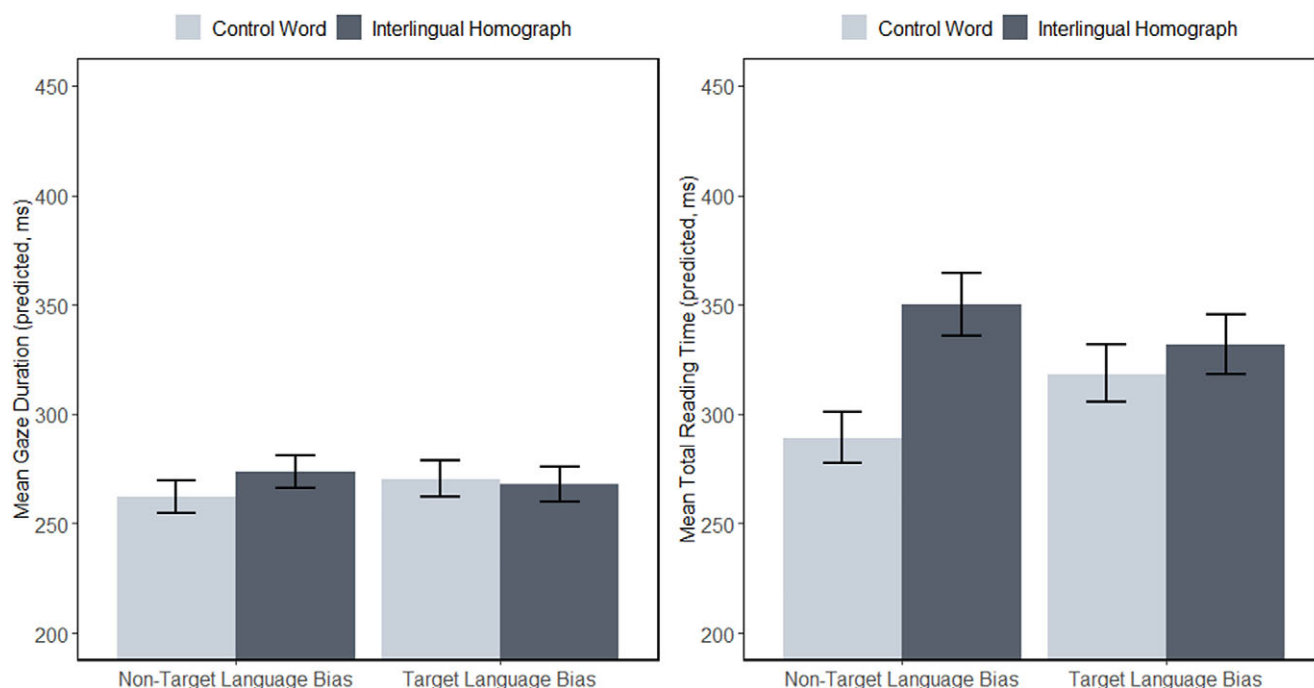

Figure 1. Predicted GDs (left) and TRTs (right) for target words under the target language versus non-target language bias conditions. Error bars reflect ±1 SEM.
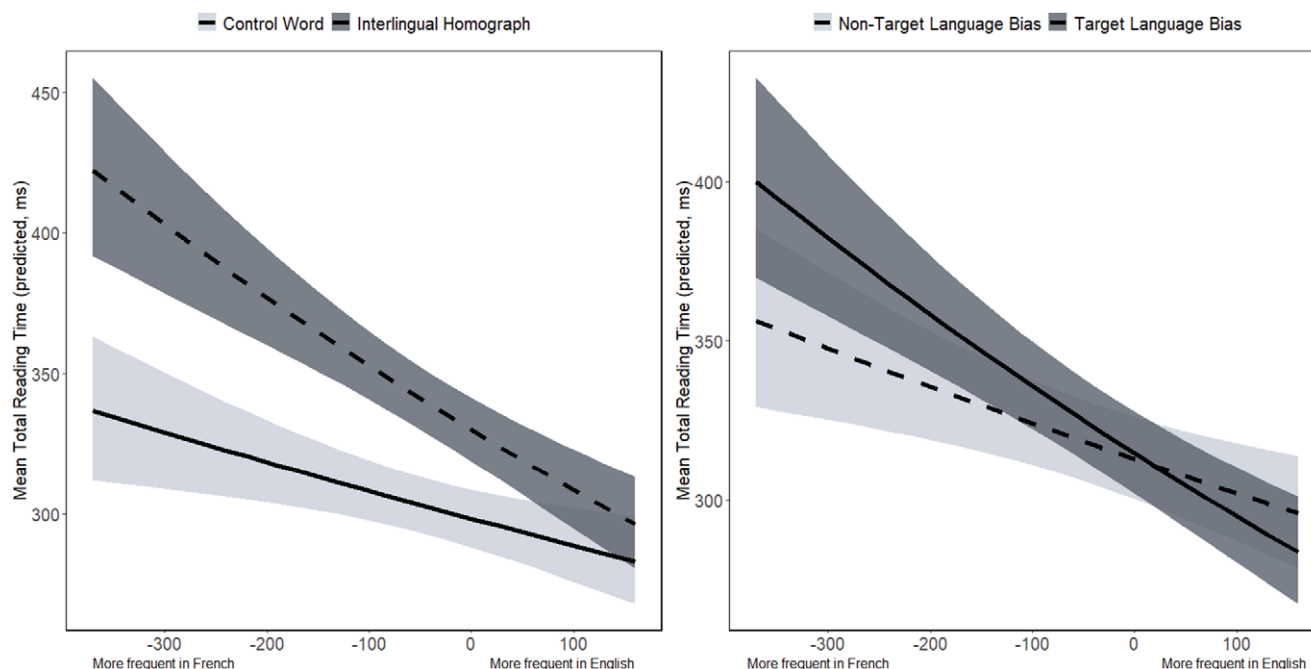
Figure 2. Predicted TRTs for target words (left) and target versus non-target language bias condition (right) as a function of cross-language word frequency. Error bands reflect ±1 SEM.

biased target language (English) meanings only. Of note, this effect did not interact with either word type (i.e., whether the target was an interlingual homograph or control) or sentence type (i.e., whether the target was strongly or moderately semantically biased toward the target word). As such, these results reflect a more general effect, where target language (English) bias and non-target language (French) bias groups processed all words as a function of relative interlingual homograph frequency (see Figure 2, right panel).

5.2.3. Interactions with individual differences in bilingual language experience. Results from early and late reading measures.

Finally, we examined how individual differences in language experience, namely L2 AoA and compartmentalized versus integrated reading habits (i.e., reading entropy), modulated the results. Two linear mixed-effects models were fitted (one for GD, one for TRT) to capture two separate interactions: (1) one involving L2 AoA and (2) another involving reading entropy. For the L2 AoA analysis, fixed effects included target word, semantic bias, target language versus non-target language bias condition, L2 AoA and their interactions. For the reading entropy analysis, fixed effects included target word, semantic bias, target language versus non-target language bias condition, reading entropy and their interactions. Thus, each model assessed the interaction of each individual difference measure while functionally controlling for the interaction of the other individual difference measure. Random effects included random intercepts for participants and items. For brevity, we only report significant findings of the highest order.

For our early reading measure (GD), we used a model with the following LME structure: gaze duration ~ [target word * semantic bias * target versus non-target language bias * scale (L2 AoA)] + [target word* semantic bias * target versus non-target language bias* scale (reading entropy)] + (1 | subject) + (1 |
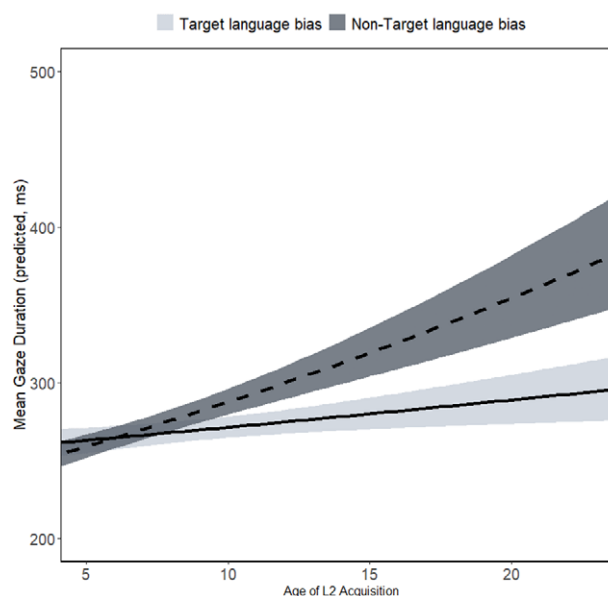

Figure 3. Predicted GDs for target words under target language versus non-target language bias conditions as a function of L2 AoA. Error bands reflect ±1 SEM.

item). We found a significant interaction between target language versus non-target language condition and L2 AoA (β = .067, SE = .031, t = 2.124, p = .036), where later L2 AoAs were associated with slower L2 (English) reading times overall. This global effect was more pronounced for the non-target language bias group (see Figure 3).

Additionally, there was a significant interaction between target language versus non-target language bias and reading entropy (β = −.070, SE = .030, t = −2.260, p = .026). Specifically, lower reading entropy was associated with slower L2 reading times for the
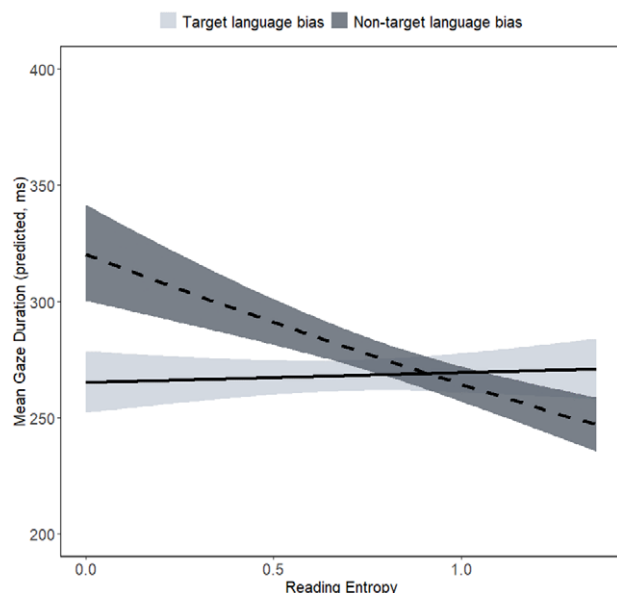
Figure 4. Predicted GDs for target words under target language versus non-target language bias conditions as a function of participants' reading entropy score. Error bands reflect ±1 SEM.

non-target language (French) bias group, while no effect was found in the target language (English) bias group (see Figure 4).

For our late reading measure (TRT), we used a model with the following LME structure: total reading time ~ [target word * semantic bias * target versus non-target language bias * scale (L2 AoA)] + [target word * semantic bias * target versus non-target language bias * scale (reading entropy)] + (1 | subject) + (1 | item). Here, we found a significant interaction between target word and L2 AoA (β = .056, SE = .017, t = −3.182, p = .001), where earlier L2 AoAs were associated with greater interlingual homograph interference (see Figure 5). No significant interactions were found regarding reading entropy in late reading measures.

6. Discussion

Guided by bilingual processing models BIA+ (Dijkstra & van Heuven 2002), the Three-Factor Model (Degani & Tokowicz,

2010) and the Language mode hypothesis (Grosjean 1998, 2001), we investigated how French–English bilinguals process interlingual homographs embedded in L2 (English) sentences that were semantically and linguistically biased. These sentential contexts biased either the target language (L2-English) meaning of the interlingual homograph (e.g., *chat* was read in an English sentence biasing the English "*informal conversation*" meaning), the non-target language (L1-French) meaning of the interlingual homograph (e.g., *chat* was read in an English sentence biasing the French "*cat*" meaning) or neither.

Our first question investigated how target language versus non-target language bias affects interlingual homograph interference in a between-subject manner. Half of the participants read sentences that, when they were highly constraining, biased interpretation toward the target language (L2-English) meaning of interlingual homographs (e.g., "informal conversation" meaning of *chat*) *or* the non-target (L1-French) meaning of interlingual homographs (e.g., "cat" meaning of *chat*) but not both. We expected that the effect of contextual language and semantic cues on interlingual homograph interference would either amplify or diminish interference, depending on the language bias condition, as context helps disambiguate the intended meaning of the target word. We expected this effect to be especially pronounced during late semantic integration (measured here through TRT), as context can reduce or eliminate the interference effect, and consequently, the language co-activation of the non-intended meaning of the words (Duyck et al., 2007; Libben & Titone, 2009; Nayakama & Archibald, 2005). Moreover, to the extent that encountering many English sentences that semantically biased non-target language (French) meanings could have induced a bilingual language mode, the target language versus non-target language bias manipulation could have a more global effect that impacted all interlingual homographs, as opposed to a more specific effect that differentially impacted interlingual homographs within particular sentential contexts (Grosjean, 1998, 2001).

The results for the core models showed that interlingual homograph interference was indeed impacted by target language versus non-target language bias, though this occurred at a global rather than at a specific level. That is, interlingual homograph interference emerged during late-stage reading (TRT), particularly in the non-target language biasing group (L1-French), while it was reduced in the target language biasing group (L2-English). These findings support the Three-Factor framework (Degani & Tokowicz, 2010),
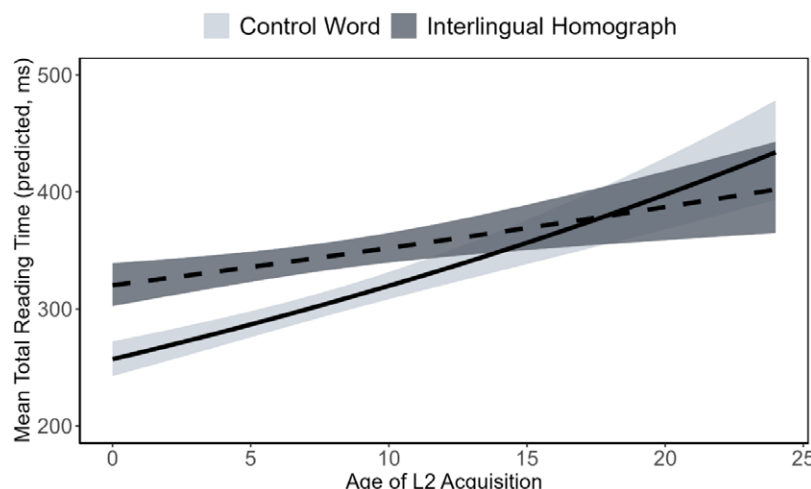

Figure 5. Predicted TRTs for control words and interlingual homographs as a function of L2 AoA. Error bands reflect ±1 SEM.

demonstrating that bilingual reading performance is shaped by the interaction of factors like semantic constraint and language context (target versus non-target meaning bias).

These findings extend previous research, which typically presented contexts that semantically biased target language meanings (i.e., English sentences biasing English meanings of interlingual homographs; see, for example, Libben & Titone, 2009). Here, the results for conditions comparable to those used in past studies (i.e., the target language bias condition, e.g., "*When Jane finished talking to the real-estate agent, she had a thorough* **chat** *with the owner*") revealed a significant interaction between the target word and semantic bias for late-stage reading (TRT). A similar pattern was also found for early-stage reading (GD), although to a lesser degree (i.e., marginal effect of target word; $p = .060$), which could have arisen because of the greater spread of relative word frequency in English and French compared to past work. Given that participants read in their L2 (English), a reduced interlingual homograph interference effect in the target language biasing condition may reflect the use of context to disambiguate the intended meaning of the target word, consistent with BIA+ (Dijkstra & van Heuven, 2002). Indeed, BIA+ suggests that context-driven activation can influence the selection of the appropriate meaning, with clear cues allowing the reader to quickly identify the correct meaning of the word. This is particularly relevant during later integration processes, but is also present, albeit to a lesser extent, in early measures (Hoversten & Traxler, 2016; Libben & Titone, 2009).

However, a crucial goal of this study was to examine how readers processed *mismatched* cues across languages (i.e., the non-target language biasing condition, e.g., "*When the veterinarian noticed the sickly animal, she had a thorough* **chat** *with the owner*"). Specifically, this involved a semantic context written in L2 (English) that misleadingly biased the L1 (French) meaning of the interlingual homograph, even though, ultimately, the intended meaning was the L2 (English). Here, results showed significantly longer later-stage reading times for interlingual homographs, suggesting readers used the context during integration and that this led to activation of the unintended L1 (French) meaning of the interlingual homograph (i.e., a cross-language garden-path effect). A similar pattern was observed for early-stage reading, though to a lesser extent. This supports previous studies that show context modulates bilingual reading behavior and cross-language activation (Duyck et al., 2007; Libben & Titone, 2009; Nakayama & Archibald, 2005).

Crucially, the results show that this contextually driven effect operated at a global rather than a specific level. Accordingly, all interlingual homographs in the non-target language bias group showed greater interference, not just those in constraining sentence contexts. Thus, the observed cross-language garden-path effects likely arose from an increased bilingual mode due to the constant encountering of English sentences biasing the opposite language (French), rather than semantic dynamics happening for each sentence individually. This aligns with the hypothesis that language activation can range from a monolingual mode (one language predominates) to a bilingual mode (both languages activated), with the level of activation influenced by contextual cues (Grosjean, 1998, 2001, 2013).

Considering that participants were instructed in their L2 (English), this account would suggest that under a target language biasing condition (i.e., when there is no conflict between semantic and language bias), readers entered an "L2-only/monolingual mode," expecting to encounter English content exclusively and inhibiting their L1 (French). This would prevent activation of the L1 (French) representation of the interlingual homographs, resulting in an absence of interference effects during early-stage reading.

The finding that interlingual homograph effects were exclusive to the non-target language biasing condition further supports the idea that different task demands elicit differential "language modes." That is, under conditions where there were conflicts between textual (presented in L2) and target language bias (biasing the L1 meaning), and where the L2 (English) meaning was ultimately intended, readers were prompted to reevaluate their "L2-only/monolingual mode." This revaluation allowed the "bilingual mode" to take over and facilitate the activation of the L1 (French) meaning of the interlingual homographs as the reader sought to integrate the correct word meaning (i.e., L2-English) during late-stage reading.

Returning to the Three-Factor Framework (Degani & Tokowicz, 2010), a key provision is that cross-language activation and the use of semantic context in bilingual reading are influenced by word-level properties (e.g., relative frequency) and participant-level factors (e.g., L2 AoA, language experiences). This leads to our second research question: how does cross-language word frequency affect cross-language effects?

We hypothesized that interlingual homographs with higher L1 (French) word frequencies would have longer reading times than those with higher L2 (English) word frequencies, as they may be harder to inhibit during L2 (English) reading. Results revealed a significant interaction between target word and cross-language word frequency scores, as well as between target versus non-target language bias condition and cross-language word frequency scores, during late-stage reading. TRTs for interlingual homographs increased with the word's relative frequency across languages, with words more frequent in the non-target (L1-French) than in the target (L2-English) language showing longer TRTs, likely due to stronger cross-language co-activation. In addition, TRTs were longer when the sentence context biased the L2 (English) meaning, as the activation of the intended L2 meaning would conflict with the strong, highly frequent L1 (French) meaning, further increasing processing difficulty. These findings support BIA+, suggesting that higher-frequency words have stronger resting activation levels, leading to increased processing difficulty (Dijkstra & van Heuven, 2002).

The "language mode" perspective may again help explain these results. If readers were in an "L2-only mode," the influence of cross-language word frequency might not emerge during early processing. In such a mode, only L2 (English) representations would be active, with L1 (French) representations suppressed. Activation of the L1 meaning may occur only during later stages of processing, particularly for homographs that are highly frequent in the non-target language (L1-French). Consistent with this, we found that interlingual homographs with higher L1 (French) than L2 (English) frequencies elicited longer reading times. This aligns with prior studies that have recurred to the language mode framework to explain comparable effects of semantic context on interlingual homograph processing (e.g., Hoversten & Traxler, 2016, 2020).

Finally, the Three-Factor Framework (Degani & Tokowicz, 2010) suggests that some readers may be more sensitive to language contexts than others. While BIA+ and language mode hypotheses acknowledge the role of individual differences in shaping bilingual reading to a certain extent, neither fully explains how language experiences, current language habits and environmental and task demands interact to shape language co-activation effects. To address our third research question investigating the interaction of individual differences on interlingual homograph interference under different sentential contexts, we explored the effects of language history (i.e., L2 AoA) and current reading habits (i.e., reading entropy) in shaping our observed findings.

We predicted that readers with earlier L2 AoAs would exhibit shorter reading times and be more likely to experience interlingual homograph interference than those with later L2 AoAs. This pattern would reflect more integrated lexical and semantic representations across languages, allowing earlier L2 AoA participants to read more efficiently. However, this same integration could also increase their susceptibility to cross-language interference, leading to greater interference effects compared to later L2 AoA readers.

The results partially confirmed these hypotheses. Although L2 (English) AoA did not specifically influence interlingual homograph processing in early-stage L2 (English) reading, we found that in early-stage reading (GD), later L2 AoAs were associated with slower L2 (English) reading times overall, particularly when reading in the non-target language (L1-French) condition. This suggests greater interference from L1 (French) when the context biases the non-target language, regardless of interlingual homograph status, for individuals who acquired their L2 later in life. This is indicative of a "monolingual mode" and consequently, stronger usage of contexts during reading. In contrast, participants with earlier L2 AoAs, whose mental representations are more integrated, may operate in a default "bilingual mode." This would reduce interference from mismatched language context cues (e.g., reading in English while the context biases the French meaning of the word).

For later-stage reading (TRT), we observed the same general L2 AoA pattern found in early-stage reading: later L2 AoAs were associated with slower reading times overall. However, an interaction between the target word and L2 AoA revealed that interlingual homograph interference effects were only present for participants with early L2 AoAs. Interestingly, this was driven by reading times for control words, whose reading times increased more steeply with later L2 AoA than those for interlingual homographs. As a result, individuals who acquired their L2 in adulthood showed similar reading times for both control and interlingual homographs. These findings align with prior research suggesting that cross-language activation may be driven by language-unique control words rather than by words that overlap across a bilingual's two languages (e.g., Libben & Titone, 2009; Pivneva et al., 2014). Our findings thus highlight differences in functional frequency: interlingual homographs, though ambiguous, are encountered in both languages due to their shared orthography, whereas control words may be encountered only in L2 (English). According to the BIA+ model's temporal delay assumption, this difference in exposure may contribute to the delayed activation of L2-only representations (Dijkstra & van Heuven, 2002).

In examining the role of current language experiences (i.e., language-related reading entropy), we predicted that participants with higher entropy scores would exhibit reduced interlingual homograph interference compared to those with lower entropy. This was based on the expectation that individuals who frequently read in multiple languages would manage contextual ambiguity better. The results showed that entropy significantly modulated early-stage reading (GD) only, and although there were no differences specific to target word type, entropy did modulate reading times depending on the type of language cue provided by the context. Specifically, we found that for those who reported a more compartmentalized use of their languages when reading (i.e., low entropy), reading times were higher under non-target language conditions compared to target language conditions. As the entropy score increased, however, reading times for words under this condition decreased.

These results suggest that bilinguals who regularly use both of their languages, particularly in reading, were faster when reading sentences with mismatching language (biasing L2-French) and

semantic (sentence was written in English) cues. This aligns with Schwartz and Kroll (2006), who also found that bilinguals with lower language competence were more sensitive to semantic context during reading. From a language mode perspective, these findings suggest that individual habits in language use in one domain (e.g., reading in one versus multiple languages) influence how flexibly readers switch between monolingual and bilingual "modes" during reading tasks.

6.1. Limitations and conclusions

These findings contribute to the empirical literature by clarifying the role of semantic and language biases in interlingual homograph processing in L2 (English) reading, along with the dynamic influences of language history and individual reading experiences. However, there are limitations to consider. First, we exclusively reported L2 (English) reading in French–English bilingual adults in a multilingual city (i.e., Montreal, Canada). Given previous research reporting differences in language co-activation effects between L1 and L2 reading (e.g., Gullifer & Titone, 2019) and the role of language exposure (e.g., Whitford & Titone, 2015), it is crucial that future research explores this paradigm in the context of L1 reading. It is possible that our current findings supporting the language mode hypothesis may only hold for L2 reading, and that the influence of L2 AoA is reduced, or not present at all, during L1 reading, rendering the "L1 mode" less susceptible to semantic and linguistic factors compared to reading in an L2.

Additionally, we ensured that participants were dominant in their L1 (French) and comparable in proficiency and language exposure; however, slight differences in specific language domains existed. For example, Table 1 indicates that participants in the target language bias condition had significantly higher proficiency in L1 (French) writing compared to those in the non-target language bias condition. Such differences may be inherent to our sample, as proficient bilinguals typically have varied academic experiences, including differential writing experiences based on the language of study. This is why we opted to examine individual linguistic experiences in a continuous fashion.

Another relevant participant-related factor not included in our study was executive control capacity, which has been shown to modulate cross-language activation during sentence reading, especially for interlingual homographs (e.g., Palma et al, 2020; Pivneva et al., 2014). Additionally, our sample consisted of French–English bilinguals with shared alphabetic scripts. Thus, it is unknown whether our findings generalize to bilinguals with different L1 and L2 scripts. This is an area worth exploring, as prior research shows that literacy in different scripts may influence L2 reading proficiency (Mansuri et al., 2025), executive control skills (e.g., Allen et al., 2021) and language co-activation effects (e.g., Coderre & van Heuven, 2014).

Finally, we did not directly examine how cross-language morphological overlap drives processing. Previous work has shown that bilinguals can access a shared morphological network when processing cognates (e.g., Mulder et al., 2015). This could also apply to interlingual homographs, as some can be considered morphological cognates (e.g., **dent** meaning "tooth" in French, morphologically related to English words like "dentist" and "denture"). This overlap suggests the possibility that the use of morphologically related words within a sentence could prime a specific meaning of the interlingual homograph. For instance, mentioning "dentistry" could activate the "hollow surface" meaning of **dent** in English. In our study, we found that two target words were morphological

cognates: *fin* and *fort.* Even in these two cases, related words in the opposite language were not utilized in either of the semantic biasing conditions. For example, the word "finish" was not used when biasing for the French meaning "end" of *fin.* Thus, future research could systematically investigate these effects. In conclusion, and with the above caveats in mind, the results of this study illustrate the dynamic, context-dependent nature of cross-language activation during L2 reading in bilingual adults. They specifically show that lexical properties, contextual, individual language experiences and reading habits jointly shape bilingual reading, and how context can be used to mitigate language co-activation in L2 reading, supporting the Three-Factor Framework (Degani & Tokowicz, 2010).

References

Allen, D., Conklin, K., & Miwa, K. (2021). Cross-linguistic lexical effects in different-script Bilingual reading are modulated by task. *International Journal of Bilingualism*, **25**, 168–188. https://doi.org/10.1177/1367006920943974

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

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (cd-rom). University of Pennsylvania, Philadelphia, PA: Linguistic Data Consortium.

Bailey, L. M., Lockary, K., & Higby, E. (2024). Cross-linguistic influence in the bilingual lexicon: Evidence for ubiquitous facilitation and context-dependent interference effects on lexical processing. *Bilingualism: Language and Cognition*, **27**(3), 495–514. https://doi.org/10.1017/S1366728923000597

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed effects models. *Frontiers in Psychology*, **4**, 328. https://doi.org/10.3389/fpsyg.2013.00328

Bates, D., Maechler, M., Bolker, B., Walker, S., (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**, 1–48. https://doi.org/10.18637/jss.v067.i01

Baum, S., & Titone, D. (2014). Moving toward a neuroplasticity view of bilingualism, executive control, and aging. *Applied Psycholinguistics*, **35**, 857–894. https://doi.org/10.1017/S0142716414000174

Beauvillain, C., & Grainger, J. (1987). Accessing interlexical homographs: Some limitations of a language-selective access. *Journal of Memory and Language*, **26**(6), 658–672. https://doi.org/10.1016/0749-596X(87)90108-2

Bergmann, C., Sprenger, S. A., & Schmid, M. S. (2015). The impact of language co-activation on L1 and L2 speech fluency. *Acta Psychologica*, **161**, 25–35. https://doi.org/10.1016/j.actpsy.2015.07.015

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

Coderre, E. L., & van Heuven, W. J. (2014). The effect of script similarity on executive control in bilinguals. *Frontiers in Psychology*, **5**, 1070. https://doi.org/10.3389/fpsyg.2014.01070

De Groot, A. M., & Nas, G. L. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language*, **30**, 90–123. https://doi.org/10.1016/0749-596X(91)90012-9

de Leeuw, E., & Bogulski, C. (2015). *L2 pronunciation proficiency, language use and age of acquisition as predictors of executive control in bilinguals*. ICPhS At: Glasgow, Scotland.

Degani, T., & Tokowicz, N. (2010). Semantic ambiguity within and across languages: An integrative review. *The Quarterly Journal of Experimental Psychology*, **63**, 1266–1303. https://doi.org/10.1080/17470210903377372

DeLuca, V., Rothman, J., Bialystok, E., & Pliatsikas, C. (2019). Redefining bilingualism as a spectrum of experiences that differentially affects brain structure and function. *Proceedings of the National Academy of Sciences*, **116**(15), 7565–7574. https://doi.org/10.1073/pnas.1811513116

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and language*, **62**, 284–301. https://doi.org/10.1016/j.jml.2009.12.003

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

Dijkstra, T., & van Heuven, W. J. (2013). The BIA model and bilingual word recognition. In J. Grainger, A. M. Jacobs, & A. Jacobs (Eds.), *Localist connectionist approaches to human cognition.* (pp. 189–225). Psychology Press.

Dijkstra, T., Van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition*, **1**, 51–66. https://doi.org/10.1017/S1366728998000121

Dijkstra, T., Wahl, A., Buytenhuijs, F., Van Halem, N., Al-Jibouri, Z., de Korte, M., & Rekké, S. (2019). Multilink: A computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, **22**, 657–679. https://doi.org/10.1017/S1366728918000287

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, **27**, 429–446. https://doi.org/10.1016/0749-596X(88)90066-6

Duyck, W., Desmet, T., Verbeke, L. P. C., & Brysbaert, M. (2004). WordGen: A tool for word selection and nonword generation in Dutch, English, German, and French. *Behavior Research Methods, Instruments & Computers*, **36**, 488–499. https://doi.org/10.3758/BF03195595

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **33**, 663–679. https://doi.org/10.1037/0278-7393.33.4.663

Elsherif, M. M., Preece, E., & Catling, J. C. (2023). Age-of-acquisition effects: A literature review. . Journal of Experimental Psychology: Learning, Memory, and Cognition, **49** (5), 812–847. https://doi.org/10.1037/xlm0001215

Gerard, L. D., & Scarborough, D. L. (1989). Language-specific lexical access of homographs by bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **15**, 305–315. https://doi.org/10.1037/0278-7393.15.2.305

Giezen, M. R., & Emmorey, K. (2016). Language co-activation and lexical selection in bimodal bilinguals: Evidence from picture–word interference. *Bilingualism: Language and Cognition*, **19**(2), 264–276. https://doi.org/10.1017/S1366728915000097

Grosjean, F. (1998). Transfer and language mode. *Bilingualism: Language and Cognition*, **1**, 175–176. https://doi.org/10.1017/S1366728998000285

Grosjean, F. (2001). The bilingual's language modes. In J. L. Nicol (Ed.), *One mind, two languages: Bilingual language processing.* (pp. 1–22). Blackwell Publishing.

Grosjean, F. (2013). Bilingual and monolingual language modes. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics.* (pp. 489–493). Blackwell Publishing.

Gullifer, J. W., Chai, X. J., Whitford, V., Pivneva, I., Baum, S., Klein, D., & Titone, D. (2018). Bilingual experience and resting-state brain connectivity: Impacts of L2 age of acquisition and social diversity of language use on control networks. *Neuropsychologia*, **117**, 123–134.

Gullifer, J. W., Kousaie, S., Gilbert, A. C., Grant, A., Giroud, N., Coulter, K., … Titone, D. (2021). Bilingual language experience as a multidimensional spectrum: Associations with objective and subjective language proficiency. *Applied Psycholinguistics*, **42**(2), 245–278. https://doi.org/10.1017/S0142716420000521

Gullifer, J. W., & Titone, D. (2019). The impact of a momentary language switch on bilingual reading: Intense at the switch but merciful downstream

for L2 but not L1 readers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **45**, 2036–2050. https://doi.org/10.1037/xlm0000695

Gullifer, J. W., & Titone, D. (2020). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*, **23**, 283–294. https://doi.org/10.1017/S1366728919000026

Gullifer, J. W., & Titone, D. (2021a). Engaging proactive control: Influences of diverse language experiences using insights from machine learning. *Journal of Experimental Psychology. General*, **150**, 414–430. https://doi.org/10.1037/xge0000933

Gullifer, J. W., & Titone, D. (2021b). Bilingualism: A neurocognitive exercise in managing uncertainty. *Neurobiology of Language*, **2**, 464–486. https://doi.org/10.1162/nol_a_00044

Hohenstein, S., Matuschek, H., & Kliegl, R. (2017). Linked linear mixed models: A joint analysis of fixation locations and fixation durations in natural reading. *Psychonomic Bulletin & Review*, **24**, 637–651. https://doi.org/10.3758/s13423-016-1138-y

Hoversten, L. J., & Traxler, M. J. (2016). A time course analysis of interlingual homograph processing: Evidence from eye movements. *Bilingualism: Language and Cognition*, **19**, 347–360. https://doi.org/10.1017/S1366728915000115

Hoversten, L. J., & Traxler, M. J. (2020). Zooming in on zooming out: Partial selectivity and dynamic tuning of bilingual language control during reading. *Cognition*, **195**, 104118. https://doi.org/10.1016/j.cognition.2019.104118

Iniesta, A., Yang, M., Beatty-Martínez, A. L., Itzhak, I., Gullifer, J. W., & Titone, D. (2025). Leveraging social network data to ground multilingual background measures: The case of general and socially based language entropy. *Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale*, **79**(1), 15–27. https://doi.org/10.1037/cep0000352

Jared, D., & Szucs, C. (2002). Phonological activation in bilinguals: Evidence from interlingual homograph naming. *Bilingualism: Language and Cognition*, **5**, 225–239. https://doi.org/10.1017/S1366728902003024

Jouravlev, O., & Jared, D. (2014). Reading Russian–English homographs in sentence contexts: Evidence from ERPs. *Bilingualism: Language and Cognition*, **17**, 153–168. https://doi.org/10.1017/S1366728913000205

Kałamała, P., Senderecka, M., & Wodniecka, Z. (2022). On the multidimensionality of bilingualism and the unique role of language use. *Bilingualism: Language and Cognition*, **25**, 471–483. https://doi.org/10.1017/S1366728921001073

Kircher, R. (2014). Thirty years after Bill 101: A contemporary perspective on attitudes towards English and French in Montreal. *Canadian Journal of Applied Linguistics/Revue Canadiennne de Linguistique Appliquée*, **17**(1), 20–50.

Lamarre, P. (2013). Catching "Montréal on the move" and challenging the discourse of unilingualism in Québec. *Anthropologica*, **55**, 41–56. http://www.jstor.org/stable/24467373

Lauro, J., & Schwartz, A. I. (2017). Bilingual non-selective lexical access in sentence contexts: A meta-analytic review. *Journal of Memory and Language*, **92**, 217–233. https://doi.org/10.1016/j.jml.2016.06.010

Li, X., Ng, K. K., Wong, J. J. Y., Lee, J. W., Zhou, J. H., & Yow, W. Q. (2021). Bilingual language entropy influences executive functions through functional connectivity and signal variability. *Brain and Language*, **222**, 105026. https://doi.org/10.1016/j.bandl.2021.105026

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**, 381–390. https://doi.org/10.1037/a0014875

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, **25**, 605–621. https://doi.org/10.1080/20445911.2013.795574

Mansouri, N., Iniesta, A., Hernández-Rivera, E., & Titone, D. (2025). *How does one's first language writing script modulate second language reading: Evidence from the English Reading Online Project.* [manuscript under review] https://osf.io/c9gbq/?view_only=f89305861ea24af7aeb1e92510a93402

Mulder, K., Dijkstra, T., & Baayen, R. H. (2015). Cross-language activation of morphological relatives in cognates: The role of orthographic overlap and task-related processing. *Frontiers in Human Neuroscience*, **9**, 16. https://doi.org/10.3389/fnhum.2015.00016

Nakayama, M., & Archibald, J. (2005). Eye-tracking and interlingual homographs. *Age*, **43**, 17–94.

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE™//A lexical database for contemporary French: LEXIQUE™. *L'année psychologique*, **101**(3), 447–462. https://doi.org/10.3406/psy.2001.1341

Palma, P., & Titone, D. (2020). Bilingual lexical access and reading. In R. R. Heredia & A. B. Cieślicka (Eds.), *Bilingual lexical ambiguity resolution.* (pp. 159–183). Cambridge University Press.

Palma, P., Whitford, V., & Titone, D. (2020). Cross-language activation and executive control modulate within-language ambiguity resolution: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **46**, 507–528. https://doi.org/10.1037/xlm0000742

Paterson, K. B., Liversedge, S. P., Filik, R., Juhasz, B. J., White, S. J., & Rayner, K. (2007). Focus identification during sentence comprehension: Evidence from eye movements. *Quarterly Journal of Experimental Psychology*, **60**, 1423–1445. https://doi.org/10.1080/17470210601100563

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**, 787–796. https://doi.org/10.1037/a0035583

Pivneva, I., Sudarshan, A., Mercier, J., Baum, S., & Titone, D. (2011). Individual Differences in Executive Function Modulate Bilingual Between-Language Activation during Reading in Healthy Older Adults: Evidence from Eye-tracking. *Procedia - Social and Behavioral Sciences*, **23**, 35–36, https://doi.org/10.1016/j.sbspro.2011.09.157

R Development Core Team (2025). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, **124**, 372–422. https://doi.org/10.1037/0033-2909.124.3.372

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**, 720–732. https://doi.org/10.1037/0096-1523.30.4.720

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language*, **55**, 197–212. https://doi.org/10.1016/j.jml.2006.03.004

Titone, D., Libben, M., Mercier, J., Whitford, V., & Pivneva, I., (2011). Bilingual lexical access during L1 sentence reading: The effects of L2 knowledge, semantic constraint, and L1-L2 intermixing. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, **37**, 1412–1431. https://doi.org/10.1037/a0024492

Titone, D., & Tiv, M. (2023). Rethinking multilingual experience through a systems framework of bilingualism. *Bilingualism: Language and Cognition*, **26**(1), 16. https://doi.org/10.1017/S1366728921001127

Tiv, M., O'Regan, E., & Titone, D. (2021). In a bilingual state of mind: Investigating the continuous relationship between bilingual language experience and mentalizing. *Bilingualism: Language and Cognition*, **24**, 918–931. https://doi.org/10.1017/S1366728921000225

Voeten, C (2023). *buildmer: Stepwise elimination and term reordering for mixed effects regression.* R package version 2.9. https://CRAN.R-project.org/package=buildmer

von Studnitz, R. E., & Green, D. (2002). Interlingual homograph interference in German-English bilinguals: Its modulation and locus of control. *Bilingualism: Language and Cognition*, **5**, 1–23. https://doi.org/10.1017/S1366728902000111

Whitford, V., Pivneva, I., & Titone, D. (2016). Eye movement methods to investigate bilingual reading. In R. R. Heredia, J. Altarriba, & A. B. Cieślicka (Eds.), *Methods in bilingual reading comprehension research.* (pp. 183–211). Springer Science + Business Media. https://doi.org/10.1007/978-1-4939-2993-1_8

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