Journal of Cognitive Psychology

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/pecp21

Does meaningful background speech modulate predictability effects during Chinese reading?

Chuanli Zang, Shuangshuang Wang, Yu Guan, Zhu Meng, Manman Zhang & Simon P. Liversedge

To cite this article: Chuanli Zang, Shuangshuang Wang, Yu Guan, Zhu Meng, Manman Zhang & Simon P. Liversedge (20 May 2024): Does meaningful background speech modulate predictability effects during Chinese reading?, Journal of Cognitive Psychology, DOI: 10.1080/20445911.2024.2352902

To link to this article: https://doi.org/10.1080/20445911.2024.2352902

Published online: 20 May 2024.

JOURNAL OF COGNITIVE PSYCHOLOGY
https://doi.org/10.1080/20445911.2024.2352902

Does meaningful background speech modulate predictability effects during Chinese reading?

Chuanli Zang [a], Shuangshuang Wang[b*], Yu Guan[b*], Zhu Meng[c], Manman Zhang[b] and Simon P. Liversedge [a]

[a]School of Psychology and Humanities, University of Central Lancashire, Preston, UK; [b]Key Research Base of Humanities and Social Sciences of the Ministry of Education, Faculty of Psychology, Tianjin Normal University, Tianjin, People's Republic of China; [c]Department of Psychology, Jiangsu Normal University, Xuzhou, People's Republic of China

ABSTRACT

Previous research indicates that background speech disrupts reading comprehension processes, but it remains unclear whether the disruption derives from semantic or phonological speech properties, and whether it affects early lexical processing or later sentence integration. Native Chinese speaking participants read sentences containing high- or low-predictability words under meaningful Chinese speech, meaningless Uyghur speech or silence conditions. Results showed that Chinese but not Uyghur speech produced increased total fixations compared to reading in silence, suggesting disruption was semantic in nature. While a standard predictability effect was comparable across background speech conditions in target word analyses, this effect disappeared in the Chinese speech condition in later measures and regions. The findings suggest that Chinese background speech may delay higher order (post-lexical) processing associated with sentence integration during reading, with implications for the Interference-by-Process hypothesis.

ARTICLE HISTORY
Received 25 October 2023
Accepted 2 May 2024

KEYWORDS
Background speech; predictability; Chinese reading

Reading is an essential process for acquiring knowledge and information but occurs in diverse environmental settings every day. Often, it takes place in silent circumstances, like a quiet library or a private office, where external visual and auditory distractions are minimal. However, such ideal circumstances are often unavailable, and therefore, much of the daily reading happens in the presence of background sounds, such as non-verbal noise (e.g. outside traffic) and speech conversations (e.g. phone calls in a shared office; TV playing in the house). Although numerous studies have demonstrated that background sounds, and particularly background speech, might interfere with an individual's cognitive processing and performance (referred to as the irrelevant speech effect, Colle & Welsh, 1976; see Vasilev et al., 2018), the precise nature of interference, particularly in the context of complex cognitive tasks like reading, remains uncertain. For example, it is unclear which specific properties of speech contribute to this interference, and how background speech disrupts reading. Does it primarily impair initial lexical processing, that is processing associated with ascertaining the identity, phonological form and semantic meaning of a word, or does it only affect later, higher order, processes involved in the integration of the word in relation to preceding sentence and discourse context? To shed light on these questions, the present study employed eye-tracking methodology to measure on-line reading behaviour in order to investigate the nature and time course of the impact of background speech on Chinese reading.

Currently, there are two representative hypotheses that provide an explicit account of which aspects of background speech disrupt reading (see Meng et al., 2020 for a review): the Phonological Interference hypothesis (Salamé & Baddeley, 1982, 1989) and the Semantic Interference hypothesis (Marsh et al., 2008, 2009; Martin et al., 1988). The former was proposed based on the working memory model. Within this model, the phonological loop consists of a phonological store for temporarily storing auditory information, and an articulatory rehearsal device for maintaining such

CONTACT Chuanli Zang czang@uclan.ac.uk School of Psychology and Humanities, University of Central Lancashire, Preston, Lancashire PR1 2HE, UK
*These authors contributed equally to this work and their order in authorship is arbitrary.

information (Baddeley, 1986). Speech input, including irrelevant background speech, has automatic access to the phonological store, whereas visually presented stimuli are converted by articulatory rehearsal process into phonological code that is also reserved in the phonological store. The presence of background speech can interfere with the encoding and retrieval of visual information in the phonological store, thus causing disruption to occur in reading. Following this account, any speech input, regardless of its meaningfulness, disrupts reading.

In contrast, the latter hypothesis posits that the disruptive effect of background speech on reading comprehension occurs at the level of semantic rather than phonological processing, as reading involves extracting meaning from the text. Thus, only when the semantic content of background speech is accessible, is reading disrupted. Martin et al. (1988) provided evidence for this hypothesis and found that compared with meaningless speech (foreign language), meaningful speech (English) produced more disruption to the comprehension of native English speakers. Furthermore, speech containing random, meaningful words was more disruptive than that containing random, meaningless nonwords. In this study, it appears that the semantic content, rather than phonological properties, of irrelevant speech interfered with processing associated with the extraction of meaning, thereby impairing comprehension. Aligning with this hypothesis, the Interference-by-Process hypothesis proposed by Marsh et al. (2008, 2009) assumes that the extraction of meaning from background speech and the formation of an interpretation of written text call on shared processes, and this causes disruption to reading comprehension. It, therefore, predicts that only meaningful background speech, but not meaningless speech, yields disruption to reading. It is apparent that meaningful speech has the potential to disrupt reading comprehension at the semantic level of individual words. If background speech triggers semantic activation of lexical entries, then semantic interference may occur when conflicting semantic representations arise from both auditory speech and visually presented words.

Previous behavioural studies investigating the impact of background speech on off-line reading comprehension have yielded mixed results with some studies reporting detrimental effects of meaningful speech to reading, whilst others report no effects (Vasilev et al., 2018). Eye-tracking methodology provides an excellent on-line tool to assess how background speech might affect different aspects of reading process and the time course of any such effects. To date, there has been some, limited, research investigating these effects during sentence reading (e.g. Cauchard et al., 2012; Hyönä & Ekholm, 2016; Meng et al., 2020; Vasilev et al., 2019; Yan et al., 2018). For example, Cauchard et al. (2012) found that meaningful background speech led to longer gaze durations (the sum of all first-pass fixations on a word before moving to another word), increased reading and rereading times compared to silence. The increase in reading time was mainly due to readers revisiting previously fixated words both within the current and preceding sentence. Cauchard et al. argued that background speech disrupted reading by causing difficulties in high-level (post-lexical) text integration. However, their findings did not distinguish between phonological and semantic levels of interference. Furthermore, their additional manipulation, wherein participants were interrupted on half of the trials by an unrelated 60-second audio story, could also have potentially influenced their reading behaviour. Hyönä and Ekholm (2016) conducted four eye-tracking experiments where participants were required to read syntactically complex Finnish sentences for comprehension in the presence of different types of background speech. Experiment 1 showed that meaningful Finnish speech did not increase fixation durations compared to meaningless Italian speech or silence, providing no evidence for either the semantic or the phonological interference hypothesis (Salamé & Baddeley, 1982, 1989). However, in their subsequent experiments, Hyönä and Ekholm found that compared with silence and coherent speech, scrambled background speech resulted in longer rereading times, and greater disruption to reading. Furthermore, the scrambling of word order from the same semantic content as the to-be-read text or from an unrelated, different text, appeared to be equally distracting, indicating that semantic processing of speech and text per se, rather than comparable content between the two is required to produce interference in reading (Marsh et al., 2008, 2009, see also Meng et al., 2020 for similar evidence).

Regardless of whether background speech effects on on-line sentence processing are phonological or semantic in nature, it seems that such interference is primarily evident in fixations made during re-reading (cf. Experiment 1 in Hyönä & Ekholm, 2016). Such eye movements are usually taken as an indication of difficulty in post-lexical linguistic processing, usually during interpretation and integration (Liversedge et al., 1998). However, there are two further studies that have directly examined whether lexical processing is disrupted in the presence of background speech. Yan et al. (2018) required participants to read Chinese sentences, each containing a high or low frequency target word, while listening to meaningful background speech (a broadcast recording of China Central Television's news programme, and its semantic content was irrelevant to the written sentences), meaningless speech (the same speech scrambled in small chunks) and silence. They obtained longer reading times, more fixations and more regressions for the text read with meaningful speech than for text read with meaningless speech and silence. In addition, low frequency words received longer reading times and less skipping than high frequency words. Interestingly, there was an interaction between frequency and background speech for the first fixation duration (the duration of the first fixation on a word, regardless of the number of fixations the word received during first-pass reading) on target words, with reliable frequency effects only in silence but not in the other two speech conditions. As the frequency effect did not occur in first fixation (other than under conditions of silence), but did occur in later reading times measures regardless of background sound, Yan et al. argued that background speech (regardless of whether meaningful or not) may disrupt early stages of lexical processing in reading.

Vasilev et al. (2019) further examined this issue in English readers but failed to find any interaction between target word frequency and background speech, and this maintained even when the frequency of all words in the sentence (rather than just the target word) were included in a supplemental analysis. Furthermore, the influence of meaningful English speech was not reliable for first-pass measures but only on second-pass measures (including refixations and regressions). These results suggest that meaningful speech may not affect initial lexical processing of words, but rather disrupts higher order processing associated with formation of a coherent sentence representation. It should be apparent, based on the discussion here, that it is currently not clear as to the precise stages of reading that are affected by meaningful speech.

The present study sought to extend previous research by examining the influence of background speech on lexical (or even pre-lexical) processing, and post-lexical integration in Chinese readers. It is well established that words that are more predictable, tend to receive shorter fixations compared to less predictable words. Moreover, predictable words are more likely to be skipped than unpredictable words (Rayner et al., 2005; Rayner & Well, 1996; see Staub, 2015 for a review). Additionally, there are studies demonstrating the influence of predictability on later measures such as regressions (saccades made from right to left) and go-past time (the sum of all fixations from the first fixation on the word until a fixation to the right of that word, e.g. Staub, 2011; Staub & Goddard, 2019). Thus, the predictability of words has an impact on eye movements starting from one of the earliest measures, word skipping and extending to later measures like regressions. And, to be clear, in order for predictability effects to occur, the reader necessarily must have evaluated the meaning of the current word in relation to preceding context. In this case, therefore, predictability may play a role in anticipating potential upcoming words, accessing and maintaining lexical information, and integrating current lexical information into the syntactic, semantic and discourse contexts (Staub, 2015). Thus, by manipulating the predictability of words in a sentence, we can examine how meaningful background speech affects the earliest through to later stages of linguistic processing through examination of whether predictability effects are diminished or delayed during Chinese sentence reading.

In the present experiment, native Chinese speaking participants were required to read Chinese sentences containing a high or low predictable word in three background speech conditions: meaningful but irrelevant Chinese speech, meaningless Uyghur speech (with no recognisable semantic information or identifiable words to Chinese speakers who do not speak Uyghur, even though it is meaningful to Uyghur speakers. Based on this, we refer to the Uyghur speech as meaningless) or silence. Uyghur and Chinese are two distinct languages with their own unique phonotactic patterns. Uyghur is a Turkic language, primarily written using a Perso-Arabic-based alphabet. It includes a different set of consonant and vowel sounds, syllable structures and phonotactic rules compared to the Chinese language. Uyghur is also an agglutinative language rich with suffixes that serve a number of different functions (Yan et al., 2014). Chinese, in contrast, has a high number of homophones and a limited number of syllable structures that typically consist of a consonant sound followed by a vowel sound. Further, Chinese has a distinct tone system that also plays a key role in the language. Given that the participants in the study were native Chinese speakers (without experience of the Uyghur language), it should be apparent that our adoption the Uyghur speech represented a background speech condition in which semantic and lexical content were unavailable to participants. According to the Phonological Interference hypothesis (Salamé & Baddeley, 1982, 1989), both Chinese and Uyghur speech might be expected to disrupt the reading process to a similar extent with disruption likely to occur at a relatively early stage of lexical processing indexed by skipping probability (the probability a word did not receive a fixation during first-pass reading) and first-pass reading time measures such as first fixation duration, single fixation duration (the duration of the first fixation on a word when only one first-pass fixation was made on a word) and gaze duration. Irrelevant speech forms (both Chinese and Uyghur) activate basic, abstract phonological representations that underpin phonological forms associated with words in the Chinese lexicon. For example Jones and Macken (1995), argued that the degree of phonotactic resemblance between the background speech (e.g. Uyghur) and phonological code associated with visual stimuli (Chinese) does not modulate the level of distraction. Thus, according to this view, it should not matter whether auditory distraction effects derive from Chinese or Uyghur background speech, those effects should be comparable. In contrast, if the interference occurs at a semantic level (or perhaps at a level of processing at which orthographic, phonological and semantic representations are activated together to attain understanding of meaning), as suggested by Martin et al. (1988) and Marsh et al. (2008, 2009), then Chinese speech should be more disruptive than Uyghur speech and the disruption might occur at a later stage of processing reflected in the measures such as go-past time and total fixation duration (the sum of all fixations on a word), because Chinese speech may trigger activation of irrelevant lexical entries and obstruct semantic representations derived from visually presented sentences. Critically, we might expect that Chinese speech produces greater disruption to reading than Uyghur, and this distraction might pose a challenge to readers in anticipating upcoming words, accessing and maintaining lexical information, or integrating words into sentential context during reading. Under such circumstances, auditory Chinese distraction effects might reduce or delay any effects of predictability compared to effects observed in the Uyghur speech and silence conditions.

Method

Participants

Ninety undergraduate students (mean age = 21 years, SD = 2; 70 females) from Tianjin Normal University participated in the eye tracking experiment. They were all native Chinese speakers with normal, or corrected-to-normal, vision. None of them were able to speak or understand Uyghur language, used in the meaningless speech condition. The research protocol was approved by the Faculty of Psychology Ethics Committee at Tianjin Normal University. Based on Westfall's (2015) power calculation method and an average effect size of $d = 0.47$ as reported by Vasilev et al. (2019), the power of our sample size (90 participants and 60 sets of target strings in total) is estimated to be 0.919. This suggests that we have sufficient power to detect an effect of background speech in our study.

Apparatus

Participants' eye movements were recorded using an SR Research Eyelink 1000 eye-tracker with a sampling rate of 1000 Hz. Sentences were displayed on a 24-inch DELL CRT monitor with a resolution of 1920 × 1080 pixels and a refresh rate of 150 Hz. At a viewing distance of 65 cm from the participant to the monitor, each Chinese character subtended an approximate visual angle of 1.1 degrees.

Materials and design

Three types of background sounds were included: meaningful Chinese speech, meaningless Uyghur speech and silence with no background sound. The speech sounds were the same as those used by Meng et al. (2020) and recorded by a native Uyghur speaker who was extremely proficient in Mandarin. The speech content was taken from a Chinese-Uyghur bilingual book (Lim, 2016) so that the semantic content of the Chinese and Uyghur speech was equivalent. The speech text was narrative and easy to understand, but it did not include dialogue, nor any semantic connection with the visually presented sentences. The intensity of both Chinese and Uyghur speech was 58–70 dB(A). The duration of all the auditory stimuli was sufficiently long to cover the period in which participants read the text (for details, see Meng et al., 2020).

Visual stimuli consisted of 60 experimental sentences with each sentence containing either a high or low predictable two-character target word. The sentence frame was identical for each pair of target words, that is, pretarget and posttarget contexts were always the same across predictability conditions (see also Hand et al., 2010). The target words were embedded in the middle part of each sentence. The predictability of target words was assessed by 20 participants, who did not take part in the subsequent eye-tracking experiment, and were asked to complete a sentence given the preceding sentence context up to but not including the target. For example, for the sentence "The children went outside to … ", the response was coded as "1" if a participant correctly provided the target word that the experimenter expected (e.g. "play") and "0" if the participant provides any alternative. The predictability of a word in a particular sentence was determined by the proportion of "1" responses for that word. The mean predictability in the current experiment was significantly higher for the high (0.69, SD = 0.15, range = 0.45–1) than the low predictability words (0.02, SD = 0.04, range = 0–0.3; $t(59) = 32.92$, $p < 0.001$). The frequency of the first and second character, and the whole target word did not differ across predictability conditions (all ts < 1.10, see Table 1, Cai & Brysbaert, 2010). Likewise, the stroke number of the first and second character (the number of strokes each character is comprised of), and the whole target word did not vary across predictability conditions (all ts < 1.64). Finally, all sentences were prescreened to ensure their naturalness. A group of 40 participants, who did not take part in the eye-tracking experiment, rated the sentence naturalness on a 5-point scale (1 = very unnatural, 5 = very natural). All sentences were rated very natural (high predictability: M = 4.2, SD = 0.4; low predictability: M = 4.1, SD = 0.4), though there was a difference between the two conditions (t = 3.36, see also Li et al., 2018 for comparable values and statistics).[1]

The experiment had a 2 (Predictability: high or low) ×3 (Background Speech: Chinese speech, Uyghur speech or silence) within-participant design. Each participant was required to read 60 experimental sentences (see Figure 1), 30 filler sentences (without high/low predictable target words and being irrelevant to the purpose of the current experiment) and 20 practice sentences (in total 110 sentences). The practice sentences were presented at the beginning of the experiment while Chinese and Uyghur speech was played interchangeably in the background through headphones, allowing participants to familiarise themselves with reading under varying background speech conditions. The remaining sentences were divided into three blocks, with each block containing 20 experimental sentences (10 sentences with high predictable words and 10 with low predictable words) and 10 filler sentences and being presented in one background speech condition. Sentences within each block were presented in a random order, but the order of the three background speech blocks and the two predictability conditions of target words were counterbalanced across six files according to a Latin square design. Each sentence under different sound conditions was read only once by each participant.

Procedure

Prior to the start of the experiment, each participant was presented with an information sheet and a written consent form. They were instructed to read the sentence silently and for comprehension, and to ignore any background sounds they may hear from headphones during the experiment because the sounds were irrelevant to their reading. Participants were then required to sit comfortably in front of the eye tracker and complete a 3-point horizontal calibration procedure (with an average calibration error below 0.20 degrees). At the beginning of each trial, a drift correction dot was presented on the left side of the screen and upon its fixation the sentence was immediately displayed. Participants were recalibrated wherever necessary. When they completed reading a sentence, they pressed a response key to terminate the display. Approximately 40% of the sentences were followed by a simple Yes/No comprehension question, to which they responded by pressing "F" (Yes) or "J" (No) on the keyboard. For example, the sentence "他放弃了国外优厚的待遇和生活条件" (meaning "He gave up the generous salary and comfortable living conditions provided abroad") was succeeded by the Yes/No question "他最终选择待在国外" (meaning "Did he choose to stay overseas", and the answer to this question should be "No"). Participants were informed that they were free to withdraw at any time (without penalty) and their data would be treated confidentially, securely stored and anonymised. After completing the experiment, they were provided comprehensive debriefing. Participants wore headphones during the whole experiment and did not report any discomfort or overloading in relation to the volume of the speech. The experiment lasted about 30 min, with a break for the participants at the end of each block.

Data analysis

We conducted both global (sentence level) and local (target word level) analyses of data (e.g. Bai et al., 2008; Meng et al., 2020; Vasilev et al., 2019; Yan et al., 2018). Global analyses are indicative of how background speech affected general eye movement behaviour during reading across the entire sentence, whereas local analyses are reflective of how high or low predictability target words were processed across background speech conditions and thus those words were defined as target areas of interest (whereas the preceding and subsequent regions were defined as pretarget and posttarget regions, see Additional Analyses). In the global analyses, we computed total sentence reading time (the sum of all fixations made on the sentence), mean fixation duration, number of fixations and number of regressive saccades. In the local analyses, we computed skipping probability (SP), first fixation duration (FFD), single fixation duration (SFD), gaze duration (GD), go-past time

[1]The difference in naturalness is very likely due to participants considering more predictable situations to also be more natural. Given the differences in naturalness, a further set of LMM analyses were conducted in which we included these ratings as a centred continuous covariate. Analyses with naturalness included as a covariate did not differ from analyses without this variable included as a covariate. For this reason, we can be sure that none of our effects was caused by differences in naturalness. For the full set of analyses including naturalness as a covariate, please see Tables A1 and A2 in the Appendix.

Table 1. Statistical properties for the high and low predictable words.

	Frequency (per million)			Number of strokes		
Predictability	First character	Second character	Target word	First character	Second character	Target word
High	696(972)	1028(1882)	83(164)	8.1(3.2)	7.7(3.0)	15.8(4.4)
Low	677(719)	790(1624)	52(152)	7.4(2.5)	8.6(3.3)	15.9(3.8)

Note: Standard Deviations in parentheses.

Predictability	Target words	Sentence
High	游泳	夏季在凉爽的海水里**游泳**是一件非常棒的事。
Low	冲浪	夏季在凉爽的海水里**冲浪**是一件非常棒的事。

Figure 1. An example sentence with high and low predictable words. The target words are in bold in this example but were presented normally in the experiment. The literal translation for the Chinese sentence "夏季 | 在凉爽的海水里 | 游泳/冲浪 | 是一件非常棒的事" is "During the summer | in cool seawater | **swimming/surfing** | is a wonderful activity to do". This translation allows one to see the context prior to the target word.

(GPT) and total fixation duration (TFD). SP is generally taken as the earliest measure for which an effect associated with the target word might be seen, as the decision to skip a word occurs during a fixation on a word earlier in the sentence than the target word; FFD, SFD and GD are first pass reading time measures that generally reflect early and relatively immediate processing when fixating on the target region. In contrast, if an effect is not observed for early first pass reading time measures, but for GPT and TFD, then this is generally taken as an indication of the variable having a relatively late influence on processing. Both the global and local measures were exported from EyeLink Data Viewer.

Linear mixed models (LMM) were conducted using the lme4 package (version 1.1-28) in R (version 4.1.0, R Development Core Team, 2021) to analyse the data. Contextual predictability, background speech and their interaction were treated as fixed factors. Participants and items were treated as crossed random factors. Three successive contrasts were set up to analyse differences between Uyghur speech and silence, differences between Chinese speech and silence, and differences between Chinese speech and Uyghur speech. Full models with maximal random effects across subjects and items were initially run for each measure and if these failed to converge, they were trimmed down starting with items, then participants, with the removal of correlations, interactions between factors, then random slopes. Fixation times were log-transformed in all analyses to increase normality.

Results

The mean comprehension accuracy was 92% (SD = 11%) in the silence condition, 90% (SD = 14%) in the Uyghur speech condition and 92% (SD = 10%) in the Chinese speech condition. There were no reliable accuracy differences across the background speech conditions (F = 0.94), indicating that participants fully understood the sentences and their comprehension was not affected by background speech. This is not surprising given that comprehension questions were quite easy to answer and required only a yes/no response (see Vasilev et al., 2019 and many other studies). The eye movement data were preprocessed in DataViewer such that fixations longer than 1200 ms and shorter than 80 ms were excluded (Bai et al., 2008; Yan et al., 2018). Furthermore, trials were removed in which sentences received less than three fixations (0.3%), or eye movement measures were above three standard deviations from each participant's mean (Global analyses: 1.0%; Local analyses: 1.6%).

Global analyses

Descriptive statistics and fixed effects estimations for the global analyses are shown in Tables 2 and 3. As can be seen, Chinese speech produced longer sentence reading times, more fixations and more regressive saccades than Uyghur speech or silence (all ts > 2.11), whereas the latter two conditions did not differ significantly in these measures (all ts < 1.89, though readers spent numerically less time reading and making slightly fewer fixations in the Uyghur than the Silence condition). It appears that the meaningless speech was no more disruptive than silence, suggesting that background speech sound per se is not disruptive, or meaningless speech is as minimally disruptive as silence. This result does not provide

Table 2. Mean (Standard Deviation) for the global eye movement measures.

	Total sentence reading time (ms)	Mean fixation duration (ms)	Number of fixations	Number of regressive saccades
Silence	3817(1202)	223(24)	11.5(4)	3.1(1.4)
Uyghur	3807(1309)	223(24)	11.4(4)	3.1(1.6)
Chinese	3978(1377)	224(24)	12.0(4)	3.4(1.6)

evidence for the Phonological Interference hypothesis (Salamé & Baddeley, 1982, 1989) which stipulates that both background speech conditions would be equally disruptive. Rather, our results indicate that meaningful rather than meaningless background speech disrupted reading through increasing the number of fixations and regressions (though, note, not increasing the average fixation duration), replicating the previous literature and providing evidence for the semantic interference hypothesis (Meng et al., 2020; Vasilev et al., 2019; Yan et al., 2018).

Local analyses

Descriptive statistics and fixed effects estimations for the target word analyses are shown in Tables 4 and 5. Consistent with the results from global analyses, there were effects of background speech only in total fixation duration with longer fixations for Chinese speech than the Uyghur speech or silence (all *ts* > 2.08), but there were no differences between the latter two conditions. Again, in line with the global analyses, the results indicate that the meaningless speech is as minimally disruptive as silence, which does not support the Phonological Interference hypothesis (Salamé & Baddeley, 1982, 1989). In contrast, our results suggest that the meaningful Chinese background speech results in greater re-reading and interferes with later stages of word processing during reading as the effect does not occur in any early first pass reading time measures (Vasilev et al., 2019; Yan et al., 2018).

Effects of word predictability were reliable in all eye movement measures (all *ts* or $|z| > 3.52$), such that low predictability words received longer fixations and were skipped less often than high predictability words during initial and later processing, replicating the standard predictability effects (Rayner et al., 2005; Rayner & Well, 1996; see Staub, 2015 for a review). However, there were no interactions between word predictability and background speech conditions in any local measures. Thus, we obtained no evidence that predictability effects are modulated by background speech during normal reading.

Additional analyses

In order to investigate whether any effects associated with processing of target words may have occurred prior to participants fixating those words, or whether delayed effects appeared on the region subsequent to the target, we carried out analyses for the pretarget and posttarget regions. Note, again, for each pair of target words the pretarget and posttarget regions were identical, making direct comparions possible across the high and low predictability conditions. Trials were removed in which eye movement measures were above three standard deviations from each participant's mean (1.2%). Eye movement measures including first pass reading time, second pass reading time (the sum of all fixations made on a region following the initial first pass time), go-past time, and total reading time were computed.

As can be seen from Tables 6 and 7, Chinese background speech resulted in longer time on the pretarget region than Uyghur speech (all *ts* > 1.97) in all eye movement measures, and longer than silence in total reading time ($t = 3.00$). There were no differences between Uyghur speech and the silence condition (all *ts* < 1.02). In addition, the predictability effect was reliable in second pass reading time and total reading time but did not interact with background speech conditions. In the post-target analyses, first pass reading time was shorter in the Chinese speech than silence condition ($t = 2.49$), and slightly shorter in the Uyghur speech than silence condition ($t = 1.93$). This aspect of our results is somewhat surprising, as we might ordinarily assume that any speech sounds, and potentially, particularly meaningful speech sounds, might be distracting (and therefore disruptive) relative to silence. Clearly, at the posttarget region during the first pass reading, this was not the case. Presumably participants terminate their initial reading prematurely at the posttarget region under background speech conditions, to facilitate their re-reading of preceding sentences for semantic integration. This assumption is supported by the increased time spent re-reading the preceding section and the posttarget region (as can be seen if we subtract first-pass reading time from go-past time) in the Chinese speech condition (701 ms), relative to the Uyghur speech (631 ms) and silence (618 ms) conditions. The predictability effect was reliable in the second pass reading time, go-past time and total reading time (all *ts* > 2.22). Interestingly, there was an interaction between predictability and background speech in the second pass reading time. Planned contrasts showed that the predictability effect was significant under silence ($b = 0.18$, $SE = 0.06$, $t = 3.00$, $p < 0.01$) and

Table 3. LMMs analyses for the global eye movement measures.

	Total sentence reading time (ms)				Mean fixation duration (ms)				Number of fixations				Number of regressive saccades			
	b	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*
Intercept	8.14	0.05	168.18	**<0.001**	5.40	0.01	476.11	**<0.001**	2.36	0.04	63.46	**<0.001**	0.94	0.05	18.79	**<0.001**
Uyghur vs. Silence	−0.02	0.01	−1.82	0.07	0.00	0.00	−0.56	0.58	−0.02	0.01	−1.88	0.06	−0.01	0.02	−0.55	0.58
Chinese vs. Silence	0.03	0.01	2.13	0.04	0.00	0.00	0.89	0.38	0.03	0.01	2.12	0.04	0.06	0.02	2.65	0.01
Chinese vs. Uyghur	0.05	0.01	4.23	<0.001	0.01	0.00	1.44	0.15	0.05	0.01	4.06	<0.001	0.07	0.02	3.46	<0.001

Note: Significant terms featured in bold, and terms approaching significance are underlined. FFD = first fixation duration; SFD = single fixation duration; GD = gaze duration; GPT = go-past time; TFD = total fixation duration; SP = skipping probability.

Uyghur speech conditions ($b = 0.12$, $SE = 0.06$, $t = 2.07$, $p < 0.05$), but not significant in the Chinese speech condition ($b = 0.02$, $SE = 0.06$, $t = 0.34$, $p = 0.74$). It appears that the disruptive influence of Chinese speech on the use of contextual predictability information, if anything, emerged comparatively late during processing. This, again, demonstrates that Chinese speech appears to negatively impact later integration stages of sentence processing.

Note that zero second pass reading times were removed from these analyses. There were substantially more instances of zero second pass reading times in the post-target (68.6%) than the pre-target (33.8%) region and this accounts for the considerable difference in mean total reading times between these two regions.

Discussion

Previous studies have demonstrated that background speech disrupts processes associated with reading comprehension (Vasilev et al., 2018 for a review). However, it remains unclear whether this disruption is attributed to the semantic or phonological properties of speech, and whether it impacts early stages of lexical processing or later stages of sentence integration. The present study employed eye tracking to record participants' eye movements while reading single Chinese sentences containing a high or low predictability word under different background speech conditions: meaningful but irrelevant Chinese speech, meaningless Uyghur speech or silence. The results from the global and local analyses are very straightforward: the meaningful Chinese speech, but not the meaningless Uyghur speech, produced increased total fixation durations, more fixations and regressions compared to reading in silence. In other words, Chinese speech was more disruptive to reading than the Uyghur speech and silence, but we obtained no evidence that Uyghur speech disrupted reading any more than the silence condition.

Our results are consistent with findings from reading of alphabetic languages (Cauchard et al., 2012; Hyönä & Ekholm, 2016; Vasilev et al., 2019) and Chinese (Meng et al., 2020; Yan et al., 2018). Further, the observed patterns offer little support for the Phonological Interference hypothesis (Salamé & Baddeley, 1982, 1989), which posits that all types of speech input, whether meaningful or meaningless, should interfere with reading by automatically accessing phonological representations stored in memory, thereby producing disruption to encoding and retrieval of linguistic representations based on orthographic information processed during reading. Our findings make reasonable sense in that there is a relatively weak connection between the nature of orthographic

Table 4. Mean (Standard Deviation) for eye movement measures on the target words.

Predictability	Background Speech	FFD (ms)	SFD (ms)	GD (ms)	GPT (ms)	TFD (ms)	SP
High	Silence	226(37)	225(37)	236(46)	284(81)	300(84)	0.32(0.21)
	Uyghur	226(41)	224(41)	235(45)	294(122)	312(106)	0.34(0.20)
	Chinese	235(49)	232(48)	246(55)	298(104)	323(94)	0.33(0.21)
Low	Silence	238(41)	236(44)	258(54)	320(124)	338(103)	0.27(0.19)
	Uyghur	237(40)	237(45)	258(53)	323(107)	333(77)	0.29(0.20)
	Chinese	236(42)	236(40)	257(60)	318(107)	354(109)	0.30(0.19)

Table 5. LMM analyses for eye movement measures on the target words.

	FFD (ms)				SFD (ms)				GD (ms)			
Predictors	*b*	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*
(Intercept)	5.40	0.02	347.38	**<0.001**	5.39	0.02	343.07	**<0.001**	5.44	0.02	302.83	**<0.001**
Predictability: Low vs. High	0.04	0.01	4.04	**<0.001**	0.05	0.01	4.51	**<0.001**	0.07	0.01	6.68	**<0.001**
Uyghur vs. Silence	0.00	0.01	−0.40	0.69	−0.01	0.01	−0.41	0.68	0.00	0.01	−0.24	0.81
Chinese vs. Silence	0.01	0.01	1.21	0.23	0.01	0.01	1.05	0.30	0.01	0.02	0.87	0.38
Chinese vs. Uyghur	0.02	0.01	1.60	0.11	0.02	0.01	1.44	0.15	0.02	0.02	1.10	0.27
Predictability × Uyghur vs. Silence	−0.02	0.02	−0.81	0.42	−0.01	0.02	−0.21	0.83	−0.01	0.03	−0.19	0.85
Predictability × Chinese vs. Silence	−0.04	0.02	−1.82	0.07	−0.03	0.02	−1.15	0.25	−0.04	0.03	−1.46	0.14
Predictability × Chinese vs. Uyghur	−0.02	0.02	−1.00	0.32	−0.02	0.02	−0.93	0.35	−0.03	0.03	−1.26	0.21

	GPT (ms)				TFD (ms)				SP			
Predictors	*b*	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*	*b*	*SE*	*z*	*p*
(Intercept)	5.58	0.03	215.95	**<0.001**	5.65	0.03	220.22	**<0.001**	−0.95	0.10	−9.67	**< 0.001**
Predictability: Low vs. High	0.10	0.01	6.40	**<0.001**	0.09	0.02	6.17	**<0.001**	−0.22	0.06	−3.53	**<0.001**
Uyghur vs. Silence	0.01	0.02	0.38	0.70	0.01	0.02	0.51	0.61	0.12	0.08	1.57	0.12
Chinese vs. Silence	0.01	0.02	0.75	0.45	0.05	0.02	2.61	**0.01**	0.10	0.08	1.35	0.18
Chinese vs. Uyghur	0.01	0.02	0.37	0.71	0.04	0.02	2.09	**0.04**	−0.02	0.08	−0.22	0.82
Predictability × Uyghur vs. Silence	−0.01	0.04	−0.27	0.79	−0.04	0.03	−1.04	0.30	−0.01	0.15	−0.03	0.97
Predictability × Chinese vs. Silence	−0.05	0.04	−1.33	0.18	−0.03	0.03	−0.73	0.47	0.04	0.15	0.27	0.79
Predictability × Chinese vs. Uyghur	−0.04	0.04	−1.05	0.30	0.01	0.03	0.31	0.76	0.05	0.15	0.31	0.76

Note: Significant terms featured in bold, and terms approaching significance are underlined. FFD = first fixation duration; SFD = single fixation duration; GD = gaze duration; GPT = go-past time; TFD = total fixation duration; SP = skipping probability.

Table 6. Mean (Standard Deviation) for eye movement measures on the pretarget and posttarget regions.

Region	Predictability	Background speech	First pass time (ms)	Second pass time (ms)	Go-past time (ms)	Total reading time (ms)
Pretarget region	High	Silence	1138(385)	530(309)	1138(385)	1502(544)
		Uyghur	1130(466)	524(381)	1130(466)	1526(729)
		Chinese	1191(538)	584(367)	1191(538)	1640(772)
	Low	Silence	1142(474)	563(407)	1142(474)	1578(742)
		Uyghur	1145(453)	591(355)	1145(453)	1570(661)
		Chinese	1192(495)	583(348)	1192(495)	1680(747)
Posttarget region	High	Silence	621(268)	524(320)	1182(604)	780(325)
		Uyghur	591(258)	506(291)	1198(718)	772(369)
		Chinese	585(260)	541(295)	1256(674)	786(341)
	Low	Silence	611(264)	583(299)	1285(676)	811(347)
		Uyghur	581(247)	569(312)	1236(678)	776(347)
		Chinese	585(245)	560(320)	1316(694)	802(357)

Note that zero second pass reading times were removed from these analyses. There were substantially more instances of zero second pass reading times in the post-target (68.6%) than the pre-target (33.8%) region and this accounts for the considerable difference in mean total reading times between these two regions.

representations and their pronunciation in written Chinese text, whilst there is a much more direct correspondence between orthography and semantic meaning (Zang et al., 2011). To be clear, unlike alphabetic languages such as English where there are often relations between the individual letters of a word and its pronunciation (e.g. C, A, T spells "cat"), in Chinese, there is little (if any) relation between the constituent strokes of a character and how it is pronounced, but a much more direct relationship between a character's semantic meaning and its orthographic form. Presumably, the disruption from the Chinese speech indicates that Chinese speech activates semantic representations associated with words, even if that speech is irrelevant and required to be ignored. Commonality of semantic processing associated with the irrelevant Chinese speech stream and the written Chinese text stream likely produces the observed disruption to reading. These findings provide support for

Table 7. LMMs analyses for eye movement measures on the pretarget and posttarget regions.

		First pass time (ms)				Second pass time (ms)				Go-past time (ms)				Total reading time (ms)			
Region	Effect	*b*	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*
Pretarget region	Intercept	6.90	0.05	134.58	**<0.001**	5.95	0.05	108.99	**<0.001**	6.90	0.05	134.58	**<0.001**	7.21	0.05	142.83	**<0.001**
	Predictability: Low vs High	0.00	0.01	0.17	0.87	0.06	0.02	2.30	0.02	0.00	0.01	0.17	0.87	0.03	0.01	3.36	<0.001
	Uyghur vs. Silence	−0.02	0.02	−1.01	0.31	−0.02	0.03	−0.63	0.53	−0.02	0.02	−1.01	0.31	−0.01	0.01	−0.96	0.34
	Chinese vs. Silence	0.02	0.02	1.07	0.28	0.06	0.03	1.61	0.11	0.02	0.02	1.07	0.28	0.05	0.02	3.00	<0.001
	Chinese vs. Uyghur	0.03	0.02	1.98	0.051	0.08	0.03	2.56	0.01	0.03	0.02	1.98	0.051	0.06	0.02	4.07	<0.001
	Predictability × Uyghur vs. Silence	0.02	0.02	0.85	0.40	0.08	0.06	1.42	0.15	0.02	0.02	0.85	0.40	0.01	0.02	0.44	0.66
	Predictability × Chinese vs. Silence	0.02	0.02	0.80	0.42	−0.01	0.06	−0.12	0.90	0.02	0.02	0.80	0.42	0.02	0.02	0.71	0.48
	Predictability × Chinese vs. Uyghur	0.00	0.02	−0.04	0.97	−0.09	0.06	−1.57	0.12	0.00	0.02	−0.04	0.97	0.01	0.02	0.27	0.78
Posttarget region	Intercept	6.17	0.05	124.43	<0.001	6.03	0.05	120.42	< 0.001	6.86	0.06	115.36	<0.001	6.47	0.06	113.03	<0.001
	Predictability: Low vs High	−0.01	0.01	−0.98	0.33	0.09	0.03	2.77	0.01	0.05	0.01	3.63	<0.001	0.03	0.01	2.23	0.03
	Uyghur vs. Silence	−0.04	0.02	−1.93	0.058	0.01	0.04	0.36	0.72	−0.03	0.02	−1.41	0.16	−0.03	0.02	−1.64	0.11
	Chinese vs. Silence	−0.06	0.02	−2.49	0.01	0.05	0.04	1.20	0.23	0.02	0.03	0.72	0.47	−0.01	0.02	−0.52	0.60
	Chinese vs. Uyghur	−0.01	0.02	−0.52	0.60	0.03	0.04	0.84	0.40	0.05	0.03	1.93	0.057	0.02	0.02	1.07	0.29
	Predictability × Uyghur vs. Silence	0.00	0.03	0.07	0.94	−0.06	0.08	−0.70	0.49	−0.03	0.03	−0.88	0.38	−0.02	0.03	−0.70	0.48
	Predictability × Chinese vs. Silence	0.02	0.03	0.72	0.47	−0.20	0.08	−2.48	0.01	−0.02	0.04	−0.60	0.55	−0.01	0.03	−0.36	0.72
	Predictability × Chinese vs. Uyghur	0.02	0.03	0.64	0.52	−0.14	0.08	−1.79	0.07	0.01	0.04	0.29	0.78	0.01	0.03	0.34	0.74

Note: Significant terms featured in bold, and terms approaching significance are underlined.

the Semantic Interference hypothesis (Martin et al.,1988) and the Interference-by-Process hypothesis (Marsh et al., 2008, 2009), and suggest that the interference effect caused by background speech in Chinese sentence reading is due to the meaningfulness of Chinese speech, and that this is semantic, rather than phonological, in nature.

It should be noted that previous research has also tried to discriminate between the Semantic Interference hypothesis and the Interference-by-Process hypothesis. For example, Hyönä and Ekholm (2016) created scrambled speech either from the same semantic content as the to-be-read text or from a different text, and found a similar disruptive effect under the two scrambled speech conditions. Hyönä and Ekholm argued that it is not the shared semantic content between background speech and text being read that causes the interference in reading, but instead shared semantic processing associated with speech and reading. In addition, Meng et al. (2020) asked participants to either search text for a noncharacter (i.e. noncharacter detection task) or read text for meaning (i.e. semantic acceptability task) while being exposed to Chinese speech, Uyghur speech or silence. Meng et al. did not observe any effects of background speech for the former task, but a substantial disruptive effect of the Chinese speech for the latter task, where participants were required to engage in semantic processing in order to construct a coherent representation of the sentence meaning. The two sources of evidence seem to demonstrate that the disruptive effect of background speech in text processing is pronounced when the speech is meaningful and the visual task requires semantic processing. These findings align more with the Interference-by-Process hypothesis (Marsh et al., 2008, 2009). Our results are consistent with Hyönä and Ekholm (2016) and Meng et al. (2020), and on the basis of their findings, we argue that commonality of semantic processing is likely the cause of our findings.

Next let us consider the impact of background speech on early stages of lexical processing or later stages of sentence integration. First, the analysis of target words revealed a standard predictability effect, with shorter fixation durations and more skipping for predictable compared to unpredictable words. These findings replicate previous studies (e.g. Rayner et al., 2005; Rayner & Well, 1996) and demonstrate that our predictability manipulation was effective. These results also demonstrate that readers were processing the text effectively and were integrating words with preceding context appropriately. Interestingly, however, our predictability effects were not modulated by background speech at all, as there were no interactions between predictability and background speech for any measures in the target word analyses. The results align with the findings of Vasilev et al. (2019), where effects of word frequency (another of the "Big Three" influences on eye movements in reading, Clifton et al., 2016), were consistent across all background speech conditions. However, our results appear to be inconsistent with the findings of Yan et al. (2018). In Yan et al.'s study, there was an interaction between frequency and background speech for first fixation duration, with a reliable frequency effect in conditions of silence but not in the presence of meaningful Chinese, or meaningless (scrambled) Chinese speech. Yan et al. argued that background speech disrupts lexical processing of words and delays the emergence of frequency effects in eye movement measures. However, the fact that both meaningful and meaningless speech showed a similar effect (and only in the first fixation duration measure) in Yan et al.'s study undermines the argument that semantic inconsistencies between background speech and written text cause disruption to lexical access (see Vasilev et al., 2019 for similar arguments). Based on the two studies and our own findings, there seems to be little solid evidence to suggest that the semantic properties of background speech exert an influence on early lexical processing during Chinese reading.

In line with the lack of an early influence of background speech on lexical processing, our target word analyses did not show any disruptive effect of Chinese speech in skipping probability or any early reading time measures such as first fixation duration, single fixation duration and gaze duration. The only reliable effects occurred in total fixation duration. This suggests that the initial processing of words within sentences remained unaffected by Chinese speech. Additionally, when analysing the post-target region in our additional analysis, a predictability effect was observed in the silence and Uyghur speech conditions, but not in the Chinese speech condition. However, this effect only emerged in the second pass reading time. These findings further indicate that irrelevant background Chinese speech does not make less predictable words more challenging to anticipate or access. Instead, Chinese background speech may simply delay or impair the integration of words into the current sentential representation (e.g. via syntactic, semantic and discourse processing) and this occurs during later stages of reading.

Before closing, we should reflect on some limitations of our experimental work and consider whether our findings have any practical implications beyond their theoretical contribution. Of course, an immediate limitation is the fact that our work was limited to Chinese reading. It remains an open question as to whether comparable effects might be obtained in languages beyond Chinese that have different linguistic characteristics. Additionally, here we have solely considered distraction effects in relation to word predictability based on sentential context. We know that there are many other linguistic variables that affect eye movement behaviour in reading and how auditory stimuli modulate effects associated with such variables are open issues for future investigations. To be clear, much further work is required before our understanding of distraction effects in reading is comprehensive. A final issue that is interesting, and perhaps moves slightly beyond effects of sentential predictability, concerns whether the specific personal relevance of the content of auditory information might determine the degree of distraction a reader experiences. For example, information about dramatic increases fees for students at the University of Central Lancashire might, presumably, be much more salient as a distractor for a reader that is a student at the University of Central Lancashire than for a student at a different university. Presumably, it should be possible to manipulate personal relevance of information independent of its sentential predictability. Of course, the current study provided no insight into this question.

It is the case that our results might be of significant practical relevance, particularly in respect of environments where reading and understanding must be achieved in the presence of background speech, such as shared offices, learning spaces and workplaces. The disruptive impact of background speech can vary depending on the meaningfulness, familiarity and degree to which the language being spoken is comprehended. In shared office settings, for instance, background speech in the native language of the listener, such as telephone conversations or on-line meetings, clearly have the potential to disrupt reading performance or general learning activities that are underpinned by reading. In these cases, the use of headphones or noise cancellation systems may help minimise such disruption. On the contrary, when background speech is meaningless or in an unfamiliar foreign language, reading performance or general learning activities are less likely to be affected. Also, influences of irrelevant background speech may differ based on individual differences including age and neurodiversity, thus making it necessary to customise work and learning environments specifically to meet the needs of diverse populations.

In summary, our study required participants to read single Chinese sentences containing high or low predictability words under conditions of meaningful Chinese speech, meaningless Uyghur speech and silence. We demonstrated that Chinese speech, but not Uyghur speech, significantly disrupted reading and this disruption was minimally associated with early aspects of lexical and linguistic processing, but more substantially associated with higher-order post-lexical stages required for the construction of coherent sentential and discourse representations. Our findings align with the Interference-by-Process hypothesis wherein commonality of semantic processing underlies the disruptive effect of background speech in text processing. In addition, our results may offer practical insights into the design of work and learning environments, prompting further consideration of the potential influence of background speech on individuals with diverse linguistic backgrounds.

Acknowledgments

We acknowledge support from ESRC Grant (ES/R003386/1).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Economic and Social Research Council: [Grant Number ES/R003386/1].

Data availability statement

All data sets and analysis scripts are publicly available at: https://osf.io/hw2sv/.

ORCID

Chuanli Zang http://orcid.org/0000-0002-9573-4968
Simon P. Liversedge http://orcid.org/0000-0002-8579-8546

References

Baddeley, A. (1986). *Working memory*. Oxford University Press.

Bai, X. J., Yan, G. L., Liversedge, S. P., Zang, C. L., & Rayner, K. (2008). Reading spaced and unspaced Chinese text: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1277–1287. https://doi.org/10.1037/0096-1523.34.5.1277

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS One*, *5*(6), e10729. https://doi.org/10.1371/journal.pone.0010729

Cauchard, F., Cane, J. E., & Weger, U. W. (2012). Influence of background speech and music in interrupted reading: An eye-tracking study. *Applied Cognitive Psychology*, *26*(3), 381–390. https://doi.org/10.1002/acp.1837

Clifton, C., Jr., Ferreira, F., Henderson, J. M., Inhoff, A. W., Liversedge, S. P., Reichle, E. D., & Schotter, E. R. (2016). Eye movements in reading and information processing: Keith Rayner's 40 year legacy. *Journal of Memory and Language*, *86*, 1–19. https://doi.org/10.1016/j.jml.2015.07.004

Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning and Verbal Behavior*, *15*(1), 17–31. https://doi.org/10.1016/S0022-5371(76)90003-7

Hand, C. J., Miellet, S., O'Donnell, P. J., & Sereno, S. C. (2010). The frequency-predictability interaction in reading: it depends where you're coming from. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(5), 1294. https://doi.org/10.1037/a0020363

Hyönä, J., & Ekholm, M. (2016). Background speech effects on sentence processing during reading: An eye movement study. *PLoS One*, *11*(3), e0152133. https://doi.org/10.1371/journal.pone.0152133

Jones, D. M., & Macken, W. J. (1995). Phonological similarity in the irrelevant speech effect: Within- or between-stream similarity? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 103–115. https://doi.org/10.1037/0278-7393.21.1.103

Li, N., Wang, S., Mo, L., & Kliegl, R. (2018). Contextual constraint and preview time modulate the semantic preview effect: Evidence from Chinese sentence reading. *Quarterly Journal of Experimental Psychology*, *71*(1), 241–249. https://doi.org/10.1080/17470218.2017.1310914

Lim, A. (2016). *Little detective* (H. Liu & L. Zhang, Chinese Translator; W. Baikere, Uyghur Translator). Kashi Uyghur Language Publishing House.

Liversedge, S. P., Paterson, K. B., & Pickering, M. J. (1998). Eye movements and measures of reading time. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 55–75). Elsevier Science Ltd.

Marsh, J. E., Hughes, R. W., & Jones, D. M. (2008). Auditory distraction in semantic memory: A process-based approach. *Journal of Memory and Language*, *58*(3), 682–700. https://doi.org/10.1016/j.jml.2007.05.002

Marsh, J. E., Hughes, R. W., & Jones, D. M. (2009). Interference by process, not content, determines semantic auditory distraction. *Cognition*, *110*(1), 23–38. https://doi.org/10.1016/j.cognition.2008.08.003

Martin, R. C., Wogalter, M. S., & Forlano, J. G. (1988). Reading comprehension in the presence of unattended speech and music. *Journal of Memory and Language*, *27*(4), 382–398. https://doi.org/10.1016/0749-596X(88)90063-0

Meng, Z., Lan, Z., Yan, G., Marsh, J. E., & Liversedge, S. P. (2020). Task demands modulate the effects of speech on text processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(10), 1892–1905. https://doi.org/10.1037/xlm0000861

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., & 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

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org

Salamé, P., & Baddeley, A. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior*, *21*(2), 150–164. https://doi.org/10.1016/S0022-5371(82)90521-7

Salamé, P., & Baddeley, A. D. (1989). Effects of background music on phonological short-term memory. *The Quarterly Journal of*

Experimental Psychology Section A, *41*(1-A), 107–122. https://doi.org/10.1080/14640748908402355

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

Staub, A., & Goddard, K. (2019). The role of preview validity in predictability and frequency effects on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(1), 110–127. https://doi.org/10.1037/xlm0000561

Vasilev, M. R., Kirkby, J. A., & Angele, B. (2018). Auditory distraction during reading: A Bayesian Meta-Analysis of a continuing controversy. *Perspectives on Psychological Science*, *13*(5), 567–597. https://doi.org/10.1177/1745691617747398

Vasilev, M. R., Liversedge, S. P., Rowan, D., Kirkby, J. A., & Angele, B. (2019). Reading is disrupted by intelligible background speech: Evidence from eye-tracking. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(11), 1–93. https://doi.org/10.1037/xhp0000680

Westfall, J. (2015). PANGEA: Power ANalysis for GEneral Anova designs [Unpublished manuscript]. http://jakewestfall. org/publications/pangea. pdf

Yan, G., Meng, Z., Liu, N., He, L., & Paterson, K. B. (2018). Effects of irrelevant background speech on eye movements during reading. *Quarterly Journal of Experimental Psychology*, *71*(6), 1270–1275. https://doi.org/10.1080/17470218.2017.1339718

Yan, M., Zhou, W., Shu, H., Yusupu, R., Miao, D., Krügel, A., & Kliegl, R. (2014). Eye movements guided by morphological structure: Evidence from the Uighur language. *Cognition*, *132*(2), 181–215. https://doi.org/10.1016/j.cognition.2014.03.008

Zang, C., Liversedge, S. P., Bai, X., & Yan, G. (2011). Eye movements during Chinese reading. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 961–978). Oxford University Press.

Appendix

Table A1. Fixed effect estimations for eye movement measures on target words when naturalness was included as a covariate.

Effect	FFD (ms)				SFD (ms)				GD (ms)			
Predictors	*b*	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*
(Intercept)	5.40	0.02	347.74	**<0.001**	5.39	0.02	343.92	**<0.001**	5.45	0.02	302.65	**<0.001**
Low vs. high	0.04	0.01	3.70	**0.00**	0.05	0.01	4.09	**0.00**	0.07	0.01	6.29	**0.00**
Uyghur vs. Silence	0.00	0.01	−0.42	0.68	−0.01	0.01	−0.42	0.67	0.00	0.01	−0.25	0.81
Chinese vs. Silence	0.01	0.01	1.20	0.23	0.01	0.01	1.04	0.30	0.01	0.02	0.87	0.39
Chinese vs. Uyghur	0.02	0.01	1.61	0.11	0.02	0.01	1.45	0.15	0.02	0.02	1.10	0.27
Predictability ×Uyghur vs. Silence	−0.02	0.02	−0.81	0.42	−0.01	0.02	−0.21	0.83	−0.01	0.03	−0.19	0.85
Predictability × Chinese vs. Silence	−0.04	0.02	−1.81	0.07	−0.03	0.02	−1.14	0.26	−0.04	0.03	−1.46	0.15
Predictability × Chinese vs. Uyghur	−0.02	0.02	−0.99	0.32	−0.02	0.02	−0.92	0.36	−0.03	0.03	−1.26	0.21
Naturalness	−0.02	0.02	−0.81	0.42	−0.02	0.02	−1.04	0.30	−0.01	0.02	−0.41	0.68

	Go-past time (ms)				TFD (ms)				SP			
Predictors	*b*	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*	*b*	*SE*	*z*	*p*
(Intercept)	5.58	0.03	216.00	**<0.001**	5.65	0.03	221.94	**<0.001**	−0.95	0.10	−9.68	**<0.001**
Low vs. high	0.09	0.02	5.93	**0.00**	0.09	0.02	5.49	**0.00**	−0.25	0.06	−3.91	**0.00**
Uyghur vs. Silence	0.01	0.02	0.45	0.65	0.01	0.02	0.49	0.63	0.12	0.08	1.58	0.11
Chinese vs. Silence	0.01	0.02	0.79	0.43	0.05	0.02	2.60	**0.01**	0.10	0.08	1.35	0.18
Chinese vs. Uyghur	0.01	0.02	0.34	0.74	0.04	0.02	2.10	**0.04**	−0.02	0.08	−0.23	0.82
Predictability × Uyghur vs. Silence	−0.01	0.04	−0.28	0.78	−0.04	0.03	−1.03	0.30	−0.01	0.15	−0.04	0.97
Predictability × Chinese vs. Silence	−0.05	0.04	−1.30	0.19	−0.02	0.03	−0.71	0.48	0.04	0.15	0.27	0.79
Predictability × Chinese vs. Uyghur	−0.04	0.04	−1.01	0.31	0.01	0.03	0.32	0.75	0.05	0.15	0.31	0.76
Naturalness	−0.01	0.03	−0.35	0.73	−0.05	0.03	−1.62	0.11	−0.23	0.12	−1.92	0.06

Note: Significant terms featured in bold, and terms approaching significance are underlined.

Table A2. Fixed effect estimations for eye movement measures on the pretarget and posttarget regions when naturalness was included as a covariate.

Region	Effect	First pass time (ms)				Second pass time (ms)				Go-past time (ms)				Total reading time (ms)			
		b	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*
Pretarget region	Intercept	6.90	0.05	134.39	**<0.001**	5.96	0.05	115.27	**<0.001**	6.88	0.13	53.64	**<0.001**	7.21	0.05	145.03	**<0.001**
	Predictability	0.00	0.01	0.23	0.82	0.02	0.03	0.79	0.50	0.00	0.01	0.23	0.82	0.02	0.01	2.34	0.02
	Uyghur vs. Silence	−0.02	0.02	−1.01	0.31	−0.02	0.03	−0.60	0.55	−0.02	0.02	−1.01	0.31	0.01	0.01	0.96	0.34
	Chinese vs. Silence	0.02	0.02	1.07	0.28	0.06	0.04	1.56	0.12	0.02	0.02	1.07	0.28	−0.05	0.02	−3.00	0.00
	Chinese vs. Uyghur	0.03	0.02	1.98	0.05	0.07	0.03	2.51	0.01	0.03	0.02	1.98	0.05	0.06	0.02	4.05	<0.001
	Predictability × Uyghur vs. Silence	0.02	0.02	0.85	0.40	0.08	0.06	1.39	0.16	0.02	0.02	0.85	0.40	0.01	0.02	0.44	0.66
	Predictability × Chinese vs. Silence	0.02	0.02	0.80	0.42	−0.01	0.06	−0.17	0.86	0.02	0.02	0.80	0.42	0.02	0.02	0.71	0.48
	Predictability × Chinese vs. Uyghur	0.00	0.02	−0.04	0.97	−0.09	0.06	−1.58	0.11	0.00	0.02	−0.04	0.97	0.01	0.02	0.27	0.79
	Naturalness	0.01	0.03	0.19	0.85	−0.27	0.05	−5.27	0.00	0.01	0.03	0.19	0.85	−0.06	0.03	−1.97	0.05
Posttarget region	Intercept	6.17	0.05	124.26	<0.001	6.04	0.05	121.33	<0.001	6.86	0.06	116.09	<0.001	6.47	0.06	113.21	<0.001
	Predictability	−0.02	0.02	−0.89	0.38	0.07	0.03	1.99	0.05	0.03	0.02	2.15	0.03	0.02	0.01	1.07	0.29
	Uyghur vs. Silence	−0.04	0.02	−1.92	0.06	0.01	0.04	0.32	0.75	−0.03	0.02	−1.72	0.09	−0.03	0.02	−1.63	0.10
	Chinese vs. Silence	−0.06	0.02	−2.50	0.01	0.05	0.04	1.16	0.25	0.02	0.02	1.02	0.31	−0.01	0.02	−0.52	0.61
	Chinese vs. Uyghur	−0.01	0.02	−0.53	0.60	0.03	0.04	0.83	0.41	0.05	0.02	2.74	0.01	0.02	0.02	1.06	0.29
	Predictability × Uyghur vs. Silence	0.00	0.03	0.07	0.95	−0.06	0.08	−0.68	0.50	−0.03	0.04	−0.87	0.38	−0.02	0.03	−0.71	0.48
	Predictability × Chinese vs. Silence	0.02	0.03	0.72	0.47	−0.20	0.08	−2.48	0.01	−0.02	0.04	−0.57	0.57	−0.01	0.03	−0.36	0.72
	Predictability × Chinese vs. Uyghur	0.02	0.03	0.65	0.51	−0.14	0.08	−1.80	0.07	0.01	0.04	0.30	0.76	0.01	0.03	0.34	0.73
	Naturalness	−0.02	0.04	−0.37	0.71	−0.17	0.07	−2.35	0.02	−0.13	0.04	−3.19	0.00	−0.08	0.04	−2.32	0.02

Note. Significant terms featured in bold, and terms approaching significance are underlined.