Submitted as a Rapid Communication

Effects of Irrelevant Background Speech on Eye Movements during Reading

Guoli Yan[1]

Zhu Meng[1]

Nina Liu[1]

Liyuan He[1]

Kevin B. Paterson[2]

1. Academy of Psychology and Behavior, Tianjin Normal University

2. Department of Neuroscience, Psychology and Behaviour, University of Leicester

Abstract

The irrelevant speech effect (ISE) refers to the impairment of visual information processing by background speech. Prior research on the ISE has focused on short-term memory for visually-presented word lists. The present research extends this work by using measurements of eye movements to examine effects of irrelevant background speech during Chinese reading. This enabled an examination of the ISE for a language in which access to semantic representations is not strongly mediated by phonology. Participants read sentences while exposed to meaningful irrelevant speech, meaningless speech (scrambled meaningful speech) or silence. A target word of high or low lexical frequency was embedded in each sentence. The results show that meaningful, but not meaningless, background speech produced increased re-reading. In addition, the appearance of a normal word frequency effect, characterised by longer fixation times on low compared to high frequency words, was delayed when meaningful or meaningless speech was present in the background. These findings show that irrelevant background speech can disrupt normal processes of reading comprehension and, in addition, that background noise can interfere with the early processing of words. The findings add to evidence showing that normal reading processes can be disrupted by environmental noise such as irrelevant background speech.

Keywords: Background speech; Irrelevant speech effects; Eye movements during reading; Chinese reading

People are exposed daily to a wide variety of environmental noise. These include not only non-verbal noise (e.g., traffic), but speech noise such as background conversations and radio and television broadcasting, often in situations in which individuals are focused on an important cognitive activity (e.g., working in a busy office or engaging in private study). Understanding how background noise, and especially background speech, might interfere with the performance of cognitive tasks is therefore an important practical concern. Studies that have investigated this issue under laboratory conditions often show that background speech interferes with performance on cognitive tasks, typically resulting in longer response times and decreased accuracy. This phenomenon is known as the irrelevant speech effect (ISE) (Colle & Welsh, 1976).

ISEs have been shown in studies using short-term memory tasks in which participants presented with visual stimuli (usually word lists) while simultaneously exposed to background speech are later asked to recall these stimuli. Disruption of recall appears independent of the meaningfulness of speech, so that meaningless speech that includes the same phonemic content as normal speech produces similar disruption as meaningful speech (LeCompte & Shaibe, 1997). But while this might suggest the acoustic properties of speech are important in disrupting memory encoding, other research shows that a repeated utterance (‘ah’) is not disruptive, whereas an auditory signal exhibiting appreciable change (‘bah’, ‘dah’, ‘gah’) produces disruption (Jones, Madden, & Miles, 1992). Speech therefore must have acoustic variation typical of normal speech to produce the ISE.

Another important domain in which the ISE might influence cognitive performance is reading (Jones, 1995). Working memory theory (Salamé & Baddeley, 1982) attributes such effects to phonological recoding processes that occur automatically during reading. Following this account, verbal material is temporarily buffered as verbatim representations in a phonological memory store. All speech input, including task-irrelevant background speech, gains obligatory access to this store, whereas text gains access by being recoded into sub-vocal speech during rehearsal. Thus, phonological codes from both sources can co-exist and interference occurs due to irrelevant speech corrupting phonological representations for text. The account predicts all speech-like input, including meaningless text, might disrupt reading, although meaningful speech may be more disruptive. By comparison, an inference-by-process account (Marsh, Hughes, & Jones, 2007, 2009) attributes the ISE to interference due to shared use of cognitive processes rather than similarity in the content of linguistic codes. This predicts disruption due to conflict between semantic processing during reading and the automatic processing of background speech. Moreover, disruption should be produced only by meaningful speech, as meaningless speech will not recruit semantic processes.

Surprisingly few studies have investigated the ISE for reading and evidence primarily comes from related tasks such as proof-reading (Halin, Marsh, Haga, Holmgren, & Sörqvist, 2013; Jones, Miles, & Page, 1990), memory for prose (Banbury & Berry, 1998; Sörqvist, Ljungberg, Ljung, 2010), or performance answering comprehension questions (Martin, Wogalter, & Forlano, 1988; Oswald, Tremblay, & Jones, 2000; Sörqvist, Halin, & Hygge, 2010). Moreover, a key study suggests the ISE does not affect the processing of words (Boyle & Coltheart, 1996), as cognitive load imposed by the syntactic complexity of sentences but not irrelevant background speech affected acceptability judgements for sentences containing a homophone spelling error (e.g., “eight” instead of “ate”) or inappropriate word (e.g., “sight” instead of “ate”).

However, two recent eye movements studies revealed an ISE during reading (Cauchard, Cane, & Weger, 2012; Hyönä & Ekholm, 2016). Both show irrelevant background speech slows reading by causing participants to re-read more. Moreover, Hyönä and Ekholm showed this disruption was greater for meaningful text and meaningless text created by re-ordering the words in meaningful text compared to background foreign language speech (in a language unfamiliar to readers). However, as these studies examined eye movement behaviour generally, it remains unclear if the ISE impairs the processing of words during reading or only higher-order comprehension processes, although such effects are predicted by theoretical accounts. Therefore, to gain a clearer indication of the ISE during reading, the present experiment examined effects when sentences contained target words that differed in their frequency of written usage.

Eye movements are highly informative about the processing of words during reading (Rayner, 2009). In particular, words that are lower in frequency, and therefore less familiar, typically receive longer fixations than higher frequency words (e.g., Rayner & Duffy, 1986). This word frequency effect is considered a hallmark of normal reading. Consequently, if the ISE can be shown to impair this effect, by disrupting its normal time-course or producing abnormally long fixations times for low compared to high frequency words, this would indicate that the ISE impairs the processing of words. We also assessed whether such effects occur independently of the meaningfulness of speech. Finally, as the present experiment was conducted in Chinese, the findings will reveal if an ISE is observed for character-based languages such as Chinese as well as alphabetic languages in previous research (Cauchard et al., 2012; Hyönä & Ekholm, 2016). A crucial aspect of written Chinese is that access to semantics may not be as strongly mediated by the phonological recoding of written input compared to alphabetic languages, even though phonological information is mandatorily activated (e.g., Perfetti & Zhang, 1991; Zhou & Marslen-Wilson, 2000). Accordingly, the present findings will reveal if an ISE is observed for a character-based language in which phonological codes may be less important for access to word meanings.

Method

The research adhered to ethical requirements of the Helsinki Declaration. Informed consent was obtained from all participants.

Participants. Participants were 42 Chinese speakers (aged 19-23 years) from Tianjin Normal University. All reported normal vision and hearing.

Stimuli & Design. We used a 3 (background speech: meaningful speech, meaningless speech, silence) × 2 (word frequency: high, low) within-participants design. Meaningful speech was created from a recording of China Central Television's evening news broadcast (16 minutes duration). Previous research has created meaningless speech conditions using foreign language speech (e.g., Banbury & Berry, 1998; Hyönä & Ekholm, 2016), reversed speech (e.g., Jones et al., 1990; Oswald et al., 2001) or by creating scrambling speech in which speech is segmented into small chunks (often corresponding to words) that are re-ordered (e.g., Hyönä & Ekholm, 2016; Martin et al., 1988). The present research adapted this latter approach by using MATLAB to segment and scramble small chunks (60ms each) of the meaningful speech. The meaningless speech therefore preserved the general acoustic variation present in meaningful speech while removing its meaning, including that of individual words. The intensity of both types of speech was 58–66 dB(A). The ambient level for the silent condition was 45 dB(A).

Sentence stimuli consisted of 72 sentence frames, each comprising 23 characters (Figure 1). These included an interchangeable high (186 counts-per-million) or low (6 counts-per-million) frequency 2-character target word selected from the SUBTLEX-CH database (Cai & Brysbaert, 2010), and matched for number of strokes ($ts<1.2$, $ps>.2$; Table 1). Target words appeared in various locations within words but never within the first or last four characters in a sentence. A cloze procedure in which 20 participants who did not participate in the experiment provided completions to sentence fragments showed high and low frequency target words were equally unpredictable in each sentence context. In addition, rating scales administered to an additional 20 participants showed sentences containing high and low frequency words were equally natural and easy to read.

The sentences were divided into 3 blocks, each presented in one background speech condition for each participant. This was counterbalanced across participants so each block was shown in each speech condition an equal number of times and the order of blocks was fully rotated. Each block began with 18 practice sentences in the same speech condition.

Apparatus & Procedure. Participants were tested individually and instructed to read silently and for comprehension. Background speech was presented over headphones. Eye movements were recorded for right eye movements during binocular viewing using an EyeLink 1000 eye-tracker (SR Research, Canada) and 1000Hz sampling rate. Sentence stimuli were presented on a high-definition monitor across a single line. At 70cm viewing distance, each character subtended 1° (and so was of normal size for reading). The eye-tracker was calibrated at the beginning of the experiment using a 3-point horizontal calibration (accuracy >.2°), and calibration was checked prior to each trial and the eye-tracker recalibrated as necessary.

At the start of each trial, a fixation cue appeared on the left side of the display screen and, once this was fixated, a sentence was displayed with the first character replacing the cue. The participant pressed a response key once they finished reading each sentence, after which the sentence disappeared. On 25% of trials, the sentence was followed by a yes/no comprehension question to which participants responded. The experiment lasted approximately 25 minutes for each participant.

Results

Mean response accuracy for comprehension questions was 97% and did not differ across conditions (p>.05). Following convention, a data reduction procedure combined short fixations (<40ms) with nearby fixations, after which fixations <80ms or >1200ms were removed. Trials in which sentences received <4 fixations were deleted, as were trials in which scores were >3SD from each participant's mean (affecting 1.2% of data).

Two sets of analyses were performed on the remaining data. Sentence-level analyses assessed reading times and eye movements for sentences, while target word-level analyses assessed eye movements for target words. For sentence-level analyses, we report total reading time (time from the onset of a sentence display until the participant pressed a response key to indicate they had finished reading), average fixation duration (mean length of fixations), number of fixations, number of regressions (backward eye movements), and forward saccade amplitude (the mean length, in characters, of forward eye movements). For target word-level analyses, we report word-skipping (probability of not fixating a word prior to a fixation to its right), first-fixation duration (length of the first fixation on a word prior to a fixation to its right), gaze duration (sum of all fixations on a word prior to a fixation to its right or a regression from the word), and total reading time (sum of all fixations on a word). Sentence-level analyses provided a general indication of effects on readers' eye movements, while word-level analyses enabled a more focused analysis of effects for target words.

Sentence-Level Analyses

Sentence-level data were analysed using one-way repeated-measures analysis of variance (ANOVA) with the factor background speech (meaningful speech, meaningless speech, silence), and error variance computed over participants (F_1) and stimuli (F_2). Pair-wise comparisons were Bonferroni-corrected.

Table 2 shows means for sentence-level measures. There were significant main effects of background speech on reading time, $F_1(2,82)=9.15$, $p<.001$, $\eta_p^2=.18$, and $F_2(2,142)=8.09$, $p<.001$, $\eta_p^2=.10$, number of fixations, $F_1(2,82)=11.07$, $p<.001$, $\eta_p^2=.21$, and $F_2(2,142)=13.43$, $p<.001$, $\eta_p^2=.16$, and number of regressions, $F_1(2,82)=9.56$, $p<.001$, $\eta_p^2=.19$, and $F_2(2,142)=18.16$, $p<.001$, $\eta_p^2=.20$. Meaningful speech produced longer reading times and more fixations and regressions than meaningless speech or silence (ps<.05), which did not differ (ps>.05). Background speech did not significantly affect average fixation duration (Fs<1) or forward saccade amplitude, $F_1(2,82)=2.78$, $p=.07$, $\eta_p^2=.06$, and $F_2(2,142)=2.07$, $p=.13$, $\eta_p^2=.03$. The findings show that meaningful but not meaningless background speech slowed reading and disrupted eye movement behaviour.

Target Word-Level Analyses

Target word-level analyses were conducted using within-participants ANOVA with factors background speech (meaningful speech, meaningless speech, silence) and word frequency (high, low), with error variance computed over participants (F_1) and stimuli (F_2). Pair-wise comparisons were Bonferroni-corrected.

Table 3 shows means for target word-level measures. No main effects of background speech were observed in first-fixation duration, gaze duration or word-skipping (Fs<1.2). However, these were observed for total reading times, $F_1(2,82)=4.57$, $p=.01$, $\eta_p^2=.10$, and $F_2(2,142)=5.18$, $p<.01$, $\eta_p^2=.07$, which were longest for meaningful speech ($p<.001$), but did not differ for meaningless speech compared to silence ($p>.05$). This finding was consistent with sentence-level evidence that meaningful but not meaningless background speech resulted in greater re-reading.

Effects of word frequency were significant in first-fixation duration, $F_1(1,41)=6.89$, $p=.01$, $\eta_p^2=.14$; $F_2(1,71)=6.43$, $p=.01$, $\eta_p^2=.08$, gaze duration, $F_1(1,41)=15.33$, $p<.001$, $\eta_p^2=.27$, and $F_2(1,71)=15.45$, $p<.001$, $\eta_p^2=.18$, total reading time, $F_1(1,41)=35.31$, $p<.001$, $\eta_p^2=.46$, and $F_2(1,71)=25.46$, $p<.001$, $\eta_p^2=.26$, and word skipping, $F_1(1,41)=12.30$, $p<.01$, $\eta_p^2=.23$, and $F_2(1,71)=11.41$, $p<.01$, $\eta_p^2=.14$. These revealed standard word frequency effects such that low frequency words were skipped less frequently and received longer fixations than high frequency words during initial and later processing. There also was a significant interaction between background speech and word frequency in first-fixation durations, $F_1(2,82)=5.64$, $p<.01$, $\eta_p^2=.12$; $F_2(2,142)=5.62$, $p<.01$, $\eta_p^2=.07$. This was due to a word frequency effect only in silence so that similar effects were not observed for meaningful or meaningless background speech. As typical word frequency effects were observed in later fixation time measures, it appears background speech (whether meaningful or meaningless) disrupted the normal processing of words by delaying the appearance of word frequency effects in readers' eye movements.

Discussion

The present experiment revealed an ISE in Chinese reading. An effect in sentence-level measures (and total reading times for target words) was consistent with findings for alphabetic languages showing that meaningful but not meaningless background speech disrupts reading by increasing re-reading (Cauchard et al., 2012; Hyönä & Ekholm, 2016). The present findings therefore provide novel evidence that such effects are observed for a character-based language such as Chinese in which access to meaning may not be strongly mediated by phonological processes. Hyönä and Ekholm attributed the ISE in their eye movement experiment to background speech interfering with text comprehension and causing readers to re-inspect earlier text to re-activate information lost from working memory. This is compatible with an interference-by-process account (e.g., Marsh, et al., 2007, 2009) which attributes the ISE to conflict between the similar processing mechanisms of two sensory channels but not similarity in content of task-relevant material and background speech. Following this view, background speech that is meaningful to the reader will produce an ISE, since only it recruits processes involved in comprehension. By comparison, background speech in a foreign language or rendered meaningless so that it no longer contains clearly identifiable words will not recruit the same comprehension processes and so not produce an ISE.

Crucially, the present research provided additional novel evidence that the ISE can disrupt the normal lexical processing of words and so delay the appearance of the word frequency effect in readers' eye movements. Disruption was observed for both meaningless and meaningful speech, which resonates with findings showing that disruption to memory for word-lists is independent of the meaningfulness of background speech (LeCompte & Shaibe, 1997). Our findings show this disruption occurred early during processing, and so is likely to occur when encoding words in short-term memory tasks. The effect we observed was also short-lived, which may explain why the ISE did not appear to influence the processing of words in studies that used less sensitive measures (Boyle & Coltheart, 1996).

Such effects may be explained by interference due to the similarity of visual and auditory stimuli (e.g., Salamé & Baddeley, 1982). That is, speech inputs, including meaningless speech, that gain obligatory access to the phonological store in working memory might interfere with recoded written input and disrupt word identification. Alternatively, lexical processing may be disrupted by the load incurred when automatic processing of unattended speech-like input recruits modality-general sub-lexical processes (Frost, 1991; Hanson, 1981), thereby slowing the processing of written words. Presently, we cannot discriminate between these accounts. Future work should therefore determine if disruption to the processing of words depends on the similarity of meaningless speech to normal speech (perhaps using spectrally-rotated speech that retains more of the phonological characteristics of normal speech, e.g., Blesser, 1972). Moreover, it will be important to establish if such effects are unique to speech-like sounds or observed more generally in environments containing background noise.

The present finding are nevertheless important in revealing that the ISE can disrupt both the lexical processing of words and higher-order comprehension during reading. These and other findings also have the important practical implication that reading is less efficient in environments containing background speech, such as shared or open-plan offices or when studying with background conversations or radio or television broadcasts. Indeed, such findings imply that reading may be optimal in environments in which silence is observed, such as a quiet office or study area or traditional library setting.

References

Banbury, S., & Berry, D.C. (1998). The disruption of speech and office-related tasks by speech and office noise. *British Journal of Psychology*, *89*, 499-517.

Blesser, B. (1972). Speech perception under conditions of spectral transformation. 1. Phonetic characteristics. *Journal of Speech and Hearing Research*, *15*, 5-41.

Boyle, R., & Coltheart, V. (1996). Effects of irrelevant sounds on phonological coding in reading comprehension and short-term memory. *The Quarterly Journal of Experimental Psychology*, 49A, 398-416.

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLOS ONE*, *5*: e10729.

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*, 381-390.

Colle, H.A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning and Verbal Behaviour, 15*, 17-31.

Frost, R. (1991). Phonetic recoding of print and its effect on the detection of concurrent speech in amplitude-modulated noise. *Cognition*, *39*, 195-214.

Halin, N., Marsh, J.E., Haga, A., Holmgren, M., & Sörqvist, P. (2013). Effects of speech on proofreading: Can task-engagement manipulations shield against distraction? *Journal of Experimental Psychology: Applied*, *20*, 69-80.

Hanson, V.L. (1981). Processing of written and spoken words: Evidence for common coding. *Memory & Cognition*, *9*, 93-100.

Hyönä, J., & Ekholm, M. (2016). Background speech effects on sentence processing during reading: An eye movement study. *PLOS ONE*, *11*: e0152133.

Jones, D.M. (1995). The fate of the unattended stimulus: Irrelevant speech and cognition. *Applied Cognitive Psychology*, *9*, 23-38.

Jones, D.M., Madden, C. & Miles, C. (1992). Privileged access by irrelevant speech: The role of changing state. *Quarterly Journal of Experimental Psychology*, *44A*, 645-669.

Jones, D.M., Miles, C., & Page, C. (1990). Disruption of proofreading by irrelevant speech: Effects of attention, arousal or memory? *Applied Cognitive Psychology*, *4*, 89-109.

Jones, D.M., & Tremblay, S. (2000). Interference in memory by process or content? A reply to Neath (2000). *Psychonomic Bulletin & Review*, *7*, 550-558.

LeCompte, D.C., & Shaibe, D.M. (1997). On the irrelevance of phonological similarity to the irrelevant speech effect. *Quarterly Journal of Experimental Psychology*, *50A*, 100-118.

Marsh, J.E., Hughes, R.W., & Jones, D.M. (2007). Auditory distraction in semantic memory: A process-based approach. *Journal of Memory and Language*, *58*, 682-700.

Marsh, J.E., Hughes, R.W., & Jones, D.M. (2009). Interference by process, not content, determines semantic auditory distraction. *Cognition*, *110*, 23-38.

Martin, R.C., Wolgalter, M.S., & Forlano, J.G. (1988). Reading comprehension in the presence of unattended speech and music. *Journal of Memory and Language*, *27*, 382-398.

Oswald, C.J.P., Tremblay, S., & Jones, D.M. (2000). Disruption of comprehension by the meaning of irrelevant sound. *Memory*, *8*, 345-350.

Perfetti, C.A., & Zhang, S. (1991). Phonemic processes in reading Chinese words. *Journal of*

*Experimental Psychology: Learning, Memory and Cogni*tion, *17*, 633-643.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*, 1457-1506.

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

Salamé, P., & Baddeley, A.D. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behaviour*, *21*, 150-164.

Sörqvist, P., Halin, N., & Hygge, S. (2010). Individual differences in susceptibility to the effects of speech on reading comprehension. *Applied Cognitive Psychology*, *24*, 67-76.

Sörqvist, P., Ljungberg, J.K., & Ljung, R. (2010). A sub-process view of working memory capacity: Evidence from effects of speech on prose memory. *Memory*, *18*, 310-326.

Zhou, X., Marslen-Wilson, W. (2000). The relative time course of semantic and phonological activation in reading Chinese. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *26*, 1245-1265.

Table 1 Mean frequency and strokes of the target words

	Mean Word Frequency (per million)	Average Number of Strokes	
		First Character	Second Character
HF Words	186	7.61	7.13
LF Words	6	7.75	7.81

Note: HF refers to high frequency, LF refers to low frequency.

Table 2a. Mean sentence-level measures*.

	Background Sound Condition		
	Silence	Meaningful Speech	Meaningless Speech
Sentence Reading Time (ms)	5405 (268)	5830 (306)	5252 (237)
Average Fixation Duration (ms)	261 (5)	261 (5)	259 (4)
Fixation Count	18.2 (.9)	19.7 (.9)	17.8 (.7)
Number of Regressions	4.5 (.4)	5.0 (.4)	4.3 (.3)
Forward Saccade Amplitude (characters)	2.1 (.1)	2.2 (.1)	2.2 (.1)

* The Standard Error of the Mean is shown in parentheses.

Table 2b. Mean word-level measures*.

	Word Frequency	Background Sound Condition		
		Silence	Meaningful Speech	Meaningless Speech
First Fixation Duration (ms)	High	247 (6)	267 (6)	261 (6)
	Low	273 (9)	266 (8)	266 (7)
Gaze Duration (ms)	High	287 (7)	305 (8)	299 (9)
	Low	326 (9)	328 (10)	320 (9)
Total Reading Time (ms)	High	426 (20)	477 (25)	446 (16)
	Low	502 (28)	548 (31)	518 (26)
Word-skipping Rate	High	.21 (.02)	.19 (.02)	.21 (.03)
	Low	.16 (.03)	.16 (.02)	.16 (.03)

* The Standard Error of the Mean is shown in parentheses.

Figure Legend

Figure 1. Example sentence with high and low frequency words underlined (NB. words were not underlined in the experiment). The sentence translates as: “Special molds are necessary to make these circular products (or accessories) which are urgently needed for the manufacturers.”

Figure 1.

High frequency: 这批厂家急需的圆形产品需要专门的模具才能够制作。

Low frequency: 这批厂家急需的圆形配件需要专门的模具才能够制作。